Socioeconomic Status as a Determinant of Cardiovascular Disease

Alison Jane Beauchamp

RN, BHSc, MPH

Thesis submitted for the Doctor of Philosophy,
Department of Epidemiology and Preventive Medicine,
Monash University

December, 2010
Notice 1

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

Notice 2

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.
# Table of contents

Abstract .................................................................................................................. iii
Declaration .............................................................................................................. v
Acknowledgements .................................................................................................. vii
List of Figures ........................................................................................................... viii
List of Tables ............................................................................................................ ix
Publications and awards.......................................................................................... x
Conference presentations ......................................................................................... xii
Abbreviations ........................................................................................................... xi
1.0 Introduction .......................................................................................................... 1
   1.1 Background to the problem .............................................................................. 1
   1.2 Research questions and thesis aims ............................................................... 4
   1.3 Summary of research activities ..................................................................... 4
2.0 Literature review .................................................................................................. 6
   2.1 Definition of socioeconomic status ............................................................... 6
   2.2 Indicators of socioeconomic status ............................................................... 7
   2.3 Measurement of health inequalities ............................................................... 10
   2.4 Cardiovascular disease .................................................................................. 13
   2.5 The association between SES and CVD ....................................................... 14
   2.6 Explaining social inequalities in CVD .......................................................... 17
   2.7 Intervening to reduce social gradients in CVD ............................................. 31
   2.8 Research and policy to address CVD inequalities in Australia ..................... 34
   2.9 Summary ....................................................................................................... 35
3.0 Methods ................................................................................................................ 37
   3.1 Theoretical and conceptual framework ......................................................... 37
   3.2 Data sources ................................................................................................... 38
   3.3 Statistical analysis ......................................................................................... 43
   3.4 Summary of research activities ..................................................................... 44
4.0 Description of CVD mortality according to SES in the MCCS, and exploration of the role of cardiovascular risk factors .............................................................................. 46
5.0 Exploring possible pathways between SES and cardiovascular risk .............. 58
6.0 Unlocking the value of cohort studies for health inequalities research .......... 93
7.0 How can we reduce health inequalities ............................................................. 135
8.0 Discussion and conclusions .............................................................................. 146
References .............................................................................................................. 162
Appendix 1 .............................................................................................................. 176
Appendix 2 .............................................................................................................. 181
Appendix 3 .............................................................................................................. 188
ABSTRACT

The health and economic burden of cardiovascular disease (CVD) is greater than that of any other disease, and it remains the leading cause of death globally. Socioeconomic inequalities in CVD are apparent in most countries of the world, and it is well-established that people from lower socioeconomic groups have higher rates of premature death and suffer greater morbidity from the disease than do those from higher socioeconomic groups. This association occurs as a gradient that runs across the entire socioeconomic spectrum, so that the lower an individual’s social and economic position, the greater their risk of developing CVD. However, while there is an extensive body of work documenting this gradient, less is known about the mechanisms and pathways by which it occurs. A more complete understanding of these pathways is of paramount importance if we are to intervene effectively to reduce the burden of CVD among lower socioeconomic groups.

This thesis aims to explore mechanisms and pathways between socioeconomic status (SES) and CVD and to describe the extent to which health care systems and targeting risk factors for the disease might potentially modify health inequalities. First, epidemiological methods are used to describe and explore pathways using data from two prospective Australian studies, the Melbourne Collaborative Cohort Study (MCCS) and the AusDiab Study. The MCCS is a large study of 41,514 men and women, the majority of whom were aged 40-69 years at baseline (1990-1994). AusDiab is a national population-based study of 11,247 adults enrolled in 1999-2000, with 6,400 participants returning for a 5-year follow up in 2004-2005. Using education level as a SES indicator, analyses are undertaken to firstly describe the socioeconomic gradient in CVD mortality in the MCCS and then quantify how much of this gradient can be explained by traditional risk factors for the disease, particularly smoking. Second, because cigarette smoking and abnormal lipids are known to be the two most important risk factors leading to CVD, an analysis of whether smoking-related changes in lipoprotein subclasses occur independently of other lifestyle risk factors for CVD is also undertaken. Third, the existence of socioeconomic gradients in most major risk factors for CVD is well-established, and using longitudinal data from AusDiab, this thesis explores the onset and development of these gradients over a 5-year time period. Longitudinal cohort studies such as the MCCS offer the ability to follow-up individuals over time, allowing for exploration of the pathways between SES and the development of CVD, and providing evidence for cause and effect relationships. Two methods by which existing cohort...
studies can be utilised more effectively for health inequalities research are described, including a method for record linkage between the MCCS and a state-wide hospital admissions dataset, and a method for retrospective geocoding of participant baseline addresses in the MCCS. Record linkage provides a means of passive follow-up of lower SES participants who may be more likely to drop out from a longitudinal study, while geocoding allows for use of geographic information systems (GIS) technology. GIS can be used in analysis of the relationships between environmental exposures and subsequent health outcomes, and an example of such analysis is provided using data from the MCCS to describe the association between neighbourhood disadvantage at baseline and subsequent fatal CVD events occurring in the cohort. Finally, this thesis explores interventions to reduce socioeconomic inequalities both in cardiovascular risk factors and in CVD itself. Broad public health measures to reduce risk factors such as smoking may have potential to reach lower socioeconomic groups, and this is illustrated using the example of legislation to reduce the effects of second-hand cigarette smoke. Current ‘best-practice’ recommendations for CVD prevention and treatment are also examined for their potential impact upon the SES-CVD gradient.

Findings from this thesis demonstrate that strong SES gradients in fatal CVD and in major risk factors for the disease still exist in contemporary Australian adults. They provide supporting evidence that these gradients in risk factors contribute significantly towards gradients in CVD, and also reinforce that SES is a predictor both of incident CVD and of incident risk factors, and temporally precedes the onset of risk. Individual risk factors of greatest importance include smoking, which among women, may lead to CVD partially through its independent action on lipoproteins. Two methods by which existing cohort studies, clinical registries and other datasets can be adapted for use in health inequalities research are described in detail, thus increasing Australia’s research capacity in this area of vital importance. Finally, this thesis also identified that, despite the greater burden of CVD among lower SES groups, little is known about the impact of best-practice preventive and treatment interventions upon the SES-CVD gradient, or about barriers to their uptake and effectiveness. As such, these best-practice interventions themselves may be contributing to inequalities in CVD. These findings have implications for health inequalities research, health policy and clinical practice and contribute significantly towards increasing our understanding of socioeconomic inequalities in CVD.
General declaration

In accordance with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes three original papers published in peer reviewed journals and three unpublished publications. The core theme of the thesis is socioeconomic inequalities in cardiovascular disease. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the Department of Epidemiology and Preventive Medicine under the supervision of Professor Andrew Tonkin, Dr Anna Peeters, Professor Gavin Turrell and Associate Professor Rory Wolfe.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.
In the case of chapters 4-7 my contribution to the work involved the following:

<table>
<thead>
<tr>
<th>Thesis chapter</th>
<th>Publication title</th>
<th>Publication status</th>
<th>Nature and extent of candidate's contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Inequalities in cardiovascular disease mortality: the role of behavioural, physiological and social risk factors</td>
<td>Published</td>
<td>Principle author, critical literature review and development of the research questions, data preparation including coordination of adjudication of fatal CVD events, statistical analysis, interpretation of results, writing of manuscript, submission to journal, response to reviewer's comments, accepts overall responsibility for the publication.</td>
</tr>
<tr>
<td>5</td>
<td>Associations between smoking status, lifestyle and lipoprotein subclasses'</td>
<td>Published</td>
<td>Critical literature review and development of the research questions, data preparation, analysis design, statistical analysis, interpretation of results, tables, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
</tr>
<tr>
<td>5</td>
<td>Incidence of cardiovascular risk factors by education level 2000-2005: the Australian Diabetes, Obesity, and Lifestyle (AusDiab) Cohort Study</td>
<td>Submitted</td>
<td>Critical literature review and development of the research questions, data preparation, analysis design, statistical analysis, interpretation of results, tables, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
</tr>
<tr>
<td>6</td>
<td>Validation of de-identified record linkage to ascertain hospital admissions in a cohort study</td>
<td>Submitted, under review</td>
<td>Critical literature review and development of the research questions, ethics applications, collection of data on non-fatal CVD events from medical records and coordination of adjudication of these events, analysis design, statistical analysis, interpretation of results, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
</tr>
<tr>
<td>6</td>
<td>Unlocking the value of cohort studies for analysis of area disadvantage and health: a method for retrospective geocoding in the MCCS</td>
<td>Submitted, under review</td>
<td>Critical literature review and development of the research questions and study protocol, ethics applications for the project, data preparation, analysis design, statistical analysis, interpretation of results including manual review of geocoded addresses, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
</tr>
<tr>
<td>7</td>
<td>Best practice for prevention and treatment of cardiovascular disease through an equity lens: a review</td>
<td>Published</td>
<td>Development of the research questions and review protocol. Literature review and interpretation of findings, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
</tr>
</tbody>
</table>

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed: ..........................................................

Date: 20/12/10
Acknowledgements

Firstly I would like to sincerely thank my four supervisors, Professor Andrew Tonkin, Dr Anna Peeters, Professor Gavin Turrell and Associate Professor Rory Wolfe. Despite their own enormous work loads, they were overwhelmingly generous with both their time and knowledge, providing constant encouragement and guidance from start to finish. Each of them brought their own individual strengths to their role as supervisor, yet were able to work together as a united team, for which I am extremely grateful.

I would also like to acknowledge and thank my colleagues at the Department of Epidemiology and Preventive Medicine, especially Linton Harris, Karen Duve, Lei Chen, Allison Byrnes, Andrea Driscoll and Dianna Magliano for their moral support and practical assistance.

My thanks also to the team at the Cancer Council Victoria for providing me with the opportunity to work with the Melbourne Collaborative Cohort Study, in particular Professors Dallas English and Graham Giles. I would also like to acknowledge the assistance of Dr Helen Kelsall with the geocoding and linkage projects.

Finally I would like to thank all my family and friends who provided kindness, humour and encouragement. Thank you especially to John and Tom for your patience, love and understanding.
List of figures

Figure 1: Measurement of cardiovascular mortality rates by area disadvantage using absolute and relative measures................................................................. 12
Figure 2: Prevalence of selected health measures by socioeconomic status 2007-08 (percent) .................................................................................................................. 16
Figure 3: Description and figure of the conceptual framework for health determinants and how they might relate to health inequalities.......................................... 19
List of tables

Table 1 Advantages and disadvantages of commonly used indicators of socioeconomic status ................................................................. 9
Table 2 Summary of common indices used to measure the magnitude of health inequalities ........................................................................ 13
Table 3 Studies that quantify the contribution of risk factors to social inequalities in CVD ........................................................................ 23
Publications and awards

Awards and grants

National Health and Medical Research Council scholarship ID number 465352

Travel grant from Monash University Faculty of Medicine, Nursing and Health Sciences to attend a course in health inequalities in Florence, June 2008

Publications during candidature

The following papers have been published during my candidature:


Tonkin A, Beauchamp A, Stevenson C. The Importance of Extinguishing Second hand Smoke *Circulation*. 2009;120:1339-1341

The following papers have been submitted for publication during my candidature:


Conference presentations

Presenting or senior author


Contributing author


## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Australian Bureau of Statistics</td>
</tr>
<tr>
<td>AMI</td>
<td>Acute myocardial infarction</td>
</tr>
<tr>
<td>AusDiab</td>
<td>The Australian Diabetes, Obesity and Lifestyle Study</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>CABG</td>
<td>Coronary artery bypass graft</td>
</tr>
<tr>
<td>CD</td>
<td>Census collection district</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary heart disease</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>FFQ</td>
<td>Food frequency questionnaire</td>
</tr>
<tr>
<td>FPG</td>
<td>Fasting plasma glucose</td>
</tr>
<tr>
<td>GIS</td>
<td>Geographical information system</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>IDI</td>
<td>International Diabetes Institute</td>
</tr>
<tr>
<td>MCCS</td>
<td>Melbourne Collaborative Cohort Study</td>
</tr>
<tr>
<td>NDI</td>
<td>National Death Index</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>OGTT</td>
<td>Oral glucose tolerance test</td>
</tr>
<tr>
<td>PAR</td>
<td>Population attributable risk</td>
</tr>
<tr>
<td>PCI</td>
<td>Percutaneous coronary intervention</td>
</tr>
<tr>
<td>SEIFA</td>
<td>Socioeconomic index for areas</td>
</tr>
<tr>
<td>SES</td>
<td>Socioeconomic status</td>
</tr>
<tr>
<td>VAED</td>
<td>Victorian admitted episodes dataset</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

1.1 Background to the problem

The association between socioeconomic disadvantage and poorer health is one of the strongest and most consistent findings in health research. In most known societies, health inequalities exist between different socioeconomic groups, between men and women, those from different ethnic groups, and from different geographic regions. The scale of the problem is enormous, [3] and inequalities occur both within and between most countries of the world. [4-6] In the poorest 20% of countries for example, an infant is more than twice as likely to die before reaching the age of one year compared to the wealthiest 20% of countries [5], while an infant born to a family in the wealthiest quintile in Tamil Nadu is almost five times more likely to survive until age one than an infant born to a family in the poorest quintile in Uttar Pradesh. [5] There are also differences in the burden of illness, such that disadvantaged people are more likely to experience the onset of illness and disability at an earlier age. [3] For example, in Australia in 2007-2008, persons aged 15 years and over from the most disadvantaged areas had a prevalence rate for diabetes over twice that of those from the least disadvantaged areas. [7] Inequalities are seen not only in health outcomes but also in health-related behaviours and exposures [7-9] and in access to quality health care. [10-12]

Such health inequalities not only result in unnecessary human suffering; but also have an enormous economic and social impact. [6] In addition, many of these health inequalities can be considered inequitable and unfair. [13, 14] That is, while some inequalities may be due to biological variations or free choice, others may result from conditions largely outside the control of the individual. In the first case it may be impossible or unethical to change the underlying causes and so the resulting health inequalities are unavoidable. In the second, the uneven distribution may be unnecessary and avoidable as well as unjust so that the resulting inequalities also lead to inequity in health. In the United States, the term “health disparities” is most commonly used, with a strong focus on ethnicity or race. In other countries, and within this
thesis, the term “health inequalities” is more frequently used. This refers to differences in health between two or more socioeconomic groups that are considered to be potentially avoidable.[6]

Health inequalities do not just affect the lowest socioeconomic groups in a society. There is clear evidence for a health gradient that runs across the entire socioeconomic spectrum. Overall, the lower an individual’s socioeconomic position the worse their health. [15, 16] For example, the Whitehall Study of British civil servants was instrumental in demonstrating that health inequalities exist across all occupational classes. [17] This social gradient in health is observed in most countries of the world, [18] and signifies that it is not just the poorest who have worse health, but also the middle classes relative to the classes above them. This ‘social gradient’ approach recognises that interventions to reduce health inequalities should ideally involve the whole socioeconomic spectrum rather than only focusing on the health of the most disadvantaged. [19, 20] The caveat to this is that where the health gap is large, targeting interventions to the most disadvantaged should be a priority. [18]

Socioeconomic inequalities are thought to result from inequalities in the social determinants of health; that is, in the underlying social, economic and environmental circumstances that occur globally, nationally and locally, affecting all individuals. [1] These include determinants such as economic or health policy, education, employment, or housing. Our understanding of the pathways between these underlying determinants and subsequent inequalities in health is most commonly based on a framework of upstream, midstream and downstream factors. Within this framework, upstream health determinants such as housing or education are thought to impact either directly or indirectly upon midstream factors including health behaviours and psychosocial processes. Disruption to any of these midstream or upstream level factors can cause disease and premature mortality further downstream. [1, 6, 20] While our understanding of the causal pathways by which such ‘disruptions’ lead to health inequalities is becoming clearer, the mechanisms are less well understood. It is vitally important that we increase our knowledge of possible mechanisms so that we know where best to effectively intervene to reduce social gradients in health. There are three main lines of enquiry by which exploration of these mechanisms and pathways is grouped; traditional pathways research, contextual or neighbourhood research, and lifecourse research. The literature that explores traditional pathways tends to focus on the role of material and psychosocial factors including health-related behaviours, risk factor exposures, and access to preventive and health care resources. Studies of contextual pathways explore the impact of neighbourhood and other area-based effects on health, while
lifecourse research considers that it is the underlying socio-environmental determinants of health experienced at different life stages that influence the development of chronic and other diseases in later life. It is important to note that these pathways are not mutually exclusive; nevertheless they offer distinct foci for most of the current research efforts to understand and explain health inequalities.

While much is known in each of these research areas, important questions still need to be answered before we can develop interventions that are based on a full understanding of the mechanisms involved. In the field of traditional pathways research for example, key questions relate to the precise role and contribution of major risk factors in determining health inequalities, the extent to which addressing these risk factors will reduce social inequalities in health, and the temporal relationships between socioeconomic status (SES) and health inequalities. It is also unclear as to the extent by which inequalities in access to healthcare contribute to social gradients in health. Within the literature into neighbourhood effects, key areas for further research relate to the definition of neighbourhoods, whether neighbourhood effects on health can be accurately separated from individual-level factors, and whether it is place of residence that affects health, or health that affects place of residence. In the area of lifecourse research, key evidence gaps relate to accurate portrayal of early childhood SES exposures, and the role of specific physical and psychosocial environmental exposures.

Of the three main areas of research into health inequalities, this thesis explores two; neighbourhood effects on health and traditional pathways to health inequalities. Using educational attainment as a measurement of SES, the thesis makes a contribution to a number of the key questions outlined above by exploring inequalities in cardiovascular disease (CVD), the leading cause of death globally. [6] Specifically, this thesis uses data from two Australian longitudinal studies to explore the contribution and role of major risk factors, access to health care, and neighbourhood effects on social gradients in CVD.

The findings from the studies presented in this thesis have added to the international literature on traditional pathways research by reinforcing the important contribution of risk factors (in particular smoking) to social inequalities in CVD, and by demonstrating that gradients in major cardiovascular risk factors are widening. This thesis also provides evidence that current health-care interventions to prevent or treat CVD may be contributing to continuing educational
inequalities in the disease, and finally provides strategies by which existing longitudinal studies can be utilised for research into health inequalities.

1.2 Research questions and thesis aims

This thesis aims to describe and explore pathways and mechanisms by which social inequalities in CVD occur, and to describe the extent to which healthcare systems and targeting risk factors can potentially modify health inequalities by answering the following questions:

- What is the association between SES (measured by educational attainment) and CVD mortality in an Australian setting, and what proportion of this association can be explained by cardiovascular risk factors
- Will addressing ‘midstream’ factors such as cardiovascular risk factors and access to health systems have an impact upon the social gradient in CVD
- Does SES (measured by educational attainment) predict the incidence of cardiovascular risk factors

1.3 Summary of research activities

- A review of the literature (chapter 2) to investigate and describe;
  a) The measurement of SES and of health inequalities,
  b) The association between SES and CVD,
  c) Possible causes and pathways to inequalities in CVD,
  d) Current and future directions for research and policy to address inequalities in CVD.

- A description of CVD mortality according to educational attainment in a large Australian cohort study, and exploration of the role of behavioural, material and psychosocial risk factors for CVD (chapter 4).

- An exploration of possible pathways between risk factors and CVD outcomes by investigating the relationship between smoking and lipoproteins (chapter 5).

- Investigation of possible pathways by examining the role of educational attainment in predicting incident CVD risk (chapter 5).
• Investigation of methods by which existing cohort studies could be utilised to capture health outcomes through record linkage (chapter 6).

• Investigation of methods by which existing cohort studies could be utilised to capture health exposures through geocoding (chapter 6).

• Investigation of future directions for health policy and systems by reviewing best practice interventions and their potential impact upon the social gradient in CVD (chapter 7).
CHAPTER 2

Literature review

This chapter comprises 8 sections. Sections 1-3 describe commonly used indicators of socioeconomic status and the measurement of health inequalities. The fourth section provides a brief overview of cardiovascular disease (CVD) and outlines the role of risk factors in the development of the disease, while section 5 describes the associations between socioeconomic status and CVD. Section 6 focuses on causal pathways between the social determinants of health and inequalities in CVD, and reviews the literature on material, psychosocial and behavioural pathways, access to health care, and the role of neighbourhood. Section 7 examines two approaches to addressing social gradients in CVD, namely the population-based approach and the high-risk approach to reducing major risk factors for the disease. Finally, section 8 gives a brief overview of research and policy directions towards addressing inequalities in CVD in Australia.

2.1 Definition of socioeconomic status

Social stratification is the hierarchical arrangement of social classes and strata within a society. In modern Western societies, such stratification can occur in many ways, including by factors such as education, occupation, income, or wealth. Unequal distribution of these factors can lead to different levels of socioeconomic advantage or disadvantage, thus defining an individual’s socioeconomic position in a society. [18, 21] Socioeconomic status itself is a highly complex phenomenon, and there are several different definitions. [21-23] Usually it is defined as a measure of an individual or group’s economic and social position relative to others. [22]
2.2 Indicators of socioeconomic status

2.2.1 Indicators of socioeconomic status: definition and theory

Indicators of SES aim to provide information about an individual’s access to social and economic resources. [24] There are two main theoretical frameworks underpinning the measurement of SES. The first is Marxism, within which social stratification is seen as the result of a struggle for resources. [21] This approach focuses on issues of power, control, and ownership. [23] Examples of indicators that reflect this approach include Wright’s scheme of occupational classification. [21] In contrast to Marx, Weber’s theory suggests that society is hierarchically stratified along the three dimensions of class, status and power, creating groups whose members share a common position that leads to shared “life chances”. [21] This latter approach has tended to use the more objective measures of SES such as income or education. [23]

Education, income or wealth, and occupation, are the most commonly used indicators of SES in epidemiological and public health research. [25] While the dimensions of SES that these indicators capture are likely to be strongly interrelated, it is also probable that each reflects different individual and societal forces associated with health. [24] This means that in addition to defining common pathways between SES and health, each indicator will also identify specific pathways. [26] For example, income reflects spending power, housing, diet, and medical care while occupation measures prestige, responsibility, control, and work exposures. [25] A summary of the advantages and disadvantages of more commonly used indicators is shown in Table 1.

2.2.2 Indicators of socioeconomic status: individual level indicators

Education level is the most commonly used individual-level measure, possibly because it is considered one of the easiest indicators to capture. [21, 22, 25] Education has been called the most basic component of SES because of its influence on future occupational opportunities and earning potential. [27] There are several possible mechanisms through which education might influence health status. For example, education provides knowledge and life skills that allow people to gain ready access to information and health promotion resources. [27] Individuals with higher education may also be more likely to have better work and economic conditions and psychological resources, although the stability of education can sometimes mask changes in an
individual’s circumstances. Education level has been shown in many studies to have a strong association with both CVD mortality [25, 28, 29] and associated risk factors. [30, 31]

Other individual-level indicators of SES include income and wealth, which directly measure material resources. As with other indicators such as education, both have a “dose-response” association with health, [32] and can influence a wide range of material circumstances with direct or indirect implications for health, including access to food, shelter, or education, and access to health services. [27] The main disadvantage with income or wealth based indicators is that they are subject to reverse causation, whereby ill-health can limit a person’s earning capacity, or cause them to spend money on health-care. The third main group of individual-level indicators includes those that are occupation-based. These are generally used to define an individual’s access to resources, and exposure to psychological risks and physical hazards. [21] Occupation-based indicators are limited in their wider use, as they are not applicable to those not in the work-force. Employment status can also be used in research studies to compare health status or behaviors between unemployed and employed people, although these are only relevant to those eligible for employment.
Table 1  Advantages and disadvantages of commonly used indicators of socioeconomic status [22, 33, 34]

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Examples</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current income</td>
<td>Household income Individual income Multiple sources (rental properties, social security, etc)</td>
<td>Precise quantification of material resources Captures change over time Useful for working populations Straightforward interpretation</td>
<td>Age dependent More unstable measure than education or occupation Higher non-response rate than other SES measures Subject to reverse causation</td>
</tr>
<tr>
<td>Wealth</td>
<td>Housing Cars Investments Inheritance</td>
<td>More strongly linked to social class than income May be more informative than income, especially during old age Accumulates across the life-course</td>
<td>Difficult to calculate Higher error rate due to sensitivity in reporting Less clear meaning in younger populations Broader measure with many different components</td>
</tr>
<tr>
<td>Education</td>
<td>Years of education Highest educational level Credentials earned (e.g., diploma, degree)</td>
<td>Easy to measure with high response rates Excludes fewer members of the population More comparable across countries than other measures Education is fairly stable in adulthood Defined early in life (less likely to be influenced by reverse causality)</td>
<td>Varies between age cohorts and countries Quality of education not considered Economic returns may differ between ethnic groups and by gender Its stability can mask important changes in individuals’ circumstances.</td>
</tr>
<tr>
<td>Occupation</td>
<td>Employment status Specific occupational group Blue-/white-collar workers</td>
<td>Additionally provides measures of occupational exposures, working conditions, latitude in decision-making and psychological stress May be transferable to households Evolves across the life-course (so more current) Strongly related to income</td>
<td>Ranking may be unclear - often wide variation in income, education and prestige within each occupational grouping Difficult to classify unemployed, retired and home duties Subject to reverse causation Classifications vary making comparison between studies difficult</td>
</tr>
<tr>
<td>Area-based indices</td>
<td>Townsend deprivation index, SEIFA scores</td>
<td>Area-based indices are thought to capture more than just the aggregated properties of individuals living in a particular area</td>
<td>More global, but less informative than individual measures in terms of specific mechanisms Subject to the “ecological fallacy”</td>
</tr>
</tbody>
</table>

Abbreviations: SES=Socioeconomic status; SEIFA=Socioeconomic index for areas
2.2.3 Indicators of socioeconomic status: area-based measures

In addition to individual-based measures, there is also increasing interest in the use of composite area-based measures as a way of capturing additional aspects of social disadvantage. [35, 36] These area-level measures are developed through aggregation of a number of individual-level socioeconomic variables to form indexes of relative advantage or disadvantage which are then applied to a residential area. Examples of aggregated measures include the proportion of unemployed in an area, the proportion in blue collar or manual occupations, and the proportion with higher education in an area. They have been applied in many countries and their use as an important, effective and reliable research tool has been well demonstrated. [36-38]

Although use of these measures is subject to the "ecological fallacy" (that of assuming all individuals in an area have the same or similar characteristics), these indices are thought to capture more than just the aggregated properties of individuals living in a particular area; they are also considered to capture contextual (i.e. physical environment) and collective (i.e. sociocultural) properties of that area. [23, 39, 40] This is demonstrated by the fact that associations between area-level disadvantage and health outcomes persist even after controlling for individual SES, suggesting that the socioeconomic characteristics of residents do not entirely explain the relationship. [41-43] Area-based measures are being increasingly used in multilevel analysis, a statistical method for separating the effects of area and individual-level characteristics on health outcomes, thus allowing for exploration of both contextual and compositional influences. [35] In Australia, the Australian Bureau of Statistics (ABS) has developed the socioeconomic indexes for areas (SEIFA) derived from each census since 1986. SEIFA is a suite of four indexes each of which summarise a number of different socioeconomic variables within an area into a single measure. [44]

2.2.4 Indicators of socioeconomic status: choice of indicators

Because different indicators measure different phenomena and different causal mechanisms, their selection should ideally be based on the underlying hypothesis of the causal pathways between SES and the health outcome. [45, 46] The indicators chosen should also be relevant for the populations and outcomes under study. For example, current occupation may not be relevant in studies that include a large proportion of elderly subjects. In reality however, the choice of indicators may be limited by whichever are available. This is because health inequalities research
is frequently undertaken using a variety of established data sources (such as existing epidemiological studies or clinical registries) in which the collection of data on SES indicators may have been decided by the study objectives and resources available at the time. [22, 23] This is potentially an important limitation in health inequalities research as several authors have identified that indicators are not interchangeable, [33, 45, 46] suggesting that for some studies, the research question may not fit the actual indicator used. It is important that researchers are aware of this limitation and not only identify the underlying generic pathways between SES and health, but also those specific to that indicator. [23] It may be also possible to identify ways in which SES indicators can be retrospectively captured in existing studies.

2.3 Measurement of health inequalities

Measurement of health inequalities is undertaken using a variety of simple or sophisticated methods, each of which can give different assessments of the magnitude, direction and trends in health inequalities. (Table 2) The choice of which measure to use is generally dependant on the overall aims of the research, in particular whether to demonstrate absolute health inequalities, or those relative to a specified reference group. [47] Current literature suggests that both absolute and relative measures should be used, as it has been shown that relative differences in health can increase even as absolute differences decrease. [48-50] This will affect whether health inequalities are, in fact, identified. [18, 51] Figure 1 illustrates this by showing CVD mortality rates in Australia between 1985-87 and 1998-2000 according to area disadvantage. Over the 15 year period there was an overall decline in CVD mortality rates, and a subsequent decrease in the absolute risk difference between the highest and lowest SES quintiles. However, over the same time period, the rate ratio increased from 1.7 to 2.1, indicating a widening of relative inequalities.
When using relative measures of health inequalities, the choice of which reference group to use should be determined by whether the aim is to show differences between two extreme groups (such as the lowest and the highest educated groups), to compare to the population average, or to measure against a pre-specified target rate. [47] **Table 2** describes common indices used for measuring health inequalities, and summarises their most appropriate application.

**Figure 1: Measurement of cardiovascular mortality rates by area disadvantage using absolute and relative measures**

*Source: Draper G, Turrell G, Oldenberg B. [2]*

CVD = Cardiovascular disease; RD = Risk difference; RR = Rate ratio
Table 2  Summary of common indices used to measure the magnitude of health inequalities [22, 34]

<table>
<thead>
<tr>
<th>Indices of effect (compare 2 contrasting groups)</th>
<th>Indices of total impact</th>
<th>Everyone has high SES health (equality by leveling up)</th>
<th>Everyone has health of average SES (equality by redistribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate difference (e.g. the absolute difference in health outcome between primary school and tertiary educated)</td>
<td>Population attributable risk (e.g. the total number of cases that would be avoided if everyone had tertiary education)</td>
<td>Index of dissimilarity (e.g. the total number of cases to be redistributed between groups in order to obtain the same average rate for all groups)</td>
<td></td>
</tr>
<tr>
<td>Rate ratio (as above, but the proportional health outcome difference)</td>
<td>Population attributable risk % (as above, but a proportion of all cases)</td>
<td>Index of dissimilarity % (as above, but as a proportion of all cases of death, disease etc in the total population)</td>
<td></td>
</tr>
<tr>
<td><strong>Simple measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>More sophisticated measures</strong></td>
<td>Index of disparity (e.g. the relative difference between the highest educated and all other education groups)</td>
<td>Slope index of inequality (the health difference as one moves from the bottom to the top of the education hierarchy)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relative index of inequality (as above, but the proportional health difference)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disproportionality measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GINI coefficient$^1$</td>
<td>Health concentration index$^2$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^1$The Gini coefficient is a number between zero and one which summarises the distribution of income across a population [52]

$^2$The Health concentration index is a measure of the magnitude of health inequality that is comparable across time periods and between countries or regions [53]

Abbreviations: SES=Socioeconomic status
2.4 Cardiovascular disease

2.4.1 Cardiovascular disease: definitions & manifestations

Cardiovascular disease (CVD) is an umbrella term which refers to all diseases of the heart and blood vessels. It includes coronary heart disease (CHD), stroke, other vascular disease and heart failure. The two major clinical manifestations of CHD are myocardial infarction (heart attack) and angina. [54] The health and economic burden of CVD is greater than that of any other disease, and it remains the leading cause of death in both developed and developing countries. In 2005, CVD accounted for 17.5 million deaths globally, the majority from low- middle-income countries. [6] In Australia, despite declining age-adjusted CVD death rates, the disease was estimated to account for a third of all deaths in 2007 [7] and in 2004-05 cost 11% of Australia’s total allocated health system expenditure. [55] The burden of CVD is expected to continue to rise in coming decades, particularly within developing countries. This is because as these countries become more industrialised, there is a shift away from a predominance of nutritional deficiencies and infectious diseases to the more chronic diseases such as CVD, cancer and diabetes. [56]

2.4.2 Cardiovascular disease: the atherosclerotic process and the role of risk factors

The primary underlying cause of CVD is atherosclerosis, which relates to an accumulation of fatty deposits and plaque formation in arterial walls. Affected arteries progressively become occluded, thereby reducing the blood flow to the organs they supply. It is most serious when it affects the blood supply to the heart (causing angina or heart attack) or to the brain, which can lead to a stroke. [54] Plaque rupture or erosion is the antecedent to coronary thrombosis and myocardial infarction or sudden death.

Specific risk factors have been identified as increasing the risk of developing atherosclerotic CVD. [54] Those that are termed ‘non-modifiable’ include age, sex, and family history. Potentially modifiable behavioural risk factors are smoking, physical inactivity and low fruit and vegetable intake. These behavioural factors contribute to the development of the biomedical risk factors of
overweight/obesity, high blood pressure, diabetes and abnormal blood lipids. [54] Psychosocial factors are also thought to contribute to the development of CVD. [57] The INTERHEART case-control study conducted in 52 countries found that a combination of nine major risk factors accounted for approximately 90% of the population attributable risk (PAR) of acute myocardial infarction (AMI). Significant factors in the INTERHEART study included abnormal lipoproteins, smoking, abdominal obesity, hypertension, psychosocial stressors, diabetes, physical inactivity, irregular consumption of fruit and vegetables, and no alcohol intake. Worldwide, the two most important risk factors were abnormal lipids and smoking, which together accounted for about two-thirds of the PAR of AMI. [58] Cigarette smoking and abnormal lipids are likely to be strongly interrelated as many studies have demonstrated that smokers have a more atherogenic lipid profile. [59-61]

### 2.4.3 Cardiovascular disease: multiple risk factors and absolute risk

While each of the major risk factors was shown in the INTERHEART study to have an independent effect on CVD, there is also evidence that the presence of multiple risk factors can accelerate the development of atherosclerosis, [62, 63] possibly in a synergistic fashion. [54] For example, a person with mild hypertension and no other risk factors may be at a lower risk of a cardiovascular event than someone with the same blood pressure in the presence of mild or moderate increased levels of other risk factors. [54] Typically, risk factors have a continuous relationship with CVD, with the likelihood of illness tending to increase as the intensity of the risk factor increases. The probability that an individual will develop CVD over a given period of time (absolute risk) depends more on the number and intensity of risk factors than on the presence of any single risk factor in isolation. [64] Tools for predicting absolute CVD risk in clinical practice have been developed from data derived from large cohort studies, such as the widely used Framingham Risk Equation, recommended for use in Australia. [65]

### 2.5 The association between SES and CVD

#### 2.5.1 The association between SES and CVD: the scale of the problem

The relationship between SES and CVD is well established. Socially disadvantaged adults have more premature death from CVD and higher disease prevalence rates. [9, 23, 66] This inequality
is independent of the socioeconomic indicator used, persists after CVD events have occurred, and is even seen in countries with publicly funded universal health care. [67] In Australia, in 2007-08 for example, Australians living in the most disadvantaged areas had a prevalence rate of 24% for CVD, compared to 17% for those living in the least disadvantaged areas (Figure 2). [7]

Most of the evidence for the association between SES and CVD comes from high-income countries, primarily because evidence from the developing world remains limited despite these countries carrying the greater burden of CVD. In low- and middle-income countries, the association is less consistent, with not all countries showing an inverse socioeconomic gradient in cardiovascular risk factors or in CVD itself. However, as these countries have become more industrialised, they are showing signs of progressing from the rich having a worse cardiovascular risk profile, to the poor. [68] For example, in India, there has been a transition over the last decade from a positive association between social advantage and CVD to an inverse one. This shift is thought to be due to several factors including increased urbanisation and greater uptake of health protective behaviours among the higher social classes. [69] Therefore, research into the relationship between SES and CVD is of international importance.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Highest SES:5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>Lowest SES:1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health risk factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily smoking</td>
<td>11.1</td>
<td>15.1</td>
<td>18.8</td>
<td>21.1</td>
<td>28.6</td>
</tr>
<tr>
<td>Sedentary exercise level</td>
<td>24.9</td>
<td>31.0</td>
<td>38.1</td>
<td>38.8</td>
<td>45.4</td>
</tr>
<tr>
<td>Risky or high risk alcohol consumption</td>
<td>12.7</td>
<td>12.6</td>
<td>13.3</td>
<td>13.6</td>
<td>10.1</td>
</tr>
<tr>
<td>Overweight or obese</td>
<td>37.9</td>
<td>41.2</td>
<td>42.7</td>
<td>42.2</td>
<td>42.5</td>
</tr>
<tr>
<td><strong>Health condition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>3.5</td>
<td>3.3</td>
<td>3.5</td>
<td>4.0</td>
<td>4.4</td>
</tr>
<tr>
<td>Diabetes (Type 2)</td>
<td>2.0</td>
<td>3.7</td>
<td>4.1</td>
<td>4.3</td>
<td>6.2</td>
</tr>
<tr>
<td>Chronic respiratory disease</td>
<td>2.2</td>
<td>2.7</td>
<td>2.8</td>
<td>3.1</td>
<td>3.3</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>17.3</td>
<td>17.1</td>
<td>21.0</td>
<td>22.2</td>
<td>23.8</td>
</tr>
<tr>
<td>Severe/profound disability</td>
<td>2.9</td>
<td>4.1</td>
<td>4.6</td>
<td>5.3</td>
<td>6.1</td>
</tr>
</tbody>
</table>

Notes
1. Data are based on persons aged 15 years and over, except smoking and alcohol consumption (18 years and over).
2. Data are age-standardised to the 2001 Australian population.
Source: AIHW analysis of NHS 2007–08.
2.5.2 The association between SES and CVD: widening social gradients in CVD

Population level declines in age-adjusted CVD mortality have occurred in most industrialised countries during the past 3-4 decades. [70] However, this decline has not occurred equitably. In Australia for example, while there has been a reduction in the absolute gap for CVD mortality, relative inequality has increased. [66] Accordingly, measured in absolute terms, those at the bottom of the SES scale are better off than they were previously. However, when measured in relative terms, health inequities are worsening because those at the bottom of the gradient are not improving as fast as those at the top. [11, 66, 71] This suggests that not only are interventions to reduce CVD inequalities urgently needed, but also a greater understanding of the mechanisms and pathways that lead to socioeconomic gradients in CVD.

2.6 Explaining social inequalities in CVD: causal pathways and the underlying determinants of health

There is extensive evidence to show that health is driven largely by the underlying social, economic and environmental determinants that affect all individuals (Figure 3). [9, 18, 72, 73] Through various pathways and mechanisms, these upstream factors influence intermediate or midstream factors, primarily people’s health behaviours and psychosocial status. In turn, these behavioural and psychosocial factors can impact upon biomedical factors, such as body weight, blood lipids, and glucose intolerance. Within this framework, inequalities in health status between population groups can be seen as resulting from social and economic inequalities further upstream. The associations between the underlying social determinants and health inequalities can thus be thought of in terms of causal ‘pathways’. While the general direction of these pathways is from upstream to downstream, they can occur in reverse, and possibly even in a cyclical fashion. For example, an individual’s health status can influence their stress levels, their employment status and their income, in turn contributing further to poor health. [7] Generally however, evidence from longitudinal studies indicates that lower SES leads towards poor health rather than vice versa. [74, 75]

Despite the fact that the association between health and socioeconomic factors is well established, the mechanisms behind this association are less clear. Since the release of the Black report in
Britain in 1980, there has been an increasing body of literature exploring how health inequalities occur, with a considerable amount of this work focusing on CVD as a health outcome. Much of this research has focused on behavioural factors, the role of material and psychosocial pathways, and access to health care. [15, 76-79] Each of these pathways is described separately in the section below; however it is important to note that they are neither distinct from each other, nor does one necessarily ‘lead’ to another. Rather, as shown in Figure 3, the underlying determinants, mechanisms and pathways interact in a complex, cyclical and possibly cumulative fashion, leading to inequalities in health outcomes further downstream.

### 2.6.1 Material pathways

The 'material pathways' theory suggests that health inequalities can be largely explained by unequal access to factors necessary for health such as adequate income and housing, healthy food, or opportunities for recreation. [80] To illustrate, low education can affect people’s ability to make choices about healthy diet or adequate physical activity, and limited financial resources will constrain behavioural choices, and may also dictate the areas in which people live. [81] Work-related factors that are associated with adverse health outcomes include occupational hazards, shift work, unemployment and gender-based disparities. [19] Low quality housing has also been associated with poorer self-rated health, [82] and environmental conditions in poorer neighbourhoods such as air pollution or traffic noise may have detrimental health effects. [74]

Material pathways have been shown in several studies to be a major contributor to relative health inequalities. [76, 79, 83, 84] A Dutch study found that material factors (type of health insurance, financial problems, and housing tenure) contributed more to educational inequalities in mortality than did psychosocial factors (life events and external locus of control) or behavioural factors (smoking habits and physical activity). Part of the contribution of material factors was via psychosocial factors and part via behavioural factors. [79]
Figure 3: Description and figure of the conceptual framework for health determinants and how they might relate to health inequalities

Adapted from Turrell et al[1]
2.6.2 Psychosocial pathways

The ‘psychosocial pathways’ model proposes that factors related to living with economic disadvantage and low social status result in stress, depression and anxiety. This can directly affect health by increasing neuroendocrine activity, producing both short-term (increased heart rate) and long term effects (obesity, increased cholesterol). [85, 86] Psychosocial factors are also thought to affect health indirectly through their influence on health behaviours. [86, 87] For example, depression may lead to “riskier” behaviours such as increased smoking or alcohol consumption. There is also some evidence that social participation and support are cardioprotective, with social networks reducing the risk of CHD in both men and women. [57, 88, 89] There are a number of postulated mechanisms for this association. Social networks could offer socioeconomic benefits; they may influence disease progression through their effect on psychosocial well-being; and larger social networks could moderate disease risk by reducing certain negative health behaviours such as smoking and physical inactivity. [86, 88, 90, 91]

The field of psychosocial pathways research has been contentious however, and other studies have found less evidence for the association between psychosocial factors and CVD. [92-94] For example, a relatively small study of 308 post-menopausal women found that the inverse association between educational attainment and sub-clinical coronary atherosclerosis was not mediated by psychosocial factors such as anxiety, depression or social support, [94] and a much larger study of 49 259 Swedish women found no association between stroke and job strain or social support. [95] A recent review found that evidence for the association between psychosocial factors and health outcomes remains limited, making it difficult to explore reasons for the different findings between studies. [96] The author comments that there is still much to be learnt about the causal role of psychosocial issues in health inequalities, in particular relating to the direct effect of psychosocial factors on health outcomes. [96]

2.6.3 Behavioural factors

As described in section 2.3.2, common behavioural risk factors for CVD are smoking, physical inactivity and low fruit and vegetable intake. It is well established that these behavioural risk factors are patterned by SES. For example, in many developed countries, smoking is more common among those with less education and income. [97-99] Exposure to second-hand cigarette smoke is also greater for lower-income workers and families of disadvantaged smokers.
SES gradients also exist in other health behaviours. A systematic review of twenty-eight cross-sectional and five longitudinal studies found consistent evidence of increased levels of leisure-time physical activity in people from higher compared to lower socioeconomic groups. Other studies have found inverse associations between SES and fruit and vegetable intake in Europe, Australia, and the United States. These behavioural factors contribute to the development of biomedical risk factors including obesity, hypertension, diabetes and abnormal blood lipids. Social patterning in biomedical risk factors is also seen. For example, diabetes incidence and prevalence has been inversely associated with occupation, education, income and neighbourhood. SES has also been shown in several systematic reviews to be a predictor of obesity. In addition, inverse social gradients have been seen for hypertension and total cholesterol.

Because it is known that these behavioural and biomedical risk factors lead to AMI, the social gradients in their prevalence and number are considered to be a major factor in the subsequent development of social gradients in CVD. This is particularly the case for smoking. Several studies have shown that differential rates of smoking by SES make a major and direct contribution to health inequalities. One author reported that up to 50% of the mortality differences between the lowest and highest socioeconomic strata among men from 4 countries could be attributed to smoking.

However, there are two main gaps in our knowledge about the role of risk factors in CVD inequalities. First, while social gradients in major risk factors such as smoking are recognised as key drivers of the SES-CVD gradient, our understanding of how or why these risk factor gradients occur is not clearly delineated. Postulated mechanisms include psychological characteristics associated with socioeconomic disadvantage, such as an external locus of control or a tendency to think about the present more than the future, stronger beliefs in the influence of chance on health and lower perception of health risks. It is also suggested that there may be cultural norms, beliefs or attitudes specific to certain socioeconomic classes that influence health behaviours such as smoking or diet. Finally, socioeconomic differences in health literacy, knowledge and skills may contribute to differential behaviours. Despite these theories however, little is actually known about how socioeconomic factors “cause” social gradients in risk factors, and whether in fact SES itself can predict the onset of such gradients.
Secondly, while the major behavioural and physiological risk factors are known to explain most of the risk of AMI at a population level, [58] they do not appear to entirely explain or account for social gradients in CVD. To illustrate, several studies have found that the SES-CVD gradient persists even after adjustment for traditional risk factors, with only 15-55% of the relative differential seen between CVD outcomes in high vs. lower SES groups explained by these risk factors. [23, 78, 120-122] This has led to suggestions that interventions focusing on traditional risk factors are unlikely to impact greatly upon health inequalities, because these risk factors do not explain the gradient. The idea that traditional risk factors provide only a limited explanation of health inequalities has not only influenced health policy and funding; it has also led to a large body of research seeking other mechanisms to explain health inequalities including neighbourhood or area effects. A summary of studies that have examined the contribution of risk factors to inequalities in CVD is shown in Table 3.

However, others argue that the apparent contradiction between the fact that the same risk factors that explain most of population level CVD do not explain the social gradient is due to the method of measuring the differential. [15, 123] This has traditionally been done using relative measures of inequality such as relative risk, and adjusting for traditional risk factors using this method shows only a small reduction in excess risk. Relative measures are only one aspect of inequality however, and as described in section 2.3, absolute measures should also be used to show the whole picture. [66]

To summarise, while their precise explanatory role is not defined or quantified it is likely that cardiovascular risk factors are central mediators in the SES-CVD gradient, and operate through both material and psychosocial pathways in a complex and synergistic fashion. For example, material inequities may result in psychosocial factors which in themselves can damage health, and both material and psychosocial pathways are likely to impact upon health behaviours. Further research is therefore required in different populations using both absolute and relative measures in order to fully investigate whether directing efforts towards reducing risk factors at a population level will reduce the absolute burden of socioeconomic inequality in CVD mortality, and the effect this will have on relative inequality. [15]
<table>
<thead>
<tr>
<th>Author &amp; Year</th>
<th>Aims</th>
<th>Country &amp; year. (N=)</th>
<th>Study design. Indicator</th>
<th>Outcome &amp; ascertainment</th>
<th>Risk factors</th>
<th>Results</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marmot 2008 [77]</td>
<td>To determine the degree to which metabolic syndrome, inflammation and health behaviours account for the social gradient in CHD.</td>
<td>5312 men, aged 39 to 63 years at baseline 1985-1988</td>
<td>Whitehall II study Employment grade</td>
<td>201 CHD events 13 yr follow up Record linkage and medical records</td>
<td>Physical exam Smoking, activity, diet, alcohol, waist, BP, glucose, CRP, Fibrinogen</td>
<td>Hazard ratio (HR) 2.2 for CHD events Adjusting for behavioural factors reduced HR by 30% (Smoking by 19%), metabolic factors by 42%.</td>
<td>Socioeconomic differences in components of the metabolic syndrome may provide the answer to how one’s SEP can lead to increased risk of CHD.</td>
</tr>
<tr>
<td>Khang 2008 [76]</td>
<td>To examine whether major CV risk factors explain relative and absolute inequalities in mortality from CVD</td>
<td>South Korea, data from 1994 health insurance database 575,377 male public servants. Cohort study Health insurance</td>
<td>CVD deaths Record linkage. 9 year follow up</td>
<td>Collected by physical exam. BP, glucose &amp; cholesterol and smoking status.</td>
<td>HR 1.38 (1.22 to 1.55) For all 4 RF, 28% HR reduction. Most effect from smoking (19% reduction) and BP (5%)</td>
<td>Policy efforts to eliminate major CVD risk factors may have a significant effect on reducing the absolute burden of inequalities in mortality.</td>
<td></td>
</tr>
<tr>
<td>McFadden 2008 [26]</td>
<td>To investigate the associations between SES and CVD mortality, and the extent to which smoking and BMI explain this.</td>
<td>England Baseline 1993-1997 22,486 men &amp; women aged 39-79 Education &amp; occupation</td>
<td>All-cause and CVD mortality 15 yr follow up using death certificates</td>
<td>Self-report Smoking &amp; BMI only</td>
<td>HR for CVD mortality 2.88 (1.62, 5.11) when education &amp; income combined. No independent effects of education. Risk factors not grouped</td>
<td>Further understanding of the mechanisms underlying the association of each socioeconomic indicator with specific health outcomes is needed</td>
<td></td>
</tr>
<tr>
<td>Kivimaki 2007 [124]</td>
<td>To examine associations between SEP and behavioural risk factors and their effect on relative &amp; absolute SES gradients in CHD</td>
<td>9337 men and 39255 women (public sector employees) aged 17-65 in Finland 2000-2002</td>
<td>Cross sectional design Income, education, occupation</td>
<td>CHD prevalence by self-report</td>
<td>Self-report Smoking, alcohol, inactivity, obesity</td>
<td>OR 2.1-2.2 times higher by income Adjustment for RF reduced OR by 13% in women, 29% in men. Absolute difference not attributed to the risk factors</td>
<td>Interventions to reduce behavioural risk factors may not completely remove SES difference in relative or absolute CHD risk, although they would lessen these effects</td>
</tr>
<tr>
<td>Author &amp; Year</td>
<td>Aims</td>
<td>Country &amp; year. (N=)</td>
<td>Study design. Indicator</td>
<td>Outcome &amp; ascertainment</td>
<td>Risk factors</td>
<td>Results</td>
<td>Conclusion</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>--------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Laaksonen 2007 [119]</td>
<td>To examine the extent to which health behaviours can account for relative educational differences in CVD mortality</td>
<td>Finland. Data from annual surveys 1979-2001 29,065 men &amp; 31,543 women</td>
<td>Prospective cohort study</td>
<td>CVD deaths Mean follow up time 11.9 years. Ascertainment through linkage</td>
<td>Self-reported Smoking, alcohol, inactivity, fat &amp; vegetable intake, BMI</td>
<td>HR 1.46 in men, 2.16 in women 54% reduction for all risk factors in men, 22% reduction in women</td>
<td>Smoking, low vegetable use and inactivity explained a substantial part of educational level differences in CVD and all-cause mortality</td>
</tr>
<tr>
<td>Kuper 2006 [93]</td>
<td>Is there a social gradient in CHD in Swedish women and can it be explained by established coronary or psychosocial factors.</td>
<td>Sweden Baseline 1991-1992 48,066 women aged 30-50</td>
<td>Prospective cohort study Education, Job strain</td>
<td>Incident fatal &amp; non fatal CHD events. Linkage follow up 11 yr</td>
<td>Self-reported Smoking, alcohol, exercise, BMI, diabetes, HPT</td>
<td>HR 3.3 all CHD events 61% reduction from combined risk factors Smoking &amp; BMI 35% each, 13% alcohol</td>
<td>Most of the educational gradient explained by coronary risk factors but not by psychosocial variables, or health-seeking behaviour. From a public health perspective, these data support the need for CVD risk reduction programs and policies that inherently incorporate social and environmental components.</td>
</tr>
<tr>
<td>Albert 2006 [125]</td>
<td>To evaluate the relationship between traditional and novel CVD risk factors, SES and incident CVD events among women.</td>
<td>N. America Year not stated 22,688 women.</td>
<td>RCT looking at effects of Vitamin E &amp; Aspirin on CVD outcomes Education, income</td>
<td>Fatal &amp; non fatal CVD events Record review and adjudication 10 year follow up</td>
<td>Smoking, diabetes, HDL, LDL, TG, HRT, HPT, alcohol, activity, BMI, family history, CRP</td>
<td>HR 0.5 with lowest educated as reference group 49% of relationship between education &amp; CVD explained by RF, but nearly all of income/CVD relationship explained</td>
<td>From a public health perspective, these data support the need for CVD risk reduction programs and policies that inherently incorporate social and environmental components.</td>
</tr>
<tr>
<td>Harald 2006 [120]</td>
<td>To investigate the association of SES, with CHD morbidity and mortality, and to analyse to which extent modifiable risk factors explain SES differences in CHD</td>
<td>Finland. Surveys were 1982, 1987, 1992 &amp; 1997 men, 10,211 women Population based survey</td>
<td>Education, occupation, income</td>
<td>CHD mortality through registry linkage. Non fatal events, not stated 5 year follow up</td>
<td>Questionnaire &amp; physical exam Smoking, alcohol, inactivity. BMI, TC, HPT</td>
<td>HR 2.00 in male manual workers for CHD mortality 69% reduction for behavioural RF, 87% for behavioural &amp; physiological. Men: Most effect from smoking. No gradient for women</td>
<td>There is still room to reduce SES differences in CHD incidence by reducing the differences in risk factor levels. The single most important factor in reducing the excess CHD risk among males is the reduction of smoking.</td>
</tr>
<tr>
<td>Author &amp; Year</td>
<td>Aims</td>
<td>Country &amp; year. (N=)</td>
<td>Study design. Indicator</td>
<td>Outcome &amp; ascertainment</td>
<td>Risk factors</td>
<td>Results</td>
<td>Conclusion</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
<td>----------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Emberson 2004 [126]</td>
<td>To estimate the amount of SES differences in CHD that can be explained by established risk factors</td>
<td>Britain, 1978 5628 men aged 40-59 Prospective observational Occupation</td>
<td>CHD events and all-cause deaths 20 yr follow up Registry linkage and records</td>
<td>Physical exam, questionnaire Cholesterol, BP, BMI, smoking, alcohol, activity,</td>
<td>HR 1.50 for non fatal CHD. Reduction in all cause mortality after adjustment for all = 55%, Smoking 24%, BP 15%,</td>
<td>Population-wide strategies to reduce CHD risk factors are likely to have greater benefits than strategies targeting inequalities in CHD.</td>
<td></td>
</tr>
<tr>
<td>Strand 2004 [78]</td>
<td>To investigate the degree to which risk factors explain the association between education &amp; CVD mortality.</td>
<td>Norway, 1974-1978 22,712 men &amp; 21,972 women aged 35-49 Prospective health survey Education</td>
<td>Fatal IHD events Ascertained through linkage 24 yr follow up</td>
<td>Self-report and physical exam. Smoking, BP inactivity, BMI, marital status</td>
<td>HR 1.33 in men, 1.72 in women 91% reduction for men and 67% for women</td>
<td>Most of excess IHD mortality in lower SES classes is mediated through known CVD risk factors. In women smoking, BP &amp; cholesterol main contributors</td>
<td></td>
</tr>
<tr>
<td>Woodward 2003 [127]</td>
<td>To explain the CHD incidence and all-cause mortality differentials in housing tenure through proven cardiovascular risk factors.</td>
<td>Scotland Baseline 1984-1987 5735 men &amp; 5850 women aged 40-59 Housing tenure</td>
<td>All-cause mortality &amp; CHD events Mean 7.7 years Record linkage</td>
<td>Questionnaire, physical exam Smoking, BP, alcohol, BMI, activity, TC, diabetes, fibrinogen, personality</td>
<td>HR 1.48 in men and 2.64 in women. Approx 75% of HR for CHD events explained by all 14 RF in men &amp; women, Classical RF explained 42% of CHD events in men, 40% in women.</td>
<td>Smoking explains around 40% of the social dimension of CHD and death in this study; smoking avoidance should go a long way toward removing the social differential in health among adults.</td>
<td></td>
</tr>
<tr>
<td>Van Lenthe 2002 [128]</td>
<td>To quantify the contribution of material and behavioural factors to educational differences in the incidence of AMI, taking into account interrelationships.</td>
<td>The Netherlands 1991 at baseline 4853 men &amp; 5289 women Prospective cohort Education</td>
<td>Incident AMI (fatal or non-fatal) 7 year follow up record linkage of hospital admission data</td>
<td>Self-reported Alcohol, smoking activity, BMI employment, financial</td>
<td>HR 1.85 for AMI Decreased by 60% after adjustment for both behavioural factors Adjusting for both behavioural &amp; material reduced HR by 100%</td>
<td>Material factors contribute more to educational differences in AMI than behavioural factors.</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: HR=Hazard ratio; CHD=Coronary heart disease; BP=Blood pressure; OR=Odds ratio; CVD=Cardiovascular disease; BMI=Body mass index; RF=Risk factors; SES=Socioeconomic status
2.6.4 Access to health care

Inequity in access to appropriate health care has long been considered one possible material pathways explanation for socioeconomic disparities in CVD mortality, although evidence is inconsistent. For example, many studies show that people with lower SES have lower rates of coronary revascularisation procedures following AMI. [129-132] These inequities in angiography and revascularisation procedures are pervasive and appear to hold whether individual or area-based indicators of SES are used. [133, 134] The same effect has been seen in population (community) based studies. A Finnish study [135] examined socioeconomic differences in access to invasive coronary procedures among two cohorts of CHD patients; those with hospitalisations due to their first MI and those with non-acute onset of angina pectoris in the community. In general, angina patients and persons of lower SES received fewer procedures during the one-year follow-up than MI patients and those of higher SES. Waiting times for coronary artery bypass graft (CABG) surgery have also been found to be inversely associated with deprivation, with patients from the lower deprivation categories who were on a waiting list for surgery in Scotland, waiting an average of 3 weeks longer than those in the highest category. [136] However, not all studies have shown inequalities in access. For example, Britton [137], using the Whitehall II cohort examined access to procedures including angiography and revascularisation. The results showed that whilst the commonly found SES-CHD gradient was present there was no difference in the use of angiography, percutaneous coronary intervention PCI) or CABG, even after adjusting for clinical need (likely to be greater for lower SES groups).

The impact of inequalities in revascularisation procedures upon inequalities in CVD outcomes is unclear. Several studies have shown that disparities in evidence-based treatments account for most of the socioeconomic gradient in CVD outcomes following an acute cardiac event such as AMI. [138-140] However, others have shown that these same treatments explain less of the social gradient than do clinical status and CVD risk factor profile on admission. [67, 141] One Canadian study conducted in a population hospitalised for AMI within a universal health care system found that although income was strongly and inversely correlated with 2-year mortality rate, this effect was greatly attenuated after adjustment for age and pre-existing cardiovascular events or risk factors. [67] Adjustment for in-hospital process factors (including revascularisation) affected the income–mortality gradient only minimally, whilst adjustment for age, sex, ethnicity,
social support, and cardiovascular risk profile together accounted for 68% of the differences in mortality rates after acute MI between high- and low-income tertiles.

The causes of inequity in revascularisation remain unclear. Factors that may influence this inequitable access include inability to pay for procedures in privatised systems. However, one study from Australia, which has universal health care, [142] found that residents of high SES areas were more likely to undergo angiography and PCI than those from more disadvantaged areas (adjusted rate ratio for angiograms for high versus low SES: 1.74, 95% confidence interval (CI) 1.37, 2.21). Geographic location may be another reason for inequality in access to revascularisation, although one author found that while more affluent neighbourhoods tended to have a greater concentration of specialised services, inequitable distribution of hospital resources did not entirely account for the effects of SES on access to procedures and on outcome after AMI. [131]

The evidence for socioeconomic disparities in prescribing patterns for statin medication is also conflicting. While some European studies found no association between social deprivation and prescribing or use of statins, [137, 143-145] others have shown that socioeconomic inequity is present. These latter studies have demonstrated that poor housing tenure, [146] education [147-149] and income [150] are associated with lower rates of prescribing, self-reported use or dispensing rates. This disparity has been seen in countries with universal health systems, but has also been seen within subsidised health programs. One study that examined the use of statins in a Medicaid population in the U.S, found that even within a system designed to provide equitable access, younger ethnic minority groups were less likely to receive statins than their white counterparts. [151]

Conversely, other studies have found that statin prescribing is higher in disadvantaged areas. Recent Australian data has shown that in 2005, people in lower socioeconomic areas were dispensed cholesterol lowering agents at a higher rate than those in the least disadvantaged area, although at 6 months, those from lower SES areas were taking significantly fewer statins than their more disadvantaged counterparts. [152] In the UK, one cross-sectional study found that higher volumes of statins were prescribed in general practices that served more deprived populations. [153] While these higher rates of prescribing are consistent with the greater prevalence of CVD seen in more disadvantaged groups, it is unclear as to how well the increased
To summarise, access to measures such as coronary revascularisation and statins is shown to be inequitable in some studies. However, the impact this has on social gradients in CVD is largely unclear. It is also been shown that lower SES patients are less likely to visit their general practitioner for preventive reasons. Therefore, further evidence for the role of health care access in determining social gradients in CVD is required, in particular to explore the cumulative effects of inequitable access at all stages of the CVD continuum ranging from prevention, to those with increased absolute risk and those with manifest disease and heart failure.

2.6.5 Mechanisms and pathways to social inequalities in CVD: the role of neighbourhood

While the mechanisms and causal pathways described above focus on individual-level processes, these processes are generally shaped by the environmental context in which individuals live. Environments can affect people through a variety of factors including social interactions and norms that govern health behaviours, and availability of resources for healthy living. Because place of residence is strongly patterned by social position, neighbourhood characteristics may therefore be important contributors to health inequalities. Two reviews, one of 86 studies, have shown that living in a disadvantaged area is associated with both a higher prevalence and incidence of disease, including CVD. The majority of studies reviewed were cross-sectional in design, but longitudinal studies have also demonstrated an area effect on CVD incidence. For example, a study of incident AMI in Rome, found that living in the most deprived areas (census blocks) was associated with an 80% increased risk of incident AMI among women and a 40% increased risk for men.

Recent interest in area-based effects on health is partly driven by the idea that individual-based factors alone cannot explain health inequalities. This area of research fits in with the broader determinants of health framework illustrated in figure 3, in which factors such as housing or planning policy are seen as underlying determinants of health. The mechanisms between place of residence and health are not clearly delineated, but are thought to fall into two broad themes. Firstly, mechanisms may be related to the physical aspects of living in a disadvantaged area, such
as environmental exposures, or limited access to resources including public transport, healthy food, and public spaces. One review found that walking as transport and walking for recreation are both associated with features of the environment such as land-use mix, pedestrian infrastructure and population density. [162] Availability of public parks has also been associated with physical activity, [163] and obesity is shown to be associated with the built environment, [164] while access to supermarkets is associated with a healthier diet. [165] The second mechanism by which area is thought to affect health relates to the social aspects of living in a disadvantaged neighbourhood. This is less well-described, but two studies have shown that social cohesion and social capital are associated with CVD incidence or mortality. [166, 167]

There are several methodological issues with area-based research, including the use of census-based measures to define area deprivation, (such as the SEIFA measures, used in Australia and described in section 2.1.3). These measures are derived from aggregation of individual-level data, and so may simply reflect the fact that deprived people are more likely to live in deprived areas. [156] Additionally, census-based areas were defined for administrative purposes, and may not adequately capture the true neighbourhood characteristics, meaning that the causal processes are not well described. [161] Further, people with poorer health may self-select into poorer neighbourhoods, and this may lead to selection bias, or reverse causation. [156] For example having chronic heart failure may restrict a person’s ability to work, thus affecting their income and forcing them to move to a cheaper neighbourhood.

Multilevel analysis using longitudinal data offers a method by which some of the above difficulties can be addressed. Multilevel statistical techniques allow for the simultaneous investigation of effects on health at different levels of aggregation. [33] As such, area-related (contextual) effects on health outcomes can be separated from individual-level (compositional) effects. Studies that have used multilevel analysis have shown residual area effects on CVD once individual differences have been accounted for, suggesting that the social and physical environment in which people live has an additional effect on their health status. [39, 40, 43, 168] In addition, the use of geographic information systems (GIS) to define areas according to the research question rather than using pre-defined administrative boundaries means that causal processes and pathways may be better delineated. [35, 169] For example, using a GIS, a study from Brazil created neighbourhoods using both the geographic boundaries between administrative neighborhoods and natural geographic
barriers. They found that these local neighborhood units had less socioeconomic heterogeneity than the administrative neighborhoods. [169]

### 2.6.6 Mechanisms and pathways to social inequalities in CVD: lifecourse effects

Although this thesis does not directly examine lifecourse effects on social inequalities in CVD, it is an area of research that is strongly linked with both risk factors and neighbourhood effects on health, and is therefore important to include in this literature review. The lifecourse approach to understanding health inequalities considers that it is the underlying socio-environmental determinants of health experienced at different life stages that influence the development of chronic diseases. [23, 72] Several systematic reviews demonstrate that childhood socioeconomic conditions are important predictors of CVD in adulthood. [170-172] For example, a Scottish study reported that adult men whose fathers were employed in manual occupations had a relative CVD mortality risk of 1.61 (95% CI 1.39, 1.88) compared to men whose fathers were employed in non-manual occupations. [173] Another British study found that men whose family did not own a car in their childhood were at increased CHD risk even after adjustment for adult social class and behaviours (hazard ratio (HR) 1.35, 95% CI 1.04, 1.75). [48]

Three different lifecourse models are described: latency, pathway and accumulation models. [172-174] The latency model considers that early life environments affect adult health independently of subsequent experiences. Pathway models are based on the idea that adverse early life environments direct children to disadvantaged life trajectories, whereby one bad experience leads to another. Cumulative effects models consider that the adverse health effects of unfavourable environments gradually accumulate over time to increase the risk of disease through a type of dose-response relationship. It is thought that CVD mostly follows a cumulative model. [174, 175] A study conducted on the Framingham Heart Study Offspring Cohort measured childhood SES as father’s education, and adulthood SES as own education and occupation. This cumulative SES was associated with a greater risk of incident CHD over a 29-year period (HR 1.82, 95% CI 1.17, 2.85 for low vs. high cumulative score). [175]
Particular issues in the field of lifecourse epidemiology include determining at what time point covariates should be measured; and whether family variables such as single- vs. two-parent home should be controlled for. [176] The latter issue is particularly complex as family variables may either contribute to or result from decreasing SES. Other potential areas for further research include the role of both physical and psychosocial environments, and the behavioral pathways that are driven by these environments. These issues are important to resolve. If the evidence continues to show associations of childhood SES and adult health, it is vitally important to know when the risk starts, what are the causal mechanisms by which childhood SES impacts upon adult CVD outcomes, and which strategies are more likely to be effective at reducing the long-term effects of low childhood SES. [176]

2.7 Intervening to reduce social gradients in CVD

While social gradients in CVD have been well-described and a significant body of work continues to explore the causes of these gradients, less work has been undertaken on the effectiveness of policies or interventions to reduce them. [177] Two main barriers to developing and evaluating effective interventions to reduce social inequalities in CVD have been identified. First our understanding of the underlying mechanisms is limited, creating uncertainty about which approaches should be adopted. [18, 178] Second, very few preventive or treatment strategies have been designed that aim specifically to target social gradients in CVD. [178]

2.7.1 Intervening to reduce social gradients in CVD; the population-based approach

There are two main approaches to primary prevention of CVD; population-based or high-risk strategies. Assessing the potential effect of both these approaches on socioeconomic inequalities in CVD is crucial. [179]

Reducing population levels of major risk factors through ‘passive’ population approaches will inevitably decrease some of the risk factor burden associated with disadvantage. [179] Several studies have recently shown that population-wide approaches to reduce smoking, blood pressure and total cholesterol have substantially reduced CVD events and mortality by up to 50%. [179-181] These population-based approaches include policies and legislation such as those aimed at
reducing smoking or improving diet, and target whole populations rather than individuals. Such strategies are considered effective at reducing absolute inequalities in risk factors and hence in CVD. [76, 182] Using data from the Whitehall study, Kivimaki estimated the potential effects of reducing major CVD risk factors by equal amounts across all occupational groups in a male population (for example, no smoking and a 10mmHg reduction in systolic blood pressure). The study estimated that this would result in an absolute reduction in the mortality gap between the lowest and highest occupational group of 67%. [182]

Population-based approaches for prevention of CVD are thought to be more equitable than interventions focused on individuals because they enable a passive response to behaviour change. For example, policies to reduce the amount of salt in food can potentially reach all individuals regardless of their level of knowledge or material resources. [179] However, some authors suggest that the level of equity achieved may depend on whether the strategy is considered ‘superficial’ or ‘radical’. [183] Those that are superficial are based on simply encouraging people to change their behaviour, and thus are less likely to be successful among lower socioeconomic groups. [184] For example, media messages to reduce smoking are likely to be less effective among the disadvantaged because of the many social, economic and cultural factors that contribute to their higher rates of smoking and which constitute barriers to change. Thus, strategies such as media campaigns may increase relative inequalities in smoking. In contrast, a radical population strategy such as legislation against smoking in public places aims to change the context or circumstances within which behaviour occurs, rather than the behaviour itself, allowing change to occur more passively. [183] As such, radical population strategies may decrease socioeconomic gradients in CVD. An example of such legislation is shown in Appendix 1.

2.7.2 **Intervening to reduce social gradients in CVD; the high-risk approach**

High risk approaches to CVD prevention are those that aim to screen individuals to identify their risk of CVD and then to treat that risk accordingly. In Australia for example, the Medicare 45 year old health check provides a free screening and treatment advice for chronic disease for adults aged 45 and over with at least one risk factor. [185] Such strategies require people to reduce their risk factors through lifestyle modifications or compliance with medication. In theory, these strategies would be appropriate among lower SES groups because of their greater burden of risk factors for CVD. In reality, however, individually-based strategies are likely to widen health
gradients, because disadvantaged individuals are often unable to make the required lifestyle changes due to their adverse social and economic circumstances. [8] Disadvantage can occur at every stage in the screening and treatment process from attendance, to utility of assessment tools and treatment adherence. [13] For example, several studies have demonstrated that the use of widely recommended screening tools such as the Framingham risk equation may also increase CVD inequalities as these tools have been shown to under-predict risk among lower SES individuals. [186] Prescribing of preventive therapies such as blood-pressure lowering medication or statins has also been shown to be inversely associated with SES. [150, 187] Therefore, because of this inequitable response to high-risk strategies, the high-risk approach to CVD prevention may widen the SES-CVD gradient.

As described earlier, while absolute reductions in the SES-CVD gradient appear achievable through some population-based approaches, others may be likely to increase relative inequalities. If looking to decrease both relative and absolute inequalities, a combination of population and targeted approaches may therefore be the most suitable. For example, findings from the United Kingdom show that educational differences in smoking rates are decreasing. [97, 188] This is thought to be because population-based policies are supported by measures that specifically target the disadvantaged, including the provision of subsidised nicotine replacement therapy and cessation services in disadvantaged areas. [188, 189] However, as stated earlier, evidence for the effectiveness of strategies such as these to reduce social gradients in cardiovascular risk factors or in CVD remains limited. This means that while there are many CVD prevention and screening strategies in widespread use, little is known about their combined effect on social gradients in the disease. This is an important gap in our knowledge, particularly when we consider the known inequalities in secondary prevention treatment. The cumulative effect of these inequalities across the continuum of disease, from those who are well through to those with chronic heart failure, may be a significant contributory factor to the SES-CVD gradient. There may also be a role for health care services and clinicians to play by recognising the importance of SES as an independent risk factor for disease, and seeking ways of incorporating it into current best-practice management of CVD.

It is important to note that the prevention strategies described above have mostly been targeted at reducing mid-stream risk factors for CVD. Some authors suggest that this approach will be the most successful at reducing CVD mortality amongst lower socioeconomic groups. [15] However,
others argue that while possibly effective in the shorter term, this effect will not to be sustainable because interventions directed towards midstream and downstream factors do not tackle the underlying social, environmental and political determinants of health further upstream. [80, 190] Therefore the “cause of the causes” will continue. [6] Many in the field of health inequalities research believe that there is now a strong case in favor of public policies targeting the social determinants of health. [177] However, because these determinants are broad, and encompass public health, economic and social factors, there is a need to ensure that approaches to reduce health inequalities are equally broad. The social determinants approach to reducing health inequalities thus requires a focus on policies, organisations and social structure. [191] It is likely therefore that a multisectoral approach that uses a balanced mix of population and individually-based strategies and passive/active strategies will be most effective at reducing inequalities in CVD.

2.8 Current directions for research and policy to address inequalities in CVD in Australia

As illustrated throughout this chapter, Australia is similar to many other developed countries in the existence of strong socioeconomic gradients in both cardiovascular risk factors and in CVD mortality and morbidity. [70] In addition, despite Australia having a universal health care system, there is also evidence of socioeconomic inequalities in access to treatment for CVD. [12, 192] Furthermore, the gap between non-indigenous and indigenous Australians is profound, with the indigenous population living, on average, 12 years less for males and 10 years less for females than the non-indigenous population. [7]

Australia has a reasonably sound knowledge base from which to act to reduce inequalities in CVD. [193] Social Health Atlases have been published since 1990 both for Australia as a whole and for states and territories. [193] In 2006, the Australian Institute of Health and Welfare (AIHW) published a report on socioeconomic inequalities in CVD in Australia, [66] and there has been a comprehensive “Health Inequalities Monitoring Series” written at Queensland University of Technology and published by the AIHW. [2, 8, 33] The majority of these mortality and morbidity reports are based on cross-sectional analyses at an ecological level, using area-based measures of disadvantage. Data sources for these reports include the mortality registries and patient morbidity data collections. Data on risk factor prevalence at a national level has
previously been obtained from a number of cross-sectional surveys including the 1999-2000 AusDiab Study, the 1995 National Nutrition Survey and the 1980, 1983 & 1989 Risk Factor Prevalence Surveys, which all collected measured data on risk factors. There is no current national risk factor prevalence survey in Australia that uses measured data. [194]

In 1999, a review of social inequalities health research in Australia identified that there was no well organised or funded critical mass of research capacity into SES and health. Much of the research at that time was identified as being investigator-initiated and published in a non-systematic way, with few Australian studies designed particularly for health inequalities research. [1] Particular areas for which further data is required have also been identified by the AIHW, and include population-prevention activities for which there is currently little routine collection of data in Australia. [195] There is also little national data on incidence of risk factors or inequalities in hospitalisation for CVD. [66] In addition, while it may be possible to obtain comprehensive data on CVD risk factors and treatment through data linkage, there is currently no national health-related data linkage strategy. Data linkage is the bringing together of two or more different data sources that relate to the same individual, enabling researchers to look at large volumes of data from various sources. As such, it would potentially provide a rich source of data for health inequalities research. [196]

2.9 Summary

The findings of this review indicate that there are significant gaps in our knowledge about the association between SES and CVD. First, while the evidence strongly suggests that social gradients in major cardiovascular risk factors contribute to social gradients in CVD, the magnitude at which this occurs is unclear. Second, there is a need to untangle some of the mechanisms and identify the most important pathways between the social determinants of health and gradients in risk factors, and between SES and CVD itself, so that we can target interventions appropriately. This includes both individual-level processes and those occurring at the area or neighbourhood-level. Third, there is a need to describe the degree of inequalities in CVD prevention and treatment strategies, to explore the impact this has on the SES-CVD gradient, and to identify key points in the disease continuum at which to intervene. Fourth, our understanding of the gradients in risk factors themselves is limited, that is, what “causes” them in the first place, and how they might progress. Finally, while generally robust, Australian research into health
inequalities is limited by a lack of data sources and a strategic approach to identifying gaps in knowledge. Few Australian datasets have been explicitly designed for study of health inequalities, and there has been no recent national risk factor prevalence survey. There is also a need to identify ways in which to increase Australia’s research capacity in order to develop appropriate interventions to reduce the SES-CVD gradient.

A final issue that is apparent throughout this review is that of causality. The majority of studies linking SES and health have been cross-sectional, and these designs cannot rule out alternative explanations, including reverse causation. [74] Increasingly, studies are using longitudinal data to explore causal direction. [197] This thesis therefore aims to use two existing longitudinal studies to examine the association between SES and CVD mortality in an Australian setting.
CHAPTER 3

Methods

This chapter aims to provide a general overview of the methods used in this thesis. It introduces the theoretical and conceptual framework underpinning the thesis and describes in detail the two prospective cohort studies used for the analyses presented in chapters 4 to 6. The chapter also provides an overview of the statistical methods used and a summary of the major research activities undertaken by the candidate during the period of this work.

3.1 Theoretical and conceptual framework

Epidemiology is the study of the distribution and determinants of states of health in populations. [198] This thesis is based on the study of social epidemiology, which brings together the fields of epidemiology and social science. Social epidemiology has been defined as the study of “how a society’s innumerable social arrangements, past and present, yield differential exposures and thus differences in health outcomes among the persons which comprise the population”. [198] Broadly speaking, social epidemiology utilises theories and concepts from the social sciences as well as epidemiological study designs, methods and considerations in order to identify pathways by which societal conditions affect health. [199]

Epidemiology itself aims to both measure the prevalence and incidence of disease in the population, and identify the relationships between a specific risk factor and disease outcome. [200] Social epidemiology has frequently been undertaken using cross-sectional analysis, which investigates the association between risk factors and disease outcomes at one point in time. The main limitation with this approach, particularly for social epidemiology which is concerned with causal pathways, is that it is difficult to ascertain whether the exposure preceded the disease outcomes, or whether the disease itself influenced the exposure. Longitudinal studies are more useful for investigating the temporal relationships between risk factors detected at study baseline and outcomes that develop over time. [200]
The underlying premise for this thesis is that socioeconomic inequalities in CVD in Australia are basically unfair, economically unsound and socially unacceptable. It will be impossible to address health inequalities without first addressing the underlying social determinants of health. Nevertheless, the evidence suggests that social gradients in risk factors and health care access may be important contributors to inequalities in CVD in Australia, and that targeting these midstream factors may help reduce the SES-CVD gradient at least in the short term. This will require a fuller understanding of the causal and temporal pathways and mechanisms that lead to social gradients in risk factors. Longitudinal data sources are therefore required that include socioeconomic indicators appropriate for the hypothesis under question, and that have obtained accurate data on CVD outcomes. This thesis will therefore use social epidemiological methods to explore the relationships between SES, cardiovascular risk factors and CVD in two Australian prospective cohort studies.

### 3.2. Data sources

**3.2.1 Data sources: The Melbourne Collaborative Cohort Study (MCCS)**

**3.2.1.1 MCCS: Study sample**

The MCCS was established by the Cancer Council Victoria with the aim of improving research into cancer and other chronic diseases including CVD. Between 1990 and 1994, 41 514 participants (24 479 women) aged 27-75 years (99% aged between 40-69 years) were recruited into this study. Subjects were recruited from the Melbourne metropolitan area using a combination of methods including personal invitations, and advertising via local media outlets. Southern European migrants were deliberately over-sampled in order to increase the variability of lifestyle and genetic factors, and were targeted via ethnic radio, social clubs and churches, telephone books and the electoral roll. All subjects provided written informed consent, and the study was approved by The Cancer Council Victoria’s Human Research Ethics Committee (see Appendix 2).

Between 2003 and 2008, attempts were made to re-interview all surviving participants; 28 240 participants were re-interviewed, with measurements and data collected as at baseline including physical measurements, socio-demographic data and behavioural and lifestyle factors. [201]
3.2.1.2 MCCS: risk factor measurements

Baseline examination included face-to-face interviews and questionnaires conducted in English, Greek or Italian, depending on the subject’s preferred language. Information was collected on all participants including demographic data, social factors, personal and family medical history, smoking, alcohol intake, physical activity and reproductive history for women. Standard operating procedures were used for physical measurements. [202] Weight was measured to the nearest 100gm using electronic digital scales, height to the nearest 1mm using a wall-mounted stadiometer, and waist and hip circumferences to the nearest 1mm using a 2 metre metal anthropometric tape measure. Three consecutive blood pressure readings were taken in a seated position over 5 minute intervals using a Dinamap 1846SX automatic blood pressure monitor, and the average of the second and third recordings used in analyses. Most subjects had blood samples taken (68% of participants were fasting) for total plasma cholesterol and glucose, with randomly selected subgroups also tested for a number of other factors including triglycerides and high-density lipoprotein cholesterol. These samples were collected into 15 ml lithium heparin vacutainers and measured immediately using a Kodak Ektachem analyser (Rochester, New York). Lipoprotein subclass particle concentrations and size were measured using nuclear magnetic resonance (NMR) spectroscopy in a small number of participants.

Smoking status was ascertained from questions modified from the Medical Research Council 1986 Respiratory Symptoms Questionnaire, [203, 204] and categorised as never having smoked, currently smoking at least seven cigarettes a week for at least the past year, or formerly smoking. The latter group were defined as either having previously smoked but now ceased, or as currently smoking <7 cigarettes a week. Self-reported alcohol intake was measured using a beverage-specific quantity frequency questionnaire. At baseline, subjects were asked “Have you ever drunk at least 12 alcoholic drinks in a year (sips and tastes don’t count)?” Those answering ‘no’ were considered to be lifetime abstainers. Those answering ‘yes’ were then asked to report the usual quantity and frequency for each beverage type (beer, wine and spirits). Measures were then converted to mean grams of alcohol per day. Physical activity over the previous 6 months was based on the intensity and number of times per week that exercise for recreation or sport was undertaken. These data were then combined and categorised to give an overall score of relative energy expenditure, based on the Compendium of Physical Activities. [205] Dietary information over the preceding 12 months was collected using a self-administered food frequency
questionnaire (FFQ), specifically developed to measure the diverse diet of the cohort. [206] The FFQ assessed the average intake of 121-items using nine possible frequency responses ranging from ‘never or less than once per month’, to ‘6 or more times per day’. This information was used to calculate average daily nutrient intake, including saturated fat intake. [206] Although not validated, studies have shown fair to moderate repeatability for the FFQ over a 12 month interval, [207] and moderate to strong correlations between dietary intake and plasma levels of fatty acids. [208]

3.2.1.3 MCCS: SES indicators

In the MCCS the only measurement of SES available at study baseline was educational attainment. This was ascertained from the question ‘What is the highest level of education you completed?’ Education levels were categorised as 1) primary only (comprising no school, or primary school only) 2) some secondary (comprising some secondary education only) 3) completed secondary (comprising completed secondary education, trade certificate or some study towards a tertiary degree), and 4) completed tertiary (comprising degree, diploma or higher).

3.2.1.4 MCCS: Measurement of CVD outcomes

All fatal CVD events occurring in the MCCS between participant baseline attendance date and 31 December 2002 were identified through linkage with the Victorian Registry of Births, Deaths and Marriages and the Australian National Death Index (NDI). Participant’s medical records and autopsy reports were retrieved from hospitals, general practitioners, nursing homes and the Victorian Institute of Forensic Medicine. Where no medical record or autopsy report was available, death certificates were used. Data relating to the event was collected onto laptop computers and collated, and then provided to panels of expert cardiologists and neurologists for adjudication. Events were categorised as CVD-related or otherwise according to pre-defined criteria. Deaths were attributed to CVD if the underlying cause of death was related to CHD, stroke or ‘other’ CVD cause, comprising indeterminate fatal CVD event, non-coronary cardiac death, other vascular death or heart failure.
The candidate was responsible for finalising this collection of data on fatal CVD events, including coordination of final adjudication, data entry and cleaning, preparation of a data dictionary and writing of a final report.

3.2.2 Data sources: Australian Diabetes, Obesity and Lifestyle cohort study (AusDiab)

3.2.2.1 AusDiab: study sample

The Australian Diabetes, Obesity and Lifestyle cohort study (AusDiab) was initiated with a population-based, stratified cluster survey of 11,247 adults from rural and urban Australia. The study was established by the International Diabetes Institute (IDI) to examine the natural history of diabetes, pre-diabetes, heart disease and kidney disease. The study began in 1999 and a five-year follow-up survey was conducted between 2004 and 2005. Methods and response rates for the baseline [209] and follow up [210] studies have previously been described in detail (see Appendix 3). Briefly, men and women aged 25 years and over were recruited from 42 urban and non-urban census collection districts (CDs) using a stratified clustered sampling method. CDs are the smallest geographic unit defined at each census, containing an average of 225 dwellings. Ethics approval was obtained from IDI and Monash University. All participants consented to participate in the study.

Recruitment of participants at baseline was conducted in two phases: (i) a household survey and (ii) a full physical examination. The household survey involved a brief questionnaire on demographic information. Participants were then invited to attend a testing site for a full physical examination. From 17,129 eligible households, 20,347 individuals completed a household questionnaire, and 11,247 (55%) agreed to participate in the physical examination, giving an estimated baseline response rate of 37%.

A five-year follow-up was conducted in 2004-2005. From the original cohort there were 10,788 participants eligible for follow-up and of these, 6,400 returned for physical examination and interviewer-administered questionnaire. Comparison between attendees and non-attendees to follow-up showed that attendees were less likely to be hypertensive or to smoke, had a higher level of educational attainment, and a smaller waist circumference at baseline. [210]
3.2.2.2 **AusDiab: risk factor measurements**

Baseline and follow-up assessments followed a similar protocol, which was based on the World Health Organisation (WHO) recommended model for diabetes and other non-communicable disease field surveys. [211] Data on medical history, lifestyle and demographic factors including age, sex, family history and current use of medication were collected by interviewer-administered questionnaires. Self-reported CVD was ascertained by asking participants if they had ever been told by a doctor or nurse that they had angina, coronary heart disease, or stroke. Smoking status was defined as 1) current daily smoker and 2) ex-smoker (now less than daily for at least the last 3 months, but used to smoke daily) and non-smoker (never smoked tobacco daily) combined.

Blood pressure was measured using Dinamap or a standard mercury sphygmomanometer in a seated position after five minutes rest. [212] Three measurements were taken at one minute intervals. The final blood pressure measure was defined as the mean of the first two readings unless the difference between these readings was greater than 10 mmHg, in which case the mean of the two closest measurements was used. Height was measured to the nearest 0.5cm without shoes, and weight was measured without shoes and excess clothing to the nearest 0.1kg. [213] Body mass index (BMI) was calculated as weight (kg)/height (m)^2.

Fasting serum total cholesterol and blood glucose measurements were measured with an Olympus AU600 analyser (Olympus Optical, Tokyo, Japan) at a central laboratory. Three indices of blood glucose were measured; fasting plasma glucose (FPG), 2-hour FPG and Haemoglobin A1c. All fasting participants excluding pregnant women, those on insulin or oral hypoglycaemic agents undertook a 75g Oral Glucose Tolerance Test (OGTT) to determine FPG and 2-hour FPG. Participants were classified as having known diabetes mellitus if they reported having physician-diagnosed diabetes mellitus and were either taking hypoglycaemic medication or had FPG ≥7.0 mmol/L or 2-hour FPG ≥11.1 mmol/L. Participants not reporting diabetes mellitus but with FPG ≥ 7.0 mmol/L or 2-hour FPG ≥11.1 mmol/L were classified as newly diagnosed diabetes mellitus.
3.2.2.3 AusDiab: SES indicators

Socioeconomic indicators used in AusDiab include years of education, household income and employment status. For this thesis, education level was used as the SES indicator of interest, and was ascertained by asking the question “Which of these describes the highest qualification you have received?” Education was categorised as secondary only (comprising those with a secondary school qualification or lower), diploma (comprising nursing or teaching qualification, trade or other certificate, associate or undergraduate diploma), and degree (comprising bachelor degree, post-graduate diploma or masters degree/doctorate).

3.3 Statistical analysis

The statistical methods used in this thesis are fully explained in each of the papers in chapters 4 to 6. Descriptive methods were used to report unadjusted baseline risk factors and demographics in both the MCCS and AusDiab. Trends in baseline variables across education categories were assessed using logistic or linear regression with education as a continuous variable. Analyses in AusDiab were conducted using sample weights to account for the sampling design of the study. Logistic and linear regression were also used to explore the relationship between smoking status and lipoprotein subclasses in a cross-sectional analyses in the MCCS. Univariate and multivariate models were presented, with the latter adjusting for lifestyle-related risk factors considered potential confounders of the smoking-lipoprotein relationship. Survival analysis using Cox proportional hazards regression was used to assess the association between education and CVD mortality in the MCCS. Models were constructed to assess the relationships between education at baseline and fatal CVD events, adjusting for demographic variables. Subsequent models adjusted for behavioural, social or physiological risk factor groups and finally for all risk factors simultaneously. Univariate logistic regression was also used to analyse the incidence of risk factors for CVD in AusDiab.

In addition, multilevel analysis was used in chapter 6 to assess whether neighbourhood disadvantage was associated with fatal CVD events in the MCCS independently of the individual-level variables of age and ethnicity, and educational attainment using a three-stage modelling approach. Marko Chain Monte Carlo simulation was used to estimate the fixed and random parameters for each model; this was implemented using the Metropolis-Hastings algorithm via
MLwiN software. [214] To achieve convergence of the simulated chains for the variance parameters the Metropolis-Hastings algorithm was implemented for 500,000 iterations.

The assumptions required for each statistical test were thoroughly examined. The distribution of all continuous measures was explored, and in cases where there was a marked deviation from the normal distribution, the variable was transformed as appropriate. Analyses were performed using Stata 9.2 and 10 (Stata Corp, College Station, Texas, USA) and MLwiN 2.17. [215]

### 3.4 Summary of major research activities undertaken by the candidate during the period of this work

#### 3.4.1 Finalising the collection and adjudication of all fatal CVD outcomes in the MCCS

- Data on over 900 fatal CVD events occurring in the MCCS since baseline were retrieved from medical records at over 30 sites in previous work led by the candidate. These data were adjudicated and categorised by expert panels of cardiologists and neurologists in a process which was also coordinated and managed by the candidate.
- During candidature, the candidate’s specific role included preparation of data for adjudication, cleaning and subsequent management of the data including the development of a data dictionary and other background documents, and writing of final reports.

#### 3.4.2 Record linkage of the MCCS with the Victorian Admitted Episodes Dataset (VAED)

- The candidate took the leading role in all stages of this particular project. Record linkage is the process of using common identifiers to match two or more datasets in order to obtain health outcomes or other data. Linkage to the VAED has not been previously undertaken with such a large cohort as the MCCS, and there was not full availability of identifying variables in either dataset. Therefore a pilot linkage of 2000 MCCS participants was undertaken to determine the most suitable process and algorithms for linkage.
• The project involved an extensive consultative process between the Victorian Department of Health, Monash University and the Cancer Council Victoria (holders of the MCCS data). This process was led by the candidate, who was also responsible for the development and review of an appropriate methodology for linkage, obtaining ethics approval, and writing of final reports.

• In addition, the candidate manually reviewed over 200 medical records of MCCS participants and collected data on 120 non-fatal CVD events. These data were collated by the candidate and subsequently adjudicated and coded by cardiologists and neurologists before being used to test the sensitivity of the pilot linkage.

• Following a preliminary linkage of the pilot sample, matched records were manually reviewed by the candidate, and a final algorithm was developed that will be used to link the entire MCCS cohort with the VAED (see chapter 6).

3.4.3 Obtaining SEIFA scores for the MCCS through retrospective geocoding of all baseline addresses

• The candidate also took the leading role in this project which aimed to retrospectively geocode all MCCS addresses at study baseline (1990-1994) in order to obtain an area-based measure of disadvantage at baseline.

• The process included scoping and identifying the most appropriate method for retrospective geocoding, sourcing holders of geographical systems software for the relevant time period (1991), obtaining ethics approval, developing memorandums of understanding between all parties, overseeing the geocoding process, manual review of unmatched addresses and writing of a final report (see chapter 6).
CHAPTER 4

Description of CVD mortality according to SES in the MCCS, and exploration of the role of cardiovascular risk factors

4.1 Chapter overview

A social gradient in CVD has been well documented in Australia and elsewhere, although the majority of Australian evidence is based on ecological studies, and examines prevalence rather than incidence. [66] It is also known that there are socioeconomic inequalities in the presence of risk factors for CVD. However as described in chapter 2, the degree to which these risk factors explain the SES-CVD gradient differs between studies and methods used. Because of this uncertainty, the extent to which targeting risk factors among disadvantaged groups will reduce social inequalities in CVD is unclear. Furthermore, as illustrated in Figure 3, behavioural risk factors such as smoking and poor nutrition are thought to precede the development of physiological risk factors such as high blood pressure. Analysis of cardiovascular risk factors within this framework may help clarify pathways between SES and CVD and allow us to identify key points for intervention. Few studies have assessed groups of risk factors in this way, particularly in both men and women. Finally, the relative contribution of risk factors to the social gradient in CVD may change in line with secular trends in risk factors. For example, the prevalence of obesity has increased markedly compared with 20 or even 10 years ago. Therefore, contemporary evidence concerning the association of risk factors and CVD inequalities is required.

This chapter aims to establish whether an educational gradient in CVD mortality exists in a contemporary, low risk Australian population using both absolute and relative measures of inequality, to describe the primary mediators of the current relationship (i.e. behavioural and
biomedical risk factors, and psychosocial factors) and to quantify how much of the gradient can be explained by these risk factors. This first analysis will be used to start addressing the broader questions of: do we fully understand the implications (magnitude) of current inequalities, and are we targeting the factors that are most important for this relationship.
Monash University

Declaration for Thesis Chapter 4

Declaration by candidate

In the case of chapter 4, the nature and extent of my contribution to the work ‘Inequalities in cardiovascular disease mortality: the role of behavioural, physiological and social risk factors’ was the following:

<table>
<thead>
<tr>
<th>Nature of contribution</th>
<th>Extent of contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principle author, critical literature review and development of the research questions, data preparation including coordination of adjudication of fatal CVD events, analysis design, statistical analysis, interpretation of results, writing of manuscript, submission to journal, response to reviewer’s comments, accepts overall responsibility for the publication.</td>
<td>75</td>
</tr>
</tbody>
</table>

The following co-authors contributed to the work.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Anna Peeters</td>
<td>Facilitated development of research questions, assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>2 Rory Wolfe</td>
<td>Commented on data, assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>3 Gavin Turrell</td>
<td>Facilitated development of research questions, assisted with interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>4 Linton R Harriss</td>
<td>Commented on data, assisted with interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>5 Graham G Giles</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>6 Dallas R English</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>7 John McNeil</td>
<td>Grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>8 Diana Magliano</td>
<td>Grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>9 Stephen Harrap</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, commented on manuscript drafts</td>
</tr>
<tr>
<td>10 Danny Liew</td>
<td>Grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>11 David Hunt</td>
<td>Adjudication of CVD events, commented on manuscript drafts</td>
</tr>
<tr>
<td>12 Andrew M. Tonkin</td>
<td>Grants used to fund research, facilitated development of research questions, interpretation of results, commented manuscript drafts</td>
</tr>
</tbody>
</table>

Candidate’s Signature

Date 20/10/10
Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate’s contribution to this work, and the nature of the contribution of each of the co-authors.

(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;

(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

(4) there are no other authors of the publication according to these criteria;

(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and

(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

| Location(s) | Department of Epidemiology & Preventive Medicine, Alfred Campus, Monash University |

[Please note that the location(s) must be institutional in nature, and should be indicated here as a department, centre or institute, with specific campus identification where relevant.]

<table>
<thead>
<tr>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature 1</td>
<td>15/12/2010</td>
</tr>
<tr>
<td>Signature 2</td>
<td>13/12/2010</td>
</tr>
<tr>
<td>Signature 3</td>
<td></td>
</tr>
<tr>
<td>Signature 4</td>
<td>18/10/10</td>
</tr>
<tr>
<td>Signature 5</td>
<td>25/10/10</td>
</tr>
<tr>
<td>Signature 6</td>
<td>31/11/2010</td>
</tr>
<tr>
<td>Signature 7</td>
<td>07/12/10</td>
</tr>
<tr>
<td>Signature 8</td>
<td></td>
</tr>
<tr>
<td>Signature 9</td>
<td>04/12/10</td>
</tr>
<tr>
<td>Signature 10</td>
<td>03/11/2010</td>
</tr>
<tr>
<td>Signature 11</td>
<td>25/11/2010</td>
</tr>
<tr>
<td>Signature 12</td>
<td>08/12/2010</td>
</tr>
</tbody>
</table>
Inequalities in cardiovascular disease mortality: the role of behavioural, physiological and social risk factors

Alison Beauchamp, Anna Peeters, Rory Wolfe, Gavin Turrell, Linton R Harriss, Graham G Giles, Dallas R English, John McNeil, Dianna Magliano, Stephen Harrap, Danny Liew, David Hunt, Andrew Tonkin

ABSTRACT

Background While the relationship between socio-economic disadvantage and cardiovascular disease (CVD) is well established, the role that traditional cardiovascular risk factors play in this association remains unclear. The authors examined the association between education attainment and CVD mortality and the extent to which behavioural, social and physiological factors explained this relationship.

Methods Adults (n = 38,355) aged 40–69 years living in Melbourne, Australia were recruited in 1990–1994. Subjects with baseline CVD risk factor data ascertained through questionnaire and physical measurement were followed for an average of 9.4 years with CVD deaths verified by review of medical records and autopsy reports.

Results CVD mortality was higher for those with primary education only, compared with those who had completed tertiary education, with an HR of 1.66 (95% CI 1.10 to 2.49) after adjustment for age, country of birth and gender. Those from the lowest educated group had a more adverse cardiovascular risk factor profile compared with the highest educated group, and adjustment for these risk factors reduced the HR to 1.18 (95% CI 0.78 to 1.77). In analysis of individual risk factors, smoking and waist circumference explained most of the difference in CVD mortality between the highest and lowest education groups.

Conclusions Most of the excess CVD mortality in lower socio-economic groups can be explained by known risk factors, particularly smoking and overweight. While targeting cardiovascular risk factors should not divert efforts from addressing the underlying determinants of health inequalities, it is essential that known risk factors are addressed effectively among lower socio-economic groups.

INTRODUCTION

Socio-economically disadvantaged adults are more likely to experience premature death from cardiovascular disease (CVD) and have higher disease prevalence rates than their more advantaged counterparts. CVD is largely preventable, and the landmark INTERHEART case–control study of acute myocardial infarction in 52 countries showed that the population attributable risk for nine independent risk factors totalled approximately 90%. While it is known that these risk factors are socially patterned, their precise role in generating socio-economic inequalities in CVD is unclear. Studies which have examined the role of traditional and non-traditional risk factors as explanations for relative socio-economic differences in CVD have varied in their findings. Therefore, the extent to which targeting risk factors among disadvantaged groups will reduce social inequalities in CVD is uncertain.

Analysis of cardiovascular risk factors in the context of their relative position on the pathway between socio-economic status (SES) and CVD is important in understanding the role they play in this association. Behavioural risk factors generally preclude physiological abnormalities, and social risk factors are thought to impact on both. Identifying the relative contribution of these groups of risk factors may not only help elucidate the mechanisms underlying the SES/CVD gradient but also inform priorities in public health policy. However, few studies have assessed physiological, behavioural or social risk factor groups in the same analysis, and no study has been identified that examined the three types of risk factors concurrently in both men and women.

It is also important to improve our understanding of the current contribution of individual risk factors to the SES/CVD gradient. Although several studies have found that smoking is the most important health behaviour explaining socio-economic differences in CVD mortality, others have found that physical activity is more important. Among physiological risk factors, findings are also inconsistent, with the relative importance of blood pressure, body mass index (BMI) and serum cholesterol varying between studies. Trends in population levels of these risk factors are changing and so contemporary evidence concerning their impact upon health inequalities is required.

We aimed to examine whether a socio-economic gradient in CVD mortality exists in a large cohort of both men and women, with accurate ascertainment of both endpoints and risk factors, and to describe and quantify the separate effects of behavioural, physiological and social risk factors on the relationship between SES and CVD.

METHODS

Subjects

The Melbourne Collaborative Cohort Study (MCCS) is a prospective study of 41,514 subjects aged 27–80 years, recruited between 1990 and
1994. Twenty-four per cent of subjects were southern European migrants, deliberately oversampled to extend the range of lifestyle factors. Details of the design, recruitment and study procedures have been published elsewhere. In brief, subjects were volunteers from metropolitan Melbourne, recruited using electoral rolls, telephone books, multicultural media outlets, community centres and churches. All subjects provided written informed consent, and the study was approved by The Cancer Council Victoria’s Human Research Ethics Committee. Baseline examination included face-to-face interviews and questionnaires conducted in the subjects’ preferred language (English, Greek or Italian). Subjects outside the age range of 40–69 years were excluded from analyses (n=577), as were those who self-reported a history of CVD (n=2448), and those missing data on education (n=10) or covariates (n=529), leaving 38,555 subjects for final analysis (15,261 men and 23,294 women). A total of 62 participants were known to have left Australia before the end of follow-up of 31 December 2002. Of these, 10 had primary education only; nine had some secondary education, nine had completed secondary education, and 34 had tertiary education.

Because of deliberate oversampling of southern European born migrants, we categorised participants according to their country of birth as follows: (1) Australia/New Zealand (NZ)/northern Europe (the latter comprising UK and The Netherlands), or (2) southern Europe (comprising Italy, Greece and Malta).

Assessment of socio-economic status
Socio-economic status was measured using educational attainment and was ascertained from the question ‘What is the highest level of education you completed?’ Education levels were categorised as (1) primary only (comprising no school, or primary school only), (2) some secondary (comprising some secondary education only), (3) completed secondary (comprising completed secondary education, trade certificate or some study towards a tertiary degree) and (4) completed tertiary (comprising degree, diploma or higher). These categories are considered to represent hierarchical stages of education, each of which has important socio-economic implications.

Baseline risk factors
Behavioural risk factors
Smoking status was ascertained from questions modified from the Medical Research Council 1986 Respiratory Symptoms Questionnaire, and categorised as either currently smoking at least seven cigarettes a week for a year, or not currently smoking. Alcohol intake was ascertained by asking participants their usual quantity and frequency of alcohol for the current decade and analysed according to quintiles of g/day, assuming a dose–response relationship. Subjects were classified as drinkers if they responded ‘yes’ to the question: ‘Have you ever drunk at least 12 alcoholic drinks in a year (sips and tastes don’t count)?’ Physical activity over the previous 6 months was based on the number of times per week that exercise for recreation or sport was undertaken at a vigorous, less vigorous or walking level only. These data were combined to give an overall score of relative energy expenditure in four categories, based on the Compendium of Physical Activities. Dietary information was collected using a self-administered food frequency questionnaire (FFQ) specifically developed for use in the MCCS. The FFQ assessed the average intake of 121 items over the previous 12 months with this information used to calculate average daily nutrient intake, including saturated fat intake. The number of times per day that fruit and vegetables were eaten was also ascertained.

Social connection
Social connection was determined by asking the number of people living in the household, and modelled as living alone versus living with others. In addition, the number of hours per week participants spent in social activities outside home or work was analysed according to the following categories: 0, 1–2, 3–4, 5–9 or 10+ h.

Physiological risk factors
Self-reported history of diabetes was ascertained from the baseline questionnaire. Standard methods were used to measure height, weight, and waist and hip circumferences. Blood pressure was measured three times after supine rest for 5 min, and the average of the second and third readings used in analysis. Blood samples for total plasma cholesterol were collected into 15 ml lithium–heparin vacutainers and measured immediately using a Kodak Ektachem analyser (Rochester, New York). In all, 68% of subjects were fasting when samples were taken.

Ascertainment of deaths
CVD-related deaths in the cohort were verified through medical record review and adjudication. In brief, CVD-related deaths occurring between participant baseline attendance date and 31 December 2002 were identified through linkage with the Victorian Registry of Births, Deaths and Marriages and the Australian National Death Index. Participant’s medical records and autopsy reports were reviewed and categorised by panels of expert cardiologists and neurologists. Where no medical record or autopsy report was available, death certificates (n=30) were used.

Fatal CVD events were classified as related to coronary heart disease (CHD), stroke or ‘other’ CVD cause, comprising indeterminate fatal CVD event, non-coronary cardiac death, other vascular death and heart failure.

Statistical analysis
Analyses were performed using Stata 9.2 (Stata Corp, College Station, Texas). Baseline risk factors and demographics are reported as means or proportions. The significance of any trend in baseline variables across education categories was assessed using logistic regression (for binary variables) or linear regression (for continuous variables) with education as a continuous variable.

Survival was described using crude mortalities. Cox proportional hazards models were constructed, and HRs and their 95% CIs calculated to assess the association between education and CVD mortality, with completed tertiary education used as the reference group. For these analyses, time at risk began at recruitment to the cohort and ended at the date of death, emigration from Australia or 31 December 2002, whichever came first. These models were extended to adjust for age, sex and country of birth, with subsequent models adjusting for behavioural, social or physiological risk factor groups and finally for all risk factors simultaneously. Tests of interaction between education and sex were not significant, and analyses are presented with males and females combined. Two-sided p values are presented, with p values <0.05 regarded as significant. The
proportional hazards assumption was tested and met for all variables used in analysis.

The percentage reduction in excess risk of CVD attributable to risk factors was calculated as:

\[
100 \times \frac{(HR_{\text{adjusted for age, sex, country of birth}} - HR_{\text{adjusted for age, sex, country of birth + risk factors}})}{(HR_{\text{adjusted for age, sex, country of birth}} - 1)}
\]

This measure provides an estimate of the explanatory contribution of risk factors to educational inequalities in CVD mortality.\(^7\)

### RESULTS

#### Baseline characteristics

relevant baseline characteristics are presented in table 1. Compared with more highly educated groups, those subjects with primary education only were more likely to be female, older and born in a southern European country. Inverse gradients were seen for most risk factors across all education categories apart from fruit, saturated fat intake, living alone, current alcohol drinking and average alcohol intake, for which positive gradients were seen. There was no significant gradient for the frequency of vegetable consumption. Cardiovascular mortality and educational attainment

There were 392 adjudicated CVD deaths during a mean follow-up time of 9.4 years per person. Of these, 239 related to CHD, 77 to stroke and 76 to ‘other’ CVD events. The crude mortality of CVD for those with primary education was more than twice that of those who had completed tertiary education (table 2). There was a trend towards an inverse gradient between education and fatal CVD events, although rates for the two intermediate categories were similar, with overlapping 95% CIs. Similar patterns were seen in population subgroups according to age, sex and country of birth.

#### Adjustment for risk factor groups

After adjusting for age, sex and country of birth, CVD mortality for those with primary education only was significantly higher than for those who had completed tertiary education (HR 1.66, 95% CI 1.10 to 2.49; table 3). For those who had completed some or all secondary education, the CVD mortality was intermediate between the primary and tertiary education categories, and an inverse gradient was observed.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Completed tertiary*</th>
<th>Completed secondary†</th>
<th>Some secondary‡</th>
<th>Primary only§</th>
<th>p Value for trend¶</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country of birth, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioural risk factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegetable intake (times/day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruit intake (times/day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saturated fat intake (g/day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current drinker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol intake, current drinkers (g/day)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical activity (% inactive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social connection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No social activity per week (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiological risk factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure (mm Hg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol (mmol/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Completed tertiary degree, diploma or higher.
†Completed secondary education, trade certificate or some study towards a tertiary degree.
§Some secondary education.
¶p Value for trend among education levels (tertiary as reference) using logistic regression for binary variables and univariate linear regression for continuous variables.

---

**Table 1** Characteristics of 38,355 subjects in the Melbourne Collaborative Cohort Study at baseline (1990–1994)
Table 2  Crude cardiovascular disease mortalities in the Melbourne Collaborative Cohort Study overall and by subgroups

<table>
<thead>
<tr>
<th>Highest level of education attained</th>
<th>Total</th>
<th>Completed tertiary*</th>
<th>Completed secondary†</th>
<th>Some secondary‡</th>
<th>Primary only§</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of cases</td>
<td>Mortality**</td>
<td>No of cases</td>
<td>Mortality (95% CI)</td>
<td>No of cases</td>
<td>Mortality (95% CI)</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>------------</td>
<td>------------------</td>
<td>------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Total</td>
<td>n= 38355</td>
<td>52</td>
<td>67 (51 to 88)</td>
<td>82</td>
<td>112 (90 to 139)</td>
</tr>
<tr>
<td>Men</td>
<td>n= 15281</td>
<td>38</td>
<td>105 (76 to 144)</td>
<td>60</td>
<td>171 (133 to 221)</td>
</tr>
<tr>
<td>Women</td>
<td>n= 23094</td>
<td>14</td>
<td>34 (20 to 57)</td>
<td>22</td>
<td>58 (38 to 88)</td>
</tr>
<tr>
<td>35–54</td>
<td>n= 18934</td>
<td>13</td>
<td>25 (14 to 43)</td>
<td>15</td>
<td>39 (24 to 65)</td>
</tr>
<tr>
<td>55–69</td>
<td>n= 19421</td>
<td>39</td>
<td>156 (114 to 213)</td>
<td>67</td>
<td>193 (152 to 245)</td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia/New Zealand/Northern Europe</td>
<td>n=28835</td>
<td>50</td>
<td>67 (51 to 89)</td>
<td>68</td>
<td>109 (86 to 138)</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>n=4520</td>
<td>2</td>
<td>64 (16 to 258)</td>
<td>14</td>
<td>131 (78 to 222)</td>
</tr>
</tbody>
</table>

*Completed tertiary degree, diploma or higher.
†Completed secondary education, trade certificate or some study towards a tertiary degree.
‡Some secondary education.
§No school, or primary school only.
**Number of fatal cardiovascular disease events.
††Difference in mortalities per 100,000 person years between lowest and highest education categories.

After further adjustment for the combined behavioural risk factors, the association between education and CVD mortality was attenuated, with a HR of 1.41 (95% CI 0.94 to 2.12) for those with only primary education, a 38% reduction in excess risk of CVD mortality. Adjustment for social connection reduced the HR for primary education by 2%, while adjustment for physiological risk factors reduced it by 45% (table 3). Adjustment for all risk factors combined reduced the HR for those with primary education only to 1.18 (95% CI 0.78 to 1.77), accounting for 73% of the excess risk of CVD mortality. Adjustment for all risk factors combined also eliminated differences in CVD mortality between the two intermediate education categories compared with those with tertiary education.

Because risk-factor distribution differed by country of birth, analyses were repeated for a subgroup comprising only those born in Australia, NZ or northern Europe. We were unable to do this analysis for those born in southern Europe because there were only two deaths due to CVD among those with tertiary education. The risk associated with primary education only was higher than that seen in the total cohort with an HR adjusted for sex and age of 1.94 (95% CI 1.21 to 3.13). In this subgroup, adjustment for behavioural risk factors reduced the HR for CVD mortality comparing the highest and lowest education groups to a greater extent than adjustment for physiological risk factors (54% and 54% reduction respectively), opposite findings to those in the total cohort. Adjustment for social connection had a greater effect than for the total cohort, with a 13% reduction in HR. Simultaneous adjustment for all three risk factor groups reduced the HR for CVD mortality for the lowest compared with the highest education group to 1.18 (95% CI 0.72 to 1.92), a reduction in excess risk of 81%.

Individual risk factors
To determine the relative explanatory role of individual risk factors on the SES/CVD gradient for the total cohort, we adjusted separately for each risk factor (table 4). Adjustment for smoking reduced the HR for CVD mortality in the lowest

Table 3  Fatal cardiovascular disease outcomes: HRs from different Cox proportional hazards models for 38,355 Melbourne Collaborative Cohort Study subjects free of cardiovascular disease at baseline, with adjustment for behavioural, social and physiological risk factor combinations

<table>
<thead>
<tr>
<th>Highest level of education attained</th>
<th>Model 1: adjusted for age, sex and country of birth</th>
<th>Model 2: adjusted for age, sex, country of birth and behavioural risk factors</th>
<th>Model 3: adjusted for age, sex, country of birth and social connection</th>
<th>Model 4: adjusted for age, sex, country of birth and physiological risk factors</th>
<th>Model 5: model 2 and model 4 combined</th>
<th>Model 6: models 2, 3 and 4 combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed tertiary*</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Completed secondary†</td>
<td>1.25 (0.88 to 1.77)</td>
<td>1.14 (0.80 to 1.62)</td>
<td>1.23 (0.86 to 1.75)</td>
<td>1.15 (0.81 to 1.64)</td>
<td>1.05 (0.74 to 1.49)</td>
<td>1.05 (0.73 to 1.49)</td>
</tr>
<tr>
<td>Some secondary‡</td>
<td>1.21 (0.87 to 1.67)</td>
<td>1.06 (0.76 to 1.47)</td>
<td>1.19 (0.86 to 1.65)</td>
<td>1.07 (0.77 to 1.48)</td>
<td>0.94 (0.68 to 1.31)</td>
<td>0.95 (0.68 to 1.31)</td>
</tr>
<tr>
<td>Primary only§</td>
<td>1.66 (1.10 to 2.49)</td>
<td>1.41 (0.94 to 2.12)</td>
<td>1.65 (1.10 to 2.47)</td>
<td>1.36 (0.90 to 2.06)</td>
<td>1.17 (0.77 to 1.76)</td>
<td>1.18 (0.78 to 1.77)</td>
</tr>
<tr>
<td>Percentage reduction from model 1 in excess risk (HR comparing primary only vs tertiary)</td>
<td>38</td>
<td>2</td>
<td>45</td>
<td>74</td>
<td>73</td>
<td>73</td>
</tr>
<tr>
<td>p Value for trend across education levels</td>
<td>0.041</td>
<td>0.219</td>
<td>0.042</td>
<td>0.297</td>
<td>0.778</td>
<td>0.730</td>
</tr>
</tbody>
</table>

*Completed tertiary degree, diploma or higher.
†Completed secondary education, trade certificate or some study towards a tertiary degree.
‡Some secondary education.
§No school, or primary school only.
**Current smoking, frequency of fruit and vegetable intake, saturated fat intake, physical activity, alcohol intake for both drinkers and non-drinkers.
††Prior diabetes, systolic blood pressure, waist circumference, cholesterol and fasting status at time of blood collection.
‡‡p Value < 0.05.
In one study, the risk of excess CHD was greater among a low-income group of Finnish men than among those in other high socio-economic groups. In this study, the risk of CHD was reduced by 118% after adjustment for 23 cardiovascular risk factors. Most other studies have found that combined risk factors play a smaller explanatory role. Differences in findings may relate to variations in the underlying social distribution of risk factors, the risk factor combinations examined and the health of the populations studied.

The relative explanatory effect of physiological factors was greater when southern European migrants were included in the analysis. This may be because this group, which comprised 85% of those with primary education, typically has a high prevalence of diabetes and other adverse physiological risk factors. Despite this, southern Europeans in our study had lower CVD mortalities than those born in Australia, NZ and northern Europe. The underlying reasons for this are unclear, although the Mediterranean diet, predominant among southern Europeans, may have been cardioprotective.

Although social connection is associated with lower rates of recurrent events in those with prevalent CVD, its role in incident CVD is less clear. We found that social factors had little contribution to CVD mortality in our study. This may be because the measures used were too crude to detect an association in this particular cohort.

**Individual risk factors**

The prevalence of current smoking at baseline was 11.5% compared with 27.6% in Victoria in 1989–1990. Despite this low prevalence, smoking was still found to be the most significant behavioural factor contributing to CVD inequality. This is consistent with many other contemporary studies. As expected, it is also consistent with another study from the same population demonstrating a substantial contribution of smoking to social inequalities in total mortality among men. The socio-economic gradient in smoking has widened in many western countries, and it is essential that this important risk factor continues to be addressed among disadvantaged groups in order to reduce their burden of CVD.

Other contemporary studies have reported that overweight or obesity influences health inequality less than blood pressure. In this present study, waist circumference had a stronger effect than blood pressure and was the most important physiological risk factor both in analysis of the total cohort and in the subgroup born in Australia, NZ and northern Europe. Waist circumference may be a more sensitive anthropometric measure, as preliminary analysis of these data found that adjustment for waist-to-hip ratio or BMI (used in the majority of other studies) had a smaller effect. An additional explanation is that the importance of weight as an explanatory risk factor for CVD inequalities has increased because of the increasing prevalence of obesity, highlighting the importance of studies such as this to provide contemporary evidence.

### Table 4 Fatal cardiovascular disease outcomes: HRs from different Cox proportional hazards models for 38 355 Melbourne Collaborative Cohort Study subjects free of cardiovascular disease at baseline, with adjustment for individual risk factors

<table>
<thead>
<tr>
<th>Highest level of education attained</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed tertiary*</td>
<td>1.00</td>
</tr>
<tr>
<td>Completed secondary†</td>
<td>1.25 (0.88 to 1.77)</td>
</tr>
<tr>
<td>Completed secondary‡</td>
<td>1.21 (0.87 to 1.67)</td>
</tr>
<tr>
<td>Primary only§</td>
<td>1.66 (1.10 to 2.49)</td>
</tr>
<tr>
<td>Percentage reduction in excess risk$</td>
<td>0.041</td>
</tr>
<tr>
<td>p Value for trend across education levels</td>
<td>0.041</td>
</tr>
</tbody>
</table>

**Notes:**
- *Completed tertiary degree, diploma or higher.
- † Completed secondary education, trade certificate or some study towards a tertiary degree.
- ‡ Some secondary education.
- § No school, or primary school only.
- $ HR comparing primary only versus tertiary.

---

**DISCUSSION**

In this very large contemporary Australian cohort, we found an inverse relationship between education attainment and CVD mortality. In the total cohort, physiological risk factors contributed more to this gradient than behavioural factors, although when southern European born migrants were excluded, behavioural factors made the larger contribution. In combination, behavioural, social and physiological factors explained almost all of the difference in CVD mortality between the tertiary and primary education categories in the total cohort, with smoking and waist circumference contributing most to this.

**Behavioural, physiological and social risk-factor groups**

Our finding that traditional risk factors can explain a substantial proportion of the difference in CVD mortality between low and high socio-economic groups has been reported by a limited number of studies. In one study, the risk of excess CHD among a low-income group of Finnish men was reduced by 118% after adjustment for 23 cardiovascular risk factors. Most other studies have found that combined risk factors play a smaller explanatory role. Differences in findings may relate to variations in the underlying social distribution of risk factors, the risk factor combinations examined and the health of the populations studied.

The relative explanatory effect of physiological factors was greater when southern European migrants were included in the analysis. This may be because this group, which comprised 85% of those with primary education, typically has a high prevalence of diabetes and other adverse physiological risk factors. Despite this, southern Europeans in our study had lower CVD mortalities than those born in Australia, NZ and northern Europe. The underlying reasons for this are unclear, although the Mediterranean diet, predominant among southern Europeans, may have been cardioprotective.

Although social connection is associated with lower rates of recurrent events in those with prevalent CVD, its role in incident CVD is less clear. We found that social factors had little contribution to CVD mortality in our study. This may be because the measures used were too crude to detect an association in this particular cohort.

**Individual risk factors**

The prevalence of current smoking at baseline was 11.5% compared with 27.6% in Victoria in 1989–1990. Despite this low prevalence, smoking was still found to be the most significant behavioural factor contributing to CVD inequality. This is consistent with many other contemporary studies. As expected, it is also consistent with another study from the same population demonstrating a substantial contribution of smoking to social inequalities in total mortality among men. The socio-economic gradient in smoking has widened in many western countries, and it is essential that this important risk factor continues to be addressed among disadvantaged groups in order to reduce their burden of CVD.

Other contemporary studies have reported that overweight or obesity influences health inequality less than blood pressure. In this present study, waist circumference had a stronger effect than blood pressure and was the most important physiological risk factor both in analysis of the total cohort and in the subgroup born in Australia, NZ and northern Europe. Waist circumference may be a more sensitive anthropometric measure, as preliminary analysis of these data found that adjustment for waist-to-hip ratio or BMI (used in the majority of other studies) had a smaller effect. An additional explanation is that the importance of weight as an explanatory risk factor for CVD inequalities has increased because of the increasing prevalence of obesity, highlighting the importance of studies such as this to provide contemporary evidence.
This study explores the relationship between education and CVD mortality within a large, contemporary and ethnically diverse sample of both men and women using a wide variety of risk factors.

What this study adds

- Traditional risk factors can explain a large part of the association between socioeconomic disadvantage and CVD.
- Consistent with several other studies, our findings show that smoking is the most important risk factor in this relationship; however, our study also shows that central adiposity is the second most important factor in explaining the increased risk of CVD among lower-educated groups.
- The relative importance of risk factors also differs between those born in southern Europe, and those born in northern Europe/Australia and New Zealand, and dietary factors may contribute to this difference.

Methodological considerations and limitations

This study was based on a large sample with accurate ascertainment of cause of death and lengthy follow-up. Education was the only SES indicator used in this study. It is possible that other indicators such as occupation or income may have contributed independently to the relationship between SES and CVD in this cohort. However, education is considered to be a robust measure of SES, and our findings are consistent with studies using other indicators. Although it is unlikely that education level changed during follow-up of this older cohort, changes in social, behavioural and physiological risk factors are possible, leading to attenuation of their effects. Because risk factors were measured at one time point only, no conclusions about temporal relationships between different groups of risk factors are possible. Despite this, our findings suggest that traditional risk factors do play an underlying mechanistic role in the gradient in CVD mortality related to SES.

This study was conducted in Melbourne, Australia, and findings may not be applicable to other regions. Generalisability is also affected by the oversampling of southern European migrants. Further, the MCCS has a low standardised mortality ratio for fatal CVD (0.41 relative to mortality levels in the state of Victoria). However, the observation of an inverse gradient in this relatively healthy cohort clearly underlines the strength of the association between education attainment and CVD.

There was a significantly greater proportion of southern European born migrants in the group with primary education only. Some of this difference could be due to measurement bias. Although differential misclassification due to language barriers is unlikely (as questionnaires were in the subjects’ preferred language), educational attainment may have different socio-economic meanings within the countries represented in this cohort. Analyses that excluded southern European born migrants showed a stronger educational gradient in CVD mortality than that seen in the total cohort. This may either reflect true differences between the two groups or be due to an underestimation of effects in the total cohort because of the inclusion of southern Europeans.

Caution needs to be used when interpreting the percentage reduction in excess risk. This concept relies on an assumption that there is no confounding of the relationship between the intermediary variable and the outcome. This is difficult to assess from these data. This measure should therefore be considered as an approximate indication of relative importance only. In addition, the fact that there was still some association between education and CVD mortality in our study after adjustment for risk factors does not preclude the possibility of residual confounding. This may have occurred because of measurement bias, particularly as behavioural and social factors were self-reported.

CONCLUSION

The decline in CVD mortality over the past decades has not occurred equitably, with a widening of relative inequality in many countries, including Australia. It has been suggested that because traditional risk factors do not account for all of the SES/CVD gradient, addressing them within disadvantaged groups will not significantly reduce health inequalities. While we agree that health inequalities cannot be fully reduced until the underlying determinants of social disadvantage are addressed, these data suggest that addressing traditional risk factors (particularly smoking and central adiposity) may strongly influence the current gradient in CVD. These findings reinforce the need for public health policies and strategies to effectively address both behavioural and physiological risk factors among lower socio-economic groups.

Acknowledgements

This study was made possible by the contribution of many people, including the original investigators and the team who recruited the participants. We would like to express our gratitude to the many thousands of Melbourne residents who continue to participate in the study.

Funding

This work was supported by the National Health and Medical Research Council (NHMRC) (ID No 208057, 334032, 396414). Further infrastructure support was provided by The Cancer Council Victoria and Monash University. AB is a PhD scholar funded by the NHMRC (ID No 465352). GT is a Senior Research Fellow funded by the NHMRC (ID No 390109). AF is funded by a VicHealth Fellowship. Cohort recruitment was funded by VicHealth and The Cancer Council Victoria.

Ethics approval

Ethics approval was provided by the Cancer Council Victoria Human Research Ethics Committee.

Provenance and peer review

Not commissioned; externally peer reviewed.

REFERENCES

Research report


Findings from this analysis suggest that reducing the burden of risk factors among lower socioeconomic groups would significantly impact upon subsequent inequalities in CVD mortality. We found a strong educational gradient in most major risk factors for CVD, with smoking identified as the most important individual risk factor explaining the difference in CVD mortality between the highest and lowest education groups. The next chapter aims to further explore these findings using two separate analyses. First, we examine a possible mechanism by which smoking may lead to CVD. Second, the association between SES itself and the onset of the gradient in risk factor prevalence and number is explored.
58

CHAPTER 5

Exploring mechanisms and pathways between socioeconomic status and cardiovascular risk

5.1 Chapter overview

In chapter 2, it was shown that while inequalities in cardiovascular risk factors are thought to be a major contributor to inequalities in health outcomes, the mechanisms and pathways by which this occurs are not clearly elucidated. The two papers presented in this chapter explore this, using very different perspectives.

The first analysis presented in this chapter investigates the relationships between smoking status, lifestyle-related risk factors and lipoprotein subclasses in the MCCS. As described earlier, the INTERHEART study found abnormal lipids and smoking to be the two risk factors accounting for the greatest proportion of the population attributable risk for AMI. [58] However, the pathways by which smoking leads to CVD are unclear. Many studies have shown that smokers have a more adverse lipoprotein profile compared to non-smokers, [216] yet it is not certain whether this association is mediated by lifestyle-related factors, known to be more prevalent among smokers than non-smokers. [217] This analysis therefore aims to examine whether smoking-related changes in lipoprotein subclasses occur independently of diet, alcohol and physical activity, in order to identify whether treatment of lipids among smokers will require more than management of these lifestyle factors. While the somewhat small sample size in this study did not allow for stratification by a measure of SES, the analysis is particularly relevant for this thesis as socioeconomic disadvantage is known to be associated with higher smoking rates, and as shown in chapter 4, smoking is the most important risk factor contributing to the social gradient in CVD.

The second analysis in this chapter aimed to determine whether SES is a predictor of the earlier onset of certain risk factors and of the number of risk factors accumulated over time. Lower SES is known to be associated with a higher prevalence and clustering of major risk factors for CVD,
leading to increased “overall” risk of the disease. However, few studies have examined either the incidence of cardiovascular risk factors or changes in cumulative risk over time according to SES, [104, 110, 218, 219] and so our understanding of how social gradients in risk factors begin and develop in individuals over time is limited. Findings from this analysis will provide further evidence for the role of SES in the development of gradients in CVD, and will also allow us to determine which risk factors we should be particularly targeting in order to reduce future social gradients in the disease.
Monash University  

Declaration for Thesis Chapter 5  

Declaration by candidate  

In the case of chapter 5, the nature and extent of my contribution to the work 'Associations between smoking status, lifestyle and lipoprotein subclasses' was the following:  

<table>
<thead>
<tr>
<th>Nature of contribution</th>
<th>Extent of contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical literature review and development of the research questions, data preparation, analysis design, statistical analysis, interpretation of results, tables, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
<td>75</td>
</tr>
</tbody>
</table>

The following co-authors contributed to the work:  

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Andrew M. Tonkin</td>
<td>Grants used to fund research, facilitated development of research questions, interpretation of results, commented on manuscript drafts,</td>
</tr>
<tr>
<td>2 Anna Peeters</td>
<td>Assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>3 Rory Wolfe</td>
<td>Commented on data, assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>4 Gavin Turrell</td>
<td>Facilitated development of research questions, commented on manuscript drafts</td>
</tr>
<tr>
<td>5 Linton R Harriss</td>
<td>Commented on data, assisted with interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>6 Graham G Giles</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>7 Dallas R English</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research</td>
</tr>
<tr>
<td>8 Alicia J Jenkins</td>
<td>Grants used to fund research, commented on data, assisted with design of analysis, interpretation of results, commented on manuscript drafts</td>
</tr>
</tbody>
</table>

Candidate's Signature

Date 20/12/18
Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

<table>
<thead>
<tr>
<th>Location(s)</th>
<th>Department of Epidemiology &amp; Preventive Medicine, Alfred Campus, Monash University</th>
</tr>
</thead>
</table>

[Please note that the location(s) must be institutional in nature, and should be indicated here as a department, centre or institute, with specific campus identification where relevant.]

<table>
<thead>
<tr>
<th>Signature 1</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8/12/2010</td>
</tr>
<tr>
<td>Signature 2</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>15/12/2010</td>
</tr>
<tr>
<td>Signature 3</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>13/12/2010</td>
</tr>
<tr>
<td>Signature 4</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>18/10/10</td>
</tr>
<tr>
<td>Signature 5</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>25/10/2010</td>
</tr>
<tr>
<td>Signature 6</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>5/11/2010</td>
</tr>
<tr>
<td>Signature 7</td>
<td>Date</td>
</tr>
<tr>
<td></td>
<td>29/11/2010</td>
</tr>
</tbody>
</table>
Associations among smoking status, lifestyle and lipoprotein subclasses

Alison Beauchamp, BHSc MPH*, Andrew Tonkin, MD, FRACP, FCSANZ, Anna Peeters, BSc(Hons), PhD, Rory Wolfe, BSc, PhD, Gavin Turrell, BA, PhD, Linton Harriss, RN, MPH, PhD, Graham G. Giles, BSc, MSc, PhD, Dallas R. English, BSc, MSc, PhD, Alicia J. Jenkins, MD, FRACP

Department of Epidemiology and Preventive Medicine, Monash University, Alfred Hospital, Melbourne 3004, Australia (Drs. Tonkin, Peeters, Wolfe, Harriss, and Giles; and A. Beauchamp); School of Public Health, Queensland University of Technology, Kelvin Grove, Brisbane, Australia (Dr. Turrell); Cancer Epidemiology Centre, Cancer Council Victoria, Drummond Street, Melbourne, Australia (Drs. Giles and English); Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, University of Melbourne, Bouverie St, Melbourne, Australia (Drs. Giles and English); and Department of Medicine, University of Melbourne, Victoria Parade, Melbourne, Australia (Dr. Jenkins)

KEYWORDS:
Atherosclerosis; Lifestyle; Lipoproteins; Smoking; Spectroscopy

BACKGROUND: The relationship between cigarette smoking and cardiovascular disease is well established, yet the underlying mechanisms remain unclear. Although smokers have a more atherogenic lipid profile, this may be mediated by other lifestyle-related factors. Analysis of lipoprotein subclasses by the use of nuclear magnetic resonance spectroscopy (NMR) may improve characterisation of lipoprotein abnormalities.

OBJECTIVE: We used NMR spectroscopy to investigate the relationships between smoking status, lifestyle-related risk factors, and lipoproteins in a contemporary cohort.

METHODS: A total of 612 participants (360 women) aged 40–69 years at baseline (1990–1994) enrolled in the Melbourne Collaborative Cohort Study had plasma lipoproteins measured with NMR. Data were analysed separately by sex.

RESULTS: After adjusting for lifestyle-related risk factors, including alcohol and dietary intake, physical activity, and weight, mean total low-density lipoprotein (LDL) particle concentration was greater for female smokers than nonsmokers. Both medium- and small-LDL particle concentrations contributed to this difference. Total high-density lipoprotein (HDL) and large-HDL particle concentrations were lower for female smokers than nonsmokers. The proportion with low HDL particle number was greater for female smokers than nonsmokers. For men, there were few smoking-related differences in lipoprotein measures.

CONCLUSION: Female smokers have a more atherogenic lipoprotein profile than nonsmokers. This difference is independent of other lifestyle-related risk factors. Lipoprotein profiles did not differ greatly between male smokers and nonsmokers.

© 2010 National Lipid Association. All rights reserved.

* Corresponding author.
E-mail address: alison.beauchamp@monash.edu
Submitted July 1, 2010. Accepted for publication September 19, 2010.

Cigarette smoking is a major risk factor for cardiovascular disease (CVD).\(^1\) Although underlying mechanisms are not fully elucidated, many investigators demonstrate
an atherogenic lipid profile for smokers, including greater total cholesterol, low-density lipoprotein (LDL)-cholesterol and triglycerides, and lower (usually vasoprotective) high-density lipoprotein (HDL)-cholesterol levels.\footnote{Beauchamp et al Smoking, lifestyle, and lipoprotein subclasses 523} It is unclear as to the extent by which relationships between smoking and lipoproteins are influenced by other lifestyle-related risk factors, such as poor nutrition, inactivity, and adiposity.\footnote{Baseline demographics and cardiovascular risk factors} Furthermore, because of differences in lipoprotein composition, patients with the same lipid concentrations may have substantially different CVD risk, particularly in the context of comorbidities such as diabetes.\footnote{Smoking status} Detailed lipoprotein subclass analysis might therefore increase understanding of the relationship between smoking, other lifestyle-related risk factors, lipoproteins and CVD risk.\footnote{Methods}

There are three major lipoprotein classes: very-low-density lipoprotein (VLDL), LDL, and HDL. Intermediate-density lipoprotein, also known as VLDL remnants, is highly atherogenic but usually exists at low levels. There are at least three main subclasses, large, medium, and small, recognized for each of VLDL, LDL, and HDL.\footnote{Baseline demographics and cardiovascular risk factors} Traditional methods of measuring lipoprotein subclasses include gradient gel electrophoresis and density-gradient ultracentrifugation, both of which are time-consuming and not clinically available. An accurate, faster and more available approach to their measurement is nuclear magnetic resonance (NMR) spectroscopy.\footnote{Smoking status} NMR studies have shown that LDL particle concentration is independently associated with and a better predictor of CVD events than LDL-cholesterol (LDL-C) concentrations\footnote{Baseline demographics and cardiovascular risk factors} and that both small LDL and large VLDL are positively associated with coronary heart disease (CHD), whereas large HDL are protective.\footnote{Baseline demographics and cardiovascular risk factors}

In the Framingham Offspring Study, NMR has been used to investigate the relationship between smoking, small LDL concentrations, and estrogen receptor alpha gene variation\footnote{Baseline demographics and cardiovascular risk factors} and also to evaluate smoking intensity effects on lipoprotein concentrations for current smokers.\footnote{Baseline demographics and cardiovascular risk factors} No NMR studies have examined differences between non-smokers, former smokers, and current smokers nor controlled for potential influences of lifestyle-related risk factors. In this Australian study, we investigated such relationships in a contemporary cohort of men and women.

**Methods**

**Study sample**

The prospective Melbourne Collaborative Cohort Study recruited 41,514 subjects ages 27 to 80 years in 1990 to 1994. Twenty-four percent were southern European migrants, deliberately oversampled to extend the range of lifestyle factors and genetic variation. Study details have been published elsewhere.\footnote{Approval from The Cancer Council Victoria’s Human Research Ethics Committee and written informed consent from participants were obtained. Baseline examination included face-to-face interviews and questionnaires in the subject’s preferred language.} A random sample of 934 fasted subjects who met the eligibility criteria of having 3 mL of of stored plasma and a fasting plasma glucose \(<\text{7.0 mmol/L (Kodak Ektachem analyzer; Rochester, NY) was selected. Participants were excluded for subsequent cancer development (n = 35); history of heart attack, stroke, angina, or diabetes (n = 71); currently taking lipid-lowering therapy (n = 17); unknown menopause status (n = 12); missing data on cigarette smoking or covariates (n = 2); or missing data on conventional lipids (n = 125). Data from 612 participants (360 women and 252 men) were analyzed. Separate analyses were undertaken for men and women because of known gender differences in lipids and NMR-determined lipoproteins.}\footnote{Smoking status}

Smoking status was ascertained from questions modified from the validated Medical Research Council 1986 Respiratory Symptoms Questionnaire.\footnote{Smoking status} Smokers were defined as people currently smoking seven or more cigarettes a week, nonsmokers as never having smoked, and former smokers as having previously smoked but now ceased, or those smoking less than seven cigarettes a week. Time since smoking cessation ranged from 1 to 25 years.

**Baseline demographics and cardiovascular risk factors**

Country of birth was categorized as: 1) Australia/New Zealand/northern Europe (primarily UK and the Netherlands), or 2) southern Europe (mainly Italy and Greece). Additional covariates for women included oral contraceptive use (never/past and current), hormone-replacement therapy (never/past and current) and menopausal status (pre- and post-).

Participants were asked their usual quantity and frequency of alcohol intake and were considered current drinkers if they responded yes to: “Have you ever drunk at least 12 alcoholic drinks in a year (sips don’t count)?” Physical activity during the previous six months was based on the number of times per week that exercise was undertaken at a vigorous, less-vigorous, or walking-level only. These data were combined to give an overall score of relative energy expenditure in four categories, based on the Compendium of Physical Activities.\footnote{Dietary information was collected by the use of a validated, self-administered food frequency questionnaire.} Dietary information was collected by the use of a validated, self-administered food frequency questionnaire.\footnote{Average intake of 121-items over the previous year was used to calculate average daily nutrient intake.Waist-hip ratio and body mass index were calculated from direct measurements by the use of standard anthropometric methods.} The second and third of three blood pressure readings measured (DINAMAP 1846SX) after 5 minutes of supine rest were used in analyses. Metabolic syndrome
Table 1  Characteristics of 612 men and women aged 40-69 years (1990-1994) in the Melbourne Collaborative Cohort Study, according to smoking status

<table>
<thead>
<tr>
<th>Smoking status</th>
<th>Women (n = 360)</th>
<th>Men (n = 252)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non (n = 262)</td>
<td>Current (n = 32)</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>53.8 (± 8.5)</td>
<td>52.4 (± 8.3)</td>
</tr>
<tr>
<td>Mediterranean born, n (%)</td>
<td>87 (33%)</td>
<td>11 (34%)</td>
</tr>
<tr>
<td>Primary only</td>
<td>82 (31%)</td>
<td>14 (44%)</td>
</tr>
<tr>
<td>Some high</td>
<td>113 (43%)</td>
<td>12 (38%)</td>
</tr>
<tr>
<td>Full high/ tertiary</td>
<td>67 (26%)</td>
<td>6 (19%)</td>
</tr>
<tr>
<td>Cigarettes/ day (mean ± SD)</td>
<td>217 (± 10)</td>
<td>–</td>
</tr>
<tr>
<td>Years since cessation</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Post-menopausal, n (%)</td>
<td>159 (61%)</td>
<td>17 (53%)</td>
</tr>
<tr>
<td>HRT use (current), n (%)</td>
<td>21 (8%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Oral contraceptives (current), n (%)</td>
<td>4 (2%)</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Physical activity, n (%)</td>
<td>138 (53%)</td>
<td>19 (59%)</td>
</tr>
<tr>
<td>Moderate/high</td>
<td>124 (47%)</td>
<td>13 (41%)</td>
</tr>
<tr>
<td>Current drinkers, n (%)</td>
<td>120 (46%)</td>
<td>18 (56%)</td>
</tr>
<tr>
<td>Alcohol gm/d (drinkers) (median, interquartile range)</td>
<td>7.7 (2.1, 15)</td>
<td>14.4 (5.1, 27.8)</td>
</tr>
<tr>
<td>Carbohydrate gm/d (mean ± SD)</td>
<td>243 (± 117)</td>
<td>231 (± 104)</td>
</tr>
<tr>
<td>Energy intake KJ/d (median, IQR)</td>
<td>8215 (6626, 10365)</td>
<td>8143 (6235,10771)</td>
</tr>
<tr>
<td>Saturated fat gm/d (mean ± SD)</td>
<td>31.7 (± 14.8)</td>
<td>34.9 (±21.4)</td>
</tr>
<tr>
<td>Blood glucose, mmol/L, mean (± SD)</td>
<td>5.36 (± 0.44)</td>
<td>5.37 (± 0.41)</td>
</tr>
<tr>
<td>Waist-hip ratio, mean (±SD)</td>
<td>0.78 (± 0.07)</td>
<td>0.78 (± 0.06)</td>
</tr>
<tr>
<td>Body mass index, mean (±SD)</td>
<td>27.3 (± 4.8)</td>
<td>25.5 (± 4.2)*</td>
</tr>
<tr>
<td>Normal (&lt;25)</td>
<td>88 (34%)</td>
<td>17 (53%)</td>
</tr>
<tr>
<td>Overweight (25-29.9)</td>
<td>113 (43%)</td>
<td>10 (31%)</td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td>61 (23%)</td>
<td>5 (16%)</td>
</tr>
<tr>
<td>Systolic BP, mmHg, mean (±SD)</td>
<td>134.4 (±22.7)</td>
<td>133.8 (±14.6)</td>
</tr>
<tr>
<td>Diastolic BP, mmHg, mean (±SD)</td>
<td>72.4 (±12.9)</td>
<td>73.4 (±11.0)</td>
</tr>
<tr>
<td>Pulse pressure, mean (±SD)</td>
<td>62.1 (± 15.3)</td>
<td>60.4 (± 12.0)</td>
</tr>
</tbody>
</table>
diagnosis was determined by the 2006 International Diabetes Federation criteria.18

**Conventional lipid measures**

Total cholesterol, HDL-cholesterol (HDL-C), and triglyceride levels were measured by the use of methods previously described.19 LDL-C was calculated using the Friedewald formula.20

**NMR-measured lipoproteins**

VLDL, LDL, and HDL subclass particle concentrations and mean particle diameters were measured with an automated NMR spectroscopic assay as previously described.7 CHD risk according to total LDL particle number was defined as follows:21 Total LDL particle number as >2000 nm/L and 1600–2000 nm/L (both associated with increased risk); small LDL particle number as >1200 nmol/L and 850–1200 nmol/L (associated with increased and borderline CHD risk respectively); and large HDL particle number as <4.0 mmol/L (also associated with increased CHD risk). Small LDL (pattern B) was defined as mean LDL diameter <20.5 nm.

**Statistical analyses**

Stata 9.2 (Stata Corp, College Station, TX) was used. Linear regression models for lipoprotein subclass measurements analysed the associations with smoking, adjusted for age, country of birth, education, alcohol, physical activity, waist-to-hip ratio, body mass index, total daily energy, carbohydrate, and saturated fat intake. Oral contraceptive use, hormone-replacement therapy, and menopausal status were also adjusted for in women. Two-sided $P$ values are presented, and values <.05 were regarded as significant. Because there was collinearity between total daily energy and carbohydrate intake, the carbohydrate variable was replaced by converting carbohydrate (g/day) to an energy equivalent and dividing by total daily energy intake. Large and medium VLDL and medium HDL followed skewed distributions. VLDL lipoproteins were therefore log transformed before analysis, and medium HDL data were analysed by the use of logistic regression. Analysis of differences in CHD risk between smokers and nonsmokers according to lipoprotein particle size and number was performed by the use of univariate logistic regression.

The percentage difference in lipoprotein concentrations for smokers compared with nonsmokers was calculated as follows:

\[
\left( \frac{\text{mean lipoprotein value in smokers} - \text{mean lipoprotein value in non-smokers}}{\text{standard deviation of lipoprotein in non-smokers}} \right) \times 100
\]
Results

Subjects

Baseline characteristics in Table 1 show differences between males and females subjects in the comparability of smokers and former smokers with nonsmokers. For women, current smokers had a lower mean body mass index than nonsmokers and were more likely to be within normal weight limits. Compared with female nonsmokers, former smokers were less likely to have been born in a Mediterranean country, to have primary education only, and to be obese. Additionally, female former smokers were more likely to drink alcohol than nonsmokers.

For men, current smokers were more likely to be Mediterranean-born, to have primary education only, and to be physically inactive than nonsmokers. Compared with male nonsmokers, former smokers were older, more likely to be Mediterranean-born, and to have received a primary education only. Male former smokers also had a greater daily alcohol intake, lower carbohydrate intake, lower saturated fat intake, and were more likely to be obese than nonsmokers.

Conventional lipids

Females

As shown in Table 1, female smokers had greater triglycerides and total cholesterol/HDL-C ratio, and a lower HDL-C than nonsmokers in both univariate analysis and after adjustment for lifestyle-related factors. LDL-C levels were greater in smokers than nonsmokers, although these differences were only apparent after multivariate adjustment.

In female former smokers, triglycerides were higher compared with nonsmokers (P = .02). HDL-C was greater in former smokers than current smokers (P = .02), and total/HDL-C ratio was lower (P = .001), with similar concentrations of both these lipids in nonsmokers and former smokers. All differences were significant after multivariate adjustment.

Males

In men, the only significant smoking-related differences were in total and LDL-cholesterol, which were both less in former smokers than in current smokers (both P = .03) after multivariate adjustment, with levels similar to those of nonsmokers. There were no significant differences in triglycerides between smoking categories.

NMR-measured lipoproteins

Females

As shown in Table 2, no differences were seen in VLDL particle concentrations, but VLDL was larger in smokers compared with nonsmokers. Median total LDL particle concentration was also higher in smokers than nonsmokers, with both medium and small LDL particle concentration contributing to this difference. The evidence for these findings was compelling, regardless of whether or not lifestyle factors were adjusted for. Median total HDL and large HDL levels were lower in smokers than for nonsmokers in both univariate and multivariate analysis.

Females

Figure 1 shows the percentage difference in crude mean values of NMR-determined lipoprotein measures in female subjects. Compared with nonsmokers, current smokers had higher values for total LDL and small LDL of 49% and 37%, respectively, and lower values of total HDL and large HDL of 41% and 33%, respectively.

Males

There were few smoking-related differences in lipoprotein particle concentration or size in males (Table 2). After multivariate analysis, large LDL concentration and LDL size were greater in current smokers compared with nonsmokers. Small VLDL concentration was lower in nonsmokers than in smokers (P = .01), and VLDL size was increased (P = .04).

Proportions at increased risk of CHD

The proportion of men and women with increased CHD risk NMR profiles is shown in Table 3. On the basis of total LDL particle levels, 22% of female smokers were at increased CHD risk compared with 4% of female nonsmokers (P < .001). A greater proportion of female smokers had small LDL particle number >1200 nmol/L and low levels of large HDL than nonsmokers.

For men, there were no statistically significant differences in the proportions of those with NMR profiles associated with increased CHD, although there was a suggestive trend towards more smokers than nonsmokers with mean LDL particle size ≤20.5 nm (P = .06); an opposite pattern to that observed in women.

Discussion

In this work we demonstrate that female smokers have a more atherogenic lipoprotein profile than nonsmokers, although this profile is less adverse in ex-smokers. Differences are seen in both conventional lipid concentrations and in NMR-determined lipoprotein profiles and are independent of lifestyle-related factors. There were few smoking-related changes in lipoproteins for males.
Table 2  Lipoprotein measures for 612 men and women aged 40-69 years (1990-1994) in the Melbourne Collaborative Cohort Study, according to smoking status *

<table>
<thead>
<tr>
<th></th>
<th>Women (n = 360)</th>
<th>Men (n = 252)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonsmokers n = 262</td>
<td>Current smokers n = 32</td>
</tr>
<tr>
<td>Total VLDL particle concentration, nmol/L</td>
<td>56.3 (37.8, 77.6)</td>
<td>53.5 (22.9, 81.2)</td>
</tr>
<tr>
<td>Large VLDL</td>
<td>0.7 (0.2, 2.2)</td>
<td>0.7 (0.2, 2.7)</td>
</tr>
<tr>
<td>Medium VLDL</td>
<td>13.4 (5.5, 22.5)</td>
<td>7.7 (3.6, 19.1)</td>
</tr>
<tr>
<td>Small VLDL</td>
<td>40.7 (28.3, 55.5)</td>
<td>39.9 (19.0, 56.6)</td>
</tr>
<tr>
<td>Total LDL particle concentration, nmol/L</td>
<td>1291 (1089, 1549)</td>
<td>1467 (1149, 750)</td>
</tr>
<tr>
<td>Large LDL</td>
<td>703 (575, 880)</td>
<td>733 (612, 883)</td>
</tr>
<tr>
<td>Medium LDL</td>
<td>105 (49, 160)</td>
<td>134 (67, 172)</td>
</tr>
<tr>
<td>Small LDL</td>
<td>423 (225, 464)</td>
<td>549 (301, 678)</td>
</tr>
<tr>
<td>Ratio: Small/total LDL</td>
<td>0.43 (0.24, 0.56)</td>
<td>0.43 (0.32, 0.64)</td>
</tr>
<tr>
<td>Total HDL particle concentration, mmol/L</td>
<td>34.7 (30.9, 38.7)</td>
<td>32.5 (28.9, 33.4)</td>
</tr>
<tr>
<td>Large HDL</td>
<td>9.2 (7.1, 11.3)</td>
<td>8.1 (5.8, 9.0)</td>
</tr>
<tr>
<td>Medium HDL</td>
<td>0.6 (0.3, 0.5)</td>
<td>1 (0, 3.5)</td>
</tr>
<tr>
<td>Small HDL</td>
<td>23.4 (19.6, 27.2)</td>
<td>22.2 (17.3, 25.4)</td>
</tr>
<tr>
<td>Ratio: small/total HDL</td>
<td>0.68 (0.60, 0.74)</td>
<td>0.69 (0.56, 0.78)</td>
</tr>
<tr>
<td>NMR Lipoprotein particle size, mean diameter in nm)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VLDL particle size</td>
<td>45.8 (42.6, 49.5)</td>
<td>49.2 (44.5, 53.7)</td>
</tr>
<tr>
<td>LDL particle size</td>
<td>21.6 (21.2, 22.1)</td>
<td>21.6 (21.2, 22)</td>
</tr>
<tr>
<td>HDL particle size</td>
<td>9.2 (9.6)</td>
<td>9.3 (8.9, 9.6)</td>
</tr>
</tbody>
</table>

HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very-low-density lipoprotein.

*Smoking status ascertained by self report.
†P value for difference between nonsmokers and smokers by the use of linear regression.
‡P value < .05 for difference between non- and former smokers by the use of multivariate linear regression.
§P value < .05 for difference between current and former smokers using multivariate linear regression.
Female analyses: Multivariate analysis adjusted for age, country of birth, education, oral contraceptive use, hormone replacement therapy and menopausal status, alcohol, physical activity, WHR, BMI, total daily energy, carbohydrate, and saturated fat intake.

Male analyses: Multivariate analysis adjusted for age, country of birth, education, alcohol, physical activity, WHR, BMI, total daily energy, carbohydrate, and saturated fat intake.

**Significance.
Current smokers

For female smokers, greater total triglycerides, LDL-C and total/HDL levels, and lower HDL-C have been previously reported. Detailed NMR-determined lipoprotein analyses demonstrated that female smokers had a greater total LDL particle concentration than nonsmokers, contributed to by both medium and small LDL. Female smokers also had lower total HDL particle concentrations than nonsmokers, primarily because of a decrease in large HDL particles. Small, dense LDL particles are more atherogenic than large, buoyant LDL particles, and large HDL particles appear to be associated with greater vasoprotection than small HDL particles. Therefore, our findings point to a more atherogenic lipoprotein profile for female smokers compared with nonsmokers. Furthermore, although the percentage differences in the median levels of these lipoprotein subclasses for female smokers are small, they still lead to large differences in the proportions of those with higher values. For example, the proportion of female smokers with increased-risk small LDL was 33% greater than that of nonsmokers.

Our results are similar to NMR substudy in The Framingham Offspring Study. An earlier British cross-sectional study in which the authors used gradient gel electrophoresis also found that lower HDL-C levels for female smokers related to lower concentrations of large HDL subfractions.

Former smokers

Female former smokers had greater HDL-C and lower total/HDL-C levels than current smokers and similar levels to nonsmokers. Other studies combining men and women have also shown that smoking cessation is associated with a return of HDL-C to levels seen for nonsmokers. Plasma HDL-C concentrations increase within 2 to 8 weeks after smoking cessation, but temporal associations could not be examined in our study because smoking cessation was measured in years only.

Relative to current smokers, female former smokers also had greater levels of total HDL and large HDL particle concentrations. In addition, both total and medium LDL particle concentrations were lower than those of current smokers, suggesting reversibility of smoking effects on these lipoprotein subclasses. No other studies have examined effects of smoking cessation on NMR-derived lipoprotein measures.

Differences between men and women

There were few statistically significant changes in conventional or NMR-determined lipoprotein profiles by smoking categories for men. Other investigators also have found that smoking-related lipoprotein changes are less for men than women. For example, the investigators from the Munster Heart Study (PROCAM) study reported similar HDL-C and triglyceride concentrations for male and female smokers, but the increases in total and LDL-cholesterol for women were two and almost four-fold (respectively) that of men.

Freeman et al found similar triglyceride levels for male current smokers and nonsmokers, although differences existed for women. In addition, HDL₂ (the larger, more buoyant and cholesterol-rich HDL subclass as measured using analytical ultracentrifugation) was lower for both men and women, but only reached statistical significance for women. These differences may relate to underlying sex differences in smoking effect mechanisms that may be hormonal in nature.

Lifestyle factors and lipids

Consistent with our findings, other investigators have shown that the adverse lipid profile of smokers is independent of potential confounding lifestyle factors. We found that smoking-related differences in lipoprotein measures remained significant after adjustment for body weight, activity, alcohol and diet, all of which are known to affect LDL or HDL levels. There were few
differences in these lifestyle factors between the smoking groups in our cohort. Therefore, our findings may be due to this homogeneity, rather than to any true lack of effect. In our study, female former smokers were more likely to be current alcohol drinkers than were nonsmokers and to also have greater levels of HDL-C, HDL particle concentration and large HDL. Moderate alcohol intake by women has been associated with increased HDL-C in other studies\(^2\) and also with greater levels of HDL particle concentration and large HDL for older adults in the Cardiovascular Health Study.\(^2\)

**Limitations**

This study was conducted in Melbourne, Australia. Although these findings may not be applicable to other regions, our results support those in the (United States) Framingham Offspring Study.\(^3\) Generalizability may also be affected by oversampling of southern Europeans. In addition, smoking and other behavioral risk factors were self-reported. Any resultant misclassification, assuming it was not differential according to smoking status, was likely to bias findings towards null effects.

Estimation of LDL-C from the Friedewald formula may be fallacious when triglyceride levels are \(>4.5\) mmol/L.\(^3\) However in the present study, only one female smoker had such a level. A dose-response between the intensity of smoking and lipids has been seen in other studies.\(^1\) However, because of the relatively small sample size in our study, we unable to investigate effects according to number of cigarettes smoked/day or other sub-group analysis such as alcohol consumption. Finally, although the cross-sectional study design means it is not possible to ascertain any causal relationship between smoking and lipoprotein, a significant body of evidence from other studies supports the causal nature of the association.\(^1\)–\(^4\)

In summary, the smoking-related differences we observed in our study occurred independently of lifestyle factors, indicating that treatment of dyslipidemia, particularly for female smokers, may require more than adherence to guidelines for activity, diet and alcohol intake. By identifying the specific contribution of lipoprotein subclasses to the adverse lipoprotein profile and subsequent increased CVD risk for smokers, our findings have important implications for the interpretation of results and management of dyslipoproteinemia.

**Acknowledgments**

The authors acknowledge the MCCS investigators and participants.

**Financial disclosure**

Grants were received from VicHealth, The Cancer Council Victoria, the National Health and Medical Research Council.
References


Monash University

Declaration for Thesis Chapter 5

Declaration by candidate

In the case of Chapter 5, the nature and extent of my contribution to the work 'Incidence of cardiovascular risk factors by education level 2000-2005: the Australian Diabetes, Obesity, and Lifestyle (AusDiab) Cohort Study' was the following:

<table>
<thead>
<tr>
<th>Nature of contribution</th>
<th>Extent of contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical literature review and development of the research questions, data preparation, analysis design, statistical analysis, interpretation of results, tables, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
<td>75</td>
</tr>
</tbody>
</table>

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Rory Wolfe</td>
<td>Commented on data, design of analysis, interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>2 Diana Magliano</td>
<td>Grants used to fund research, commented on data and manuscript drafts</td>
</tr>
<tr>
<td>3 Gavin Turrell</td>
<td>Commented on data, interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>4 Andrew M. Tonkin</td>
<td>Commented on data, interpretation of results, commented on manuscript drafts, commented on manuscript drafts, commented on manuscript drafts</td>
</tr>
<tr>
<td>5 Jonathan Shaw</td>
<td>Grants used to fund research, design and conduct of AusDiab study, commented on manuscript drafts</td>
</tr>
<tr>
<td>6 Anna Peeters</td>
<td>Facilitated development of research questions, design of analysis, interpretation of results, commented on manuscript drafts</td>
</tr>
</tbody>
</table>

Candidate’s Signature

Date: 20/12/10
Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

<table>
<thead>
<tr>
<th>Location(s)</th>
<th>Department of Epidemiology &amp; Preventive Medicine, Alfred Campus, Monash University</th>
</tr>
</thead>
</table>

[Please note that the location(s) must be institutional in nature, and should be indicated here as a department, centre or institute, with specific campus identification where relevant.]

<table>
<thead>
<tr>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature 1</td>
<td>13/12/2010</td>
</tr>
<tr>
<td>Signature 2</td>
<td>16/12/2010</td>
</tr>
<tr>
<td>Signature 3</td>
<td></td>
</tr>
<tr>
<td>Signature 4</td>
<td>8/12/2010</td>
</tr>
<tr>
<td>Signature 5</td>
<td>16/12/2010</td>
</tr>
<tr>
<td>Signature 6</td>
<td>15/12/2010</td>
</tr>
</tbody>
</table>
INCIDENCE OF CARDIOVASCULAR RISK FACTORS BY EDUCATION LEVEL 2000-2005: THE AUSTRALIAN DIABETES, OBESITY, AND LIFESTYLE (AUSDIAB) COHORT STUDY

Corresponding Author: Alison Beauchamp
Postal address: Department of Epidemiology & Preventive Medicine
Monash University, Alfred Hospital
Melbourne 3004, Australia
Telephone +61 (0)3 9903 0555
Fax +61 (0)3 9903 0556

Authors:
Alison Beauchamp¹ BHSc MPH; Rory Wolfe¹ BSc PhD; Dianna J Magliano ¹² BAppS (Hons) MPH PhD;
Gavin Turrell³ BA PhD; Andrew Tonkin¹ MBBS MD FRACP FCSANZ; Jonathan Shaw¹² MD MRCP (UK)
FRACP; Anna Peeters¹ BSc(Hons) PhD.

¹. Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia,
². International Diabetes Institute, Melbourne, Australia
³. School of Public Health, Queensland University of Technology, Brisbane, Australia
ABSTRACT

BACKGROUND Lower socioeconomic status (SES) is associated with a higher prevalence of major risk factors for cardiovascular disease (CVD). However, few longitudinal studies have examined the association between SES and CVD risk factors over time. We aimed to determine whether SES, using education as a proxy, is associated with the onset of CVD risk factors over 5 years in an Australian adult cohort study.

METHODS Participants in the Australian Diabetes, Obesity and Lifestyle study (AusDiab) study aged 25 years and over who attended both baseline and 5-year follow-up examinations (n=5967) were categorised according to educational attainment. Cardiovascular risk factor data at both time points were ascertained through questionnaire and physical measurement.

RESULTS Women with lower education had a greater risk of progressing from normal weight to overweight or obesity than those with higher education (age-adjusted OR 1.57, 95% CI 1.06-2.31). Both men and women with lower education were more likely to develop diabetes (age-adjusted OR from higher education 1.75, 95% CI 1.14-2.71 and 3.01, 95% CI 1.26-7.20, respectively). A lower level of education was associated with a greater number of risk factors accumulated over time in women (OR of progressing from having two or less risk factors at baseline to three or more at follow up, 2.04, 95% 1.32-3.14).

CONCLUSION In this Australian population-based study, lower educational attainment was associated with an increased risk of developing both individual and total CVD risk factors over a 5-year period. These findings suggest that SES inequalities in CVD will persist into the future.

KEYWORDS Socioeconomic status, risk factor incidence, cardiovascular disease, diabetes, obesity
BACKGROUND

The burden of cardiovascular disease (CVD) is significantly greater among lower socioeconomic groups [1-3]. While the mechanisms and pathways between social disadvantage and increased risk of CVD are not fully understood, major risk factors for atherothrombotic disease are thought to play a significant role. Tobacco smoking, abnormal lipids, high blood pressure, diabetes and abdominal obesity in combination account for up to 90% of the population attributable risk of acute myocardial infarction [4]. In addition, many longitudinal studies have reported that social gradients in the prevalence of these and other risk factors account for a significant proportion of the social gradient in CVD [5, 6]. People from lower socioeconomic groups also tend to have a higher number of cardiovascular risk factors, leading to an increased overall risk of CVD among the more disadvantaged. [7]

Despite this evidence, understanding of how social gradients in CVD risk develop in individuals over time remains limited. Prospective studies describing the incidence of cardiovascular risk factors according to socioeconomic status (SES) are few, and findings are inconsistent. While several studies have found that the incidence of hypertension is higher among lower socioeconomic groups, [8-10] others have shown that these associations vary according to age, race and gender [11-13]. Findings are equally inconsistent for incident obesity, [14-16] and diabetes [17-20]. In addition, while it is known that risk factors in combination are more closely associated with CVD risk than single factors in isolation, [21] and that lower SES groups tend to have a greater number of cardiovascular risk factors, [7] few longitudinal studies have examined whether SES is associated with the accumulation of cardiovascular risk factors over time. [22] Further evidence for the association between SES and the incidence of both individual and cumulative risk factors will make an important contribution to our knowledge of which factors to target in order to reduce future inequalities in CVD.

Using data from an Australian adult cohort study, we aimed to determine whether SES is a predictor of incident cardiovascular risk factors, and whether SES is associated with the number of risk factors which develop over time.
METHODS

The Australian Diabetes, Obesity and Lifestyle study (AusDiab) is a population-based, stratified cluster survey of 11,247 adults aged 25 years or older in 1999-2000. Methods and response rates have been described previously. [23] A five-year follow-up was conducted in 2004-2005. From the original cohort there were 10,788 participants eligible for follow-up and of these, 6400 returned for physical examination and interviewer-administered questionnaire. For this analysis we excluded participants missing baseline data on education (n=47), diabetes or CVD (n=72), and baseline or follow-up data on smoking, systolic blood pressure, cholesterol, body mass index (BMI), and medication use (n=314). Following these exclusions, there was a total of 5967 participants who had attended both baseline and follow-up examinations. Ethics approval was obtained from the International Diabetes Institute and Monash University, Melbourne. All participants consented to participate in the study.

Education level was ascertained by asking the question “Which of these describes the highest qualification you have received?” Education was categorised as secondary only (comprising those with a secondary school qualification), diploma (comprising nursing or teaching qualification, trade certificate or undergraduate diploma), and degree (comprising bachelor degree, post-graduate diploma or masters degree/doctorate). These categories are considered to represent hierarchical stages of education, each of which has important socioeconomic implications. [24]

Baseline and follow-up assessments followed a similar protocol. Data were collected by interviewer-administered questionnaires on age, sex and current use of antihypertensive and lipid-lowering medications. Self-reported CVD was ascertained by asking if participants had been told by a doctor or nurse that they had angina, coronary heart disease, or stroke. Smoking status was defined as 1) current daily smoker and 2) ex-smoker (now smoking less than daily for at least the last 3 months, but used to smoke daily) and non-smoker (never smoked tobacco daily) combined.

Blood pressure was measured using a Dinamap or a standard mercury sphygmanometer. [25] Height and weight were measured using standard methods, [25, 26] and BMI was calculated as weight (kg)/height (m)². [26] Fasting serum total cholesterol was
measured with an Olympus AU600 analyser (Olympus Optical, Tokyo, Japan) at a central laboratory.[27]

Classification of diabetes status has been described elsewhere.[27] Briefly, participants were classified as having known diabetes mellitus if they reported having physician-diagnosed diabetes mellitus and were either taking hypoglycaemic medication or had fasting plasma glucose (FPG) ≥ 7.0 mmol/L or a 2-hour plasma glucose (PG) ≥ 11.1 mmol/L. Participants not reporting diabetes mellitus but with FPG ≥ 7.0 mmol/L or 2-hour PG ≥ 11.1 mmol/L were classified as having newly diagnosed diabetes mellitus.

**Statistical analysis**

Analyses were conducted using sample weights to account for the sampling design of the study.[23] For continuous risk factor variables, the significance of any trend across educational categories was assessed using linear regression. For dichotomous variables, educational trends were assessed using logistic regression. Two-sided p values are presented, with p-values <0.05 regarded as significant.

For each individual risk factor at baseline, we created “not at risk” groups according to baseline measurement or use of prescription medication for that risk factor. The cut-point for being considered “not at risk” for hypertension was a baseline systolic blood pressure reading of <140 mmHg or a baseline diastolic blood pressure reading of <90mmHg, and not on antihypertensive medication. For cholesterol, the cut-point for being “not at risk” was <5.5mmol/l and not on cholesterol-lowering medication at baseline, and for BMI, the cut-point was <25kg/m². We used logistic regression to analyse the incidence of risk in those participants designated “not at risk” according to these cut-points. The odds ratios represent the odds of progressing from being “not at risk” at baseline to “at risk” at follow-up for that risk factor within each education category, relative to the highest educated group. Two models are presented; model 1 (unadjusted), and model 2 (adjusted for age as a continuous variable).

The cumulative number of risk factors was calculated by adding the numbers of those “at risk” for hypertension, cholesterol, BMI, diabetes and smoking using the cut-points described above. Logistic regression was used to examine the association between education and the likelihood of having a particular number of risk factors at baseline, and of increasing the number of risk
factors between baseline and follow-up. Tests of interaction between education and sex were significant for diabetes and therefore results of sex-stratified analyses are presented.

RESULTS

Baseline characteristics of study participants are shown in Table 1. Compared to the lower educated groups, those with degree level education were younger and less likely to have prior CVD. Inverse gradients were observed for most risk factors, reflecting a more adverse risk factor profile for those with lower education compared to those with higher education. No gradients were seen in diastolic blood pressure or cholesterol in men. The numbers of risk factors at baseline (out of a total number of 5) also differed according to education for both men and women. Those with higher education had fewer risk factors, while those with lower education had a greater number.

In men “not at risk” at baseline (Table 2), no educational gradients in systolic and diastolic blood pressure, cholesterol or BMI were observed. Conversely, among women “not at risk” inverse educational gradients were seen for the majority of risk factors. Positive gradients were seen in the proportions of both men and women who were “not at risk” for all risk factors, reflecting a greater likelihood of being “at risk” if in a lower education category.

Baseline characteristics were compared between participants who attended follow-up and those who did not (data not shown). Compared to those who did attend, non-attendees were on average 2 years older. Mean levels of systolic blood pressure, cholesterol, and BMI were similar between the two groups. The prevalence of smoking in male non-attendees according to educational category was 28%, 27% and 14% for secondary, diploma and degree-educated groups, respectively, and for female non-attendees was 23%, 20% and 14%, respectively. Diabetes prevalence in male non-attendees was 12%, 11% and 6% for secondary, diploma and degree-educated groups, respectively, and 10%, 6% and 4%, respectively for female non-attendees. The proportions in each education category who were eligible for follow-up but did not attend were 44% (n=1 628) with secondary education, 39% (n=1 551) with trade certificate or diploma, and 32% (n=545) with tertiary education. For men, the odds of not attending (when eligible) were 1.60 (95% CI 1.34-1.92) times greater for those with secondary compared to
tertiary education and 1.30 (95% CI 1.10-1.54) times greater for those with diploma compared to tertiary education. For women, the odds of not attending (when eligible) were 1.65 (95% CI 1.40-1.95) and 1.32 (95% CI 1.11-1.58), respectively.

**Incidence of risk factors**

We estimated the proportions of those who progressed to being “at risk” at follow-up according to the pre-defined cut-points described earlier (data not shown). In this unadjusted analysis, significant inverse educational gradients were seen for incident hypertension in women and for incident diabetes in both men and women. Among women, 14% of those with secondary education developed incident hypertension compared to 12% of women with diploma level and 9% of women with degree level education. For diabetes, 7% of men with secondary education developed incident diabetes compared to 3% of men with diploma level and 3% with degree level education. For women, the corresponding proportions were 4%, 3% and 1%, respectively.

The likelihood of progression to incident risk according to education level is shown in **Table 3**. In age-adjusted analyses, women with secondary education and those with diploma level education were more likely to progress to increased risk for BMI (age-adjusted OR compared to degree level education 1.57 (95% CI 1.06-2.31) and 1.72 (95% CI 1.25-2.37) respectively). Both men and women with secondary and diploma level education were also more likely to develop incident diabetes than were those with higher education (age-adjusted OR from degree level education 1.75 (95% CI 1.14-2.71) and 3.01 (95% CI 1.26-7.20) respectively).

**Change in smoking status**

The proportions of current smokers decreased between baseline and follow up in most educational groups, with the greatest decreases seen in those in the lowest education categories (data not shown). Among men with secondary education, there was a 3% decrease in smoking prevalence, a 2% decrease in those with a diploma, and a 1% decrease in those with tertiary education. Among women, the decreases in smoking prevalence were 2% each for those with secondary and diploma education, while among women with a degree, smoking prevalence increased by 2%. Participants with lower education were more likely to stop smoking between baseline and follow up than those with higher education, with age-adjusted
odds ratios for smoking cessation for the lowest compared to the highest education category 2.39 (95% CI 1.13-4.65) for men, and 3.83 (95% CI 1.46-10.02) for women.

**Total number of incident risk factors**

*Figure 1* shows the age-adjusted odds ratios for the likelihood of progressing to an increased number of risk factors at follow-up according to educational attainment. Among women, the likelihood of progressing from a lower to a higher number of risk factors over time followed an educational-based pattern, with those from the lowest educated group significantly more likely than those with a degree to progress from having two or less risk factors at baseline to having three or more at follow up (age-adjusted OR 2.04 (95% 1.32-3.14). For men, the corresponding odds ratio was 1.34 (95% CI 0.83-2.17).

**DISCUSSION**

**Overall findings**

This study found that lower education was associated with an increased probability of developing major CVD risk factors over a 5 year period. Lower SES was positively associated with the onset of overweight or obesity in women, and with the incidence of diabetes in both men and women. The likelihood of accumulating a higher number of risk factors between baseline and follow up was greater for lower compared to higher educated women.

**Strengths and limitations**

This contemporary study was undertaken on a large sample with accurate measurement of risk factors. However, there was a significant loss to follow-up, and so the sample may not be representative of the Australian population, limiting the generalisability of our findings. While non-attendees were more likely to be smokers, this was the case for all educational groups, and therefore would be unlikely to affect our findings concerning the relationship between education and smoking.

While we used one indicator only to measure SES, education is considered a robust measure, and has been more strongly associated with risk factor change than has either income or occupation. [28] Educational attainment is generally a stable indicator as it may be less likely to
change after adulthood than other indicators. [24] We repeated all analyses using occupation and income (data not shown) and found similar trends to those seen when using education, reinforcing the strong associations between social disadvantage and higher CVD risk.

We described risk using categories rather than continuous measurements. There is known to be a continuous relationship between risk factor levels and risk of disease; as a risk factor progressively increases, so too does the risk of developing CVD. [29] Our results may therefore not present the most comprehensive picture of risk accumulation.

**Individual risk factors**

SES has been shown previously to be a predictor of obesity, [14, 30] and is inversely associated with mean weight gain in some studies. [15, 31] Consistent with our findings, several longitudinal studies report that lower SES is associated with incident overweight or obesity over time. [32, 33] We also found that lower educated women of normal weight were more likely to progress to incident overweight or obesity than were those with higher education. This may be driven by the higher mean BMI observed at baseline among women with lower education. There was no educational gradient seen in mean BMI among men of normal weight, possibly accounting for the differences seen in our study between men and women. Gender-related differences in the social gradient in BMI gain have also been reported by others, although findings are conflicting. [34-37] Overweight or obesity has previously been associated with other risk factors such as diabetes and systolic blood pressure. [38] Therefore, the contribution of BMI to future socioeconomic gradients in CVD is potentially of great importance, particularly among women.

Consistent with our results, other studies also report that diabetes incidence is inversely associated with SES.[17-19, 39, 40] The greater incidence of diabetes among lower educated women in our study may be related to the co-existing increases seen in overweight and obesity. Several studies report an attenuation of the effect of SES on incident diabetes after adjusting for BMI, [18, 20, 39, 40] suggesting that obesity is an important mediator in this relationship. However, in our study, the social gradient in diabetes is much stronger than that seen for BMI implying that other factors may play a part.
There was no educational gradient seen in incident hypertension in either men or women. Few studies have examined incident hypertension according to SES, with most finding that education, income and neighbourhood are all predictors of onset of hypertension.[9, 11, 13, 28] The lack of a significant finding in our study may reflect small numbers of incident hypertension.

While there was no educational gradient seen in mean systolic blood pressure among men who were “not at risk” for hypertension at baseline, a gradient was apparent when the total population was examined (including those on anti-hypertensive medication). This may indicate educational differences in the treatment of hypertension among men.

We found the prevalence of smoking declined across all socioeconomic groups, apart from among women with degree level education. Similar to other studies, the pattern was one of a greater decrease among the lower educated. [41-43] These findings are likely to reflect secular trends in smoking due to the effect of public health policies such as increased tobacco taxation. This strategy is considered one of the most effective deterrents to smoking, and has been shown to be effective among lower SES groups in some settings including Australia. [44] Overall our findings are encouraging, as smoking has previously been shown to contribute to approximately 30% of the excess risk of CVD mortality among lower SES groups. [45]

Multiple risk factors

It is known that having more than one risk factor can accelerate the development of atherosclerosis and CVD mortality [46, 47], and that disadvantage is inversely associated with the number of risk factors present.[7] The pattern of a smaller number of risk factors in higher SES groups and a greater number in lower SES groups seen in our study has been seen by others; [7, 48] however, few studies have examined the accumulation of risk factors over time according to SES. [22, 48] One United States study, utilising 20 years of data from the National Health and Nutrition Epidemiologic Follow-up Study reported that education was associated with the accumulation of behavioural risk factors for CVD, namely smoking, alcohol and obesity. [22] Our results also showed that women with lower education were more likely than those with higher education to accumulate three or more risk factors over time, potentially contributing to continued socioeconomic gradients in CVD in the future.
Implications

The implications of our findings for the future burden of CVD among lower SES groups are significant. Not only do people with less education carry a greater burden of individual CVD risk factors, they are also more likely to progress to an overall increased risk. This suggests that current inequalities in CVD mortality and morbidity will continue into the future. These findings have important implications for health policy, promotion, and practice, and reinforce the need to direct intervention efforts towards reducing SES differences in chronic disease.

CONCLUSION

Our findings provide evidence for the strong association between SES and incident risk factors for CVD in a cohort of Australian men and women. Lower education was associated with an increased risk of developing both individual and accumulated CVD risk factors over a 5-year period. These findings suggest that educational inequalities in CVD will continue into the future. Interventions to reduce CVD should be tailored so that they are effective among lower socioeconomic groups.

Acknowledgements

The AusDiab study co-coordinated by the Baker IDI Heart and Diabetes Institute, gratefully acknowledges the generous support given by:

National Health and Medical Research Council (NHMRC grant 233200), Australian Government Department of Health and Ageing, Abbott Australasia Pty Ltd, Alphapharm Pty Ltd, AstraZeneca, Bristol-Myers Squibb, City Health Centre-Diabetes Service-Canberra, Department of Health and Community Services - Northern Territory, Department of Health and Human Services – Tasmania, Department of Health – New South Wales, Department of Health – Western Australia, Department of Health – South Australia, Department of Human Services – Victoria, Diabetes Australia, Diabetes Australia Northern Territory, Eli Lilly Australia, Estate of the Late Edward Wilson, GlaxoSmithKline, Jack Brockhoff Foundation, Janssen-Cilag,, Kidney Health Australia, Marian & FH Flack Trust, Menzies Research Institute, Merck Sharp & Dohme,
Novartis Pharmaceuticals, Novo Nordisk Pharmaceuticals, Pfizer Pty Ltd, Pratt Foundation, Queensland Health, Roche Diagnostics Australia, Royal Prince Alfred Hospital, Sydney, Sanofi Aventis, Sanofi Synthelabo. Also, for their invaluable contribution to the set-up and field activities of AusDiab, we are enormously grateful to A Allman, B Atkins, S Bennett, A Bonney, S Chadban, M de Courten, M Dalton, D Dunstan, T Dwyer, H Jahangir, D Jolley, D McCarty, A Meehan, N Meinig, S Murray, K O’Dea, K Polkinghorne, P Phillips, C Reid, A Stewart, R Tapp, H Taylor, T Whalen, F Wilson and Paul Zimmet.

**Competing interests**

None declared

**Funding**

This work was supported by the National Health and Medical Research Council (NHMRC) (ID No. 465352 for AB, ID No. 390109 for GT, ID No. 586623 for JS). AP is funded by a VicHealth Fellowship.
References


### TABLE 1: Baseline risk factor measurements and number of risk factors in 5,967 AusDiab men and women participants, by educational attainment

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Secondary (n=751)</td>
<td>Diploma (n=1,359)</td>
</tr>
<tr>
<td>Age mean (SD)</td>
<td>54.0 (13.0)</td>
<td>52.2 (12.7)</td>
</tr>
<tr>
<td>Prior CVD n (%)</td>
<td>83 (11%)</td>
<td>125 (9%)</td>
</tr>
<tr>
<td>Systolic BP mean (SD)</td>
<td>134.1 (16.8)</td>
<td>132.2 (16.5)</td>
</tr>
<tr>
<td>Diastolic BP mean (SD)</td>
<td>75.2 (10.6)</td>
<td>74.6 (10.7)</td>
</tr>
<tr>
<td>Total cholesterol mean (SD)</td>
<td>5.64 (1.03)</td>
<td>5.72 (1.05)</td>
</tr>
<tr>
<td>Body mass index mean (SD)</td>
<td>27.5 (4.1)</td>
<td>27.2 (3.9)</td>
</tr>
<tr>
<td>Fasting blood glucose mean (SD)</td>
<td>5.8 (1.3)</td>
<td>5.7 (1.0)</td>
</tr>
<tr>
<td>Diabetes prevalence n (%)</td>
<td>80 (11%)</td>
<td>107 (8%)</td>
</tr>
<tr>
<td>Smoking prevalence n (%)</td>
<td>135 (18%)</td>
<td>175 (13%)</td>
</tr>
<tr>
<td>Proportion with no risk factors* n (%)</td>
<td>49 (7%)</td>
<td>126 (9%)</td>
</tr>
<tr>
<td>Proportion with one risk factor* n (%)</td>
<td>177 (24%)</td>
<td>340 (25%)</td>
</tr>
<tr>
<td>Proportion with two risk factors* n (%)</td>
<td>262 (35%)</td>
<td>509 (37%)</td>
</tr>
<tr>
<td>Proportion with three risk factors* n (%)</td>
<td>197 (26%)</td>
<td>291 (21%)</td>
</tr>
<tr>
<td>Proportion with four risk factors* n (%)</td>
<td>62 (8%)</td>
<td>89 (7%)</td>
</tr>
<tr>
<td>Proportion with five risk factors* n (%)</td>
<td>4 (1%)</td>
<td>4 (0.3%)</td>
</tr>
</tbody>
</table>

*P for trend across education categories using linear regression. Abbreviations CVD=cardiovascular disease; BP=blood pressure

* Risk factors defined as follows: hypertension (systolic blood pressure ≥ 140mmHg or diastolic blood pressure ≥ 90mmHg or on blood pressure lowering medication); hypercholestaemia (total cholesterol ≥ 5.5mmol/l or on cholesterol lowering medication); overweight or obese (BMI, ≥ 25 kg/m²); diabetes (known or newly diagnosed); smoking (current smoker).
<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Highest level of education</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Secondary (n=751)</td>
<td>Diploma (n=1359)</td>
</tr>
<tr>
<td></td>
<td>Degree (n=614)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P for trend*</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>Proportion “not at risk” n (%)</td>
<td>445 (59%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>891 (66%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>448 (73%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Systolic BP</td>
<td>mean (SD) if “not at risk”</td>
<td>124.4 (9.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>123.5 (9.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>123.9 (9.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.37</td>
</tr>
<tr>
<td>Diastolic BP</td>
<td>mean (SD) if “not at risk”</td>
<td>71.5 (9.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71.2 (8.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71.8 (8.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.56</td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>Proportion “not at risk” n (%)</td>
<td>266 (35%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>488 (36%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>260 (42%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.007</td>
</tr>
<tr>
<td>Body mass index</td>
<td>mean (SD) if “not at risk”</td>
<td>4.73 (0.52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.81 (0.48)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.76 (0.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Proportion “not at risk” n (%)</td>
<td>207 (28%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>394 (29%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>210 (34%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td>Fasting blood glucose</td>
<td>mean (SD) if “not at risk”</td>
<td>23.0 (1.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23.0 (1.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.9 (1.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.26</td>
</tr>
<tr>
<td>Smoking</td>
<td>Proportion “not at risk” n (%)</td>
<td>621 (82%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1184 (87%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>569 (93%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Not at risk defined for each risk factor as follows: For hypertension, systolic blood pressure < 140mmHg or diastolic blood pressure < 90mmHg & not on blood pressure lowering medication; for cholesterol < 5.5mmol/l & not on treatment; BMI, < 25 kg/m$^2$ or less; for diabetes, no known or newly diagnosed diabetes; for fasting blood glucose ≤6.0mmol/L; for smoking former or non-smoker

*P for trend across education categories using linear regression: Abbreviations; BP=blood pressure; SD=standard deviation
TABLE 3: Risk of incident risk factors at 5-year follow-up in men and women AusDiab participants according to educational attainment

<table>
<thead>
<tr>
<th></th>
<th>Men Model 1 - unadjusted</th>
<th>Men Model 2 – adjusted for baseline age</th>
<th>Women Model 1 - unadjusted</th>
<th>Women Model 2 – adjusted for baseline age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Secondary</td>
<td>Diploma</td>
<td>Degree</td>
<td>Secondary</td>
</tr>
<tr>
<td></td>
<td>odds ratio (95% CI)</td>
<td>odds ratio (95% CI)</td>
<td>ref</td>
<td>odds ratio (95% CI)</td>
</tr>
<tr>
<td>Hypertension: &quot;Not at risk* at baseline to at risk* at follow-up</td>
<td>1.09 (0.83, 1.41)</td>
<td>1.08 (0.84, 1.38)</td>
<td>1.00</td>
<td>0.93 (0.70, 1.24)</td>
</tr>
<tr>
<td>Total cholesterol: &quot;Not at risk* at baseline to at risk * at follow-up</td>
<td>0.97 (0.61, 1.55)</td>
<td>0.81 (0.52, 1.25)</td>
<td>1.00</td>
<td>0.94 (0.59, 1.49)</td>
</tr>
<tr>
<td>Body mass index: &quot;Not at risk* at baseline to at risk* at follow-up</td>
<td>0.93 (0.60, 1.46)</td>
<td>1.14 (0.75, 1.71)</td>
<td>1.00</td>
<td>1.03 (0.65, 1.64)</td>
</tr>
<tr>
<td>No diabetes at baseline to diabetes at follow-up</td>
<td>2.05 (1.32, 3.18)</td>
<td>0.88 (0.47, 1.65)</td>
<td>1.00</td>
<td>1.75 (1.14, 2.71)</td>
</tr>
</tbody>
</table>

*Not at risk defined for each risk factor as follows: : For hypertension, systolic blood pressure < 140mmHg or diastolic blood pressure < 90mmHg & not on blood pressure lowering medication; for cholesterol < 5.5mmol/l & not on treatment; BMI, < 25 kg/m² or less; for diabetes, no known or newly diagnosed diabetes

* At risk defined for each risk factor as follows: For hypertension, systolic blood pressure ≥ 140mmHg or diastolic blood pressure ≥ 90mmHg or on blood pressure lowering medication; for cholesterol, ≥ 5.5mmol/l or on cholesterol lowering medication; for BMI, ≥ 25 kg/m²; for diabetes, known or newly diagnosed diabetes.

Abbreviations CI = confidence intervals
Figure 1: Logistic regression analysis of the association between educational status and change from less than or equal to 2 risk factors* at baseline (2000) to greater than 2 risk factors at follow up (2005)**

* Risk factors defined as follows: hypertension (systolic blood pressure ≥ 140mmHg or diastolic blood pressure ≥ 90mmHg or on blood pressure lowering medication); hypercholesteraemia (total cholesterol ≥ 5.5mmol/l or on cholesterol lowering medication); overweight or obese (BMI, ≥ 25 kg/m²); diabetes (known or newly diagnosed); smoking (current smoker).

** Adjusted for age
The findings from the first analysis presented in this chapter highlighted that female smokers had a more atherogenic lipid profile compared to female non-smokers, and that this association was independent of lifestyle-related factors including alcohol intake, physical inactivity, body mass index, and dietary patterns. Due the sample size, we were unable to stratify the analysis by a measure of SES. However, lower SES is known to be associated with higher smoking rates, and as shown previously, smoking is major contributing risk factor to the social gradient in CVD. Thus the relevance of these findings for this thesis are that treatment of lipoproteins among smokers may require more than management of lifestyle factors, and because of their higher rates of smoking, aggressive treatment of lipids among lower SES groups may be warranted.

The second analysis was undertaken using data from the AusDiab cohort study, and found that lower education was associated with the incidence of overweight/obesity in women, and of diabetes in both men and women over a 5-year period. Lower education was also associated both with a greater number of risk factors at baseline for men and women, and with a higher likelihood of developing a greater number of incident risk factors over time in women. Thus findings from this analysis provide further evidence for the role of SES in the development of gradients in CVD. These findings also highlight the importance of targeting obesity and diabetes prevention strategies towards lower SES groups in order to reduce future social gradients in CVD.

The second analysis in this chapter used longitudinal data to explore individual-level changes in risk factors over time, thus making a valuable contribution to the health inequalities literature. Because cohort or longitudinal data allows us to examine changes in risk factors or disease status at an individual level, it means that the impact of ageing, treatment effects and secular trends on socioeconomic gradients in risk factors can also be explored. This is particularly useful in analysis of the relationship between individual SES and CVD risk as this relationship is likely to be a dynamic one which changes over time. The next chapter presents two different methods by which existing cohort studies can be adapted to be more applicable for use in health inequalities research.
CHAPTER 6

Unlocking the value of cohort studies for health inequalities research

6.1 Chapter overview

Prospective cohort studies potentially offer a wealth of data for health inequalities research. In particular, the ability to follow up individuals over time means that pathways between the social determinants of health and health outcomes can be explored over the lifecourse. In addition, the temporal sequencing of exposures and outcomes can be described, providing evidence for cause and effect relationships. However, within many countries including Australia, there is limited availability of longitudinal data to examine health inequalities, and few studies have collected data on socioeconomic exposures or analysed CVD outcomes according to SES.

This chapter discusses two methods by which existing longitudinal studies can be utilised more effectively to examine associations between SES and health outcomes. In chapter 5, we showed that lower educated subjects in a cohort study were less likely to continue participation than higher educated subjects, meaning they were lost from analyses of any health outcomes. The first paper in this chapter therefore addresses whether electronic record linkage between cohort studies and hospital administrative datasets is a viable method for increasing passive follow-up of study participants. We describe the methodology for a pilot record linkage of the MCCS with the Victorian Admitted Episodes Dataset (VAED) without using names and addresses.

The second analysis in this chapter examines whether baseline addresses from an existing cohort study can be retrospectively geocoded. Geocoding is a first step to obtaining data on environmental and neighbourhood exposures, and as discussed in chapter 2, these characteristics may be important contributors to health inequalities in disease. This paper describes the steps
involved in retrospective geocoding of baseline (1990-1994) addresses in the MCCS, and the subsequent assigning of a time-appropriate measure of area disadvantage to those addresses. This data was used to describe fatal CVD events occurring in the cohort by quintile of area disadvantage.
Declaration for Thesis Chapter 6

Declaration by candidate

In the case of chapter 6, the nature and extent of my contribution to the work 'Validation of de-identified record linkage to ascertain hospital admissions in a cohort study' was the following:

<table>
<thead>
<tr>
<th>Nature of contribution</th>
<th>Extent of contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical literature review and development of the research questions, ethics applications for the project, data preparation including collection of data on non-fatal CVD events, from medical records and coordination of adjudication of these events, analysis design, statistical analysis, interpretation of results including manual review of linked cases, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
<td>80</td>
</tr>
</tbody>
</table>

The following co-authors contributed to the work.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Andrew M. Tonkin</td>
<td>Grants used to fund research, facilitated development of project and research questions, interpretation of results, commented on manuscript drafts,</td>
</tr>
<tr>
<td>2 Helen Kelsall</td>
<td>Facilitated development of project including ethics approvals, assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>3 Vijaya Sundararajan</td>
<td>Commented on data, assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>4 Dallas R English</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, facilitated development of project, assisted with design of analysis, interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>5 Rory Wolfe</td>
<td>Commented on data, assisted with design of analysis, commented on manuscript drafts</td>
</tr>
<tr>
<td>6 Gavin Turrell</td>
<td>Interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>7 Graham G Giles</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>8 Anna Peeters</td>
<td>Assisted with design of analysis and interpretation of results, commented on manuscript drafts</td>
</tr>
</tbody>
</table>

Candidate’s Signature   Date 2011/10
Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate’s contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

**Location(s)**

| Department of Epidemiology & Preventive Medicine, Alfred Campus, Monash University |

[Please note that the location(s) must be institutional in nature, and should be indicated here as a department, centre or institute, with specific campus identification where relevant.]

| Signature 1 | 8.12.2010 |
| Signature 2 | 13.12.2010 |
| Signature 3 | 5.11.2010 |
| Signature 4 | 13.12.2010 |
| Signature 5 | 25.11.2010 |
| Signature 6 | |
| Signature 7 | 15.12.2010 |
| Signature 8 | |

---------------------------------------------
VALIDATION OF DE-IDENTIFIED RECORD LINKAGE TO ASCERTAIN HOSPITAL ADMISSIONS IN A COHORT STUDY

Alison Beauchamp* BHSc MPH; Andrew Tonkin1 MBBS MD FRACP FCSANZ; Helen Kelsall MB BS, MPH, MHlthSc, PhD, FAFPHM; Vijaya Sundararajan2, 3 BA(Hons) MD MPH FACP; Dallas R English 4, 5 BSc MSc PhD; Lalitha Sundaresan2 MBBS, MPH; Rory Wolfe1 BSc PhD; Gavin Turrell6 BA PhD; Graham G Giles1, 4, 5 BSc MSc PhD; Anna Peeters1 BSc(Hons) PhD;

1. Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia,

2. Victorian Data Linkages, Department of Health, Melbourne, Australia

3. Department of Medicine, Monash Medical Centre, Monash University, Melbourne, Australia

4. Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia

5. Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, University of Melbourne, Australia

6. School of Public Health, Queensland University of Technology, Brisbane, Australia

* Corresponding author

Email addresses:
AB: Alison.Beauchamp@monash.edu
AT: Andrew.Tonkin@monash.edu
HK: Helen.Kelsall@monash.edu
VS: Vijaya.Sundararajan@health.vic.gov.au
DE: Dallas.English@cancervic.org.au
LS: Lalitha.Sundaresan@health.vic.gov.au
RW: Rory.Wolfe@monash.edu
GT: g.turrell@qut.edu.au
GG: Graham.Giles@cancervic.org.au
AP: Anna.Peeters@monash.edu
ABSTRACT

Background Cohort studies can provide valuable evidence of cause and effect relationships but are subject to loss of participants over time, limiting the validity of findings. Computerised record linkage offers a passive and ongoing method of obtaining health outcomes from existing routinely collected data sources. However, the quality of record linkage is reliant upon the availability and accuracy of common identifying variables. We sought to develop and validate a method for linking a cohort study to a state-wide hospital admissions dataset with limited availability of unique identifying variables.

Methods A sample of 2000 participants from a cohort study (n=41 514) was linked to a state-wide hospitalisations dataset in Victoria, Australia using the national health insurance (Medicare) number and demographic data as identifying variables. Availability of the health insurance number was limited in both datasets; therefore linkage was undertaken both with and without use of this number and agreement tested between both algorithms. Sensitivity was calculated for a sub-sample of 101 participants with a hospital admission confirmed by medical record review.

Results Of the 2000 study participants, 85% were found to have a record in the hospitalisations dataset when the national health insurance number and sex were used as linkage variables and 92% when demographic details only were used. When agreement between the two methods was tested the disagreement fraction was 9%, mainly due to “false positive” links when demographic details only were used. A final algorithm that used multiple combinations of identifying variables resulted in a match proportion of 87%. Sensitivity of this final linkage was 95%.

Conclusions High quality record linkage of cohort data with a hospitalisations dataset that has limited identifiers can be achieved using combinations of a national health insurance number and demographic data as identifying variables.
BACKGROUND

Cohort studies are a valuable source of information for epidemiological research, primarily because information about potential risk factors is collected before the outcomes of interest occur (1). For example, the long-standing Framingham cohort study was critical in demonstrating the relationship between certain risk factors and the development of cardiovascular disease (CVD) events during follow-up (2-4).

The quality of evidence from cohort studies relies on complete and accurate ascertainment of outcomes such as myocardial infarction or stroke. Information about these and other health outcomes can be collected in a variety of ways, including medical record review and self-report from participants. While the former is considered the "gold standard" (5-6), it is particularly resource intensive for large cohorts. In addition, over longer periods of time, medical records may be difficult to locate or may be destroyed according to legislative requirements. Self-report from participants has been shown to have varying accuracy (5-10), and is subject to "loss to follow-up", an inherent problem and source of bias in cohort studies. Specific groups at risk of loss to follow-up include those of lower socioeconomic status and those with poorer health, often the groups of major interest to epidemiological research.

An alternative method of obtaining health outcome data for cohort studies is computerised record linkage (11-12). This is the process of using common identifiers to link the cohort data with health services administrative or other datasets, for example to identify whether study participants have been admitted to hospital, and for which medical conditions (9, 13). Record linkage enables the optimal use of data from cohort studies because it permits "passive" follow-up, so that even if participants have been lost to follow-up from the study, information on specific health outcomes can still be obtained (11).

There are two main methods for computerised record linkage, probabilistic and deterministic. Probabilistic record linkage links records based on the statistical probability that common identifiers belong to the same person (11). Deterministic linkage links two records based on complete agreement between the common identifiers (12). Deterministic linkage is particularly suited to linkage of individual level data where accuracy is important and where data quality within the various datasets is high (14-15).

Linkage of cohort data with health services administrative datasets has been routine in several countries including the United Kingdom (13, 16), Canada (17-18), and Sweden (19) for many
years. It is important to note that for many of these studies, particularly those that use deterministic methods, linkage is based on unique identifiers such as social security number or health record number that are used by individuals throughout a lifetime and across the spectrum of health and social services (11, 20-21). However, it is not always possible to use unique identifiers, either because they are not fully available within the datasets, or because legislation requires that anonymity of records is maintained. In these situations, combinations of non-unique identifiers must be used such as date of birth, sex, and postal code. A key issue for researchers then becomes one of accuracy; that is, whether the combination of identifying variables used is sufficiently precise to identify the correct person, but not so broad as to incorrectly match to another person who has the same demographic data. In addition, it is essential to make allowance for mistakes in data entry particularly when using deterministic methods of linkage in which records are linked only if they match exactly (22). Therefore, when developing a method for linking two or more datasets with limited identifying information, validation of the linkage is vital.

We planned to link a cohort study with a state-wide hospital admissions dataset in order to obtain data on incident CVD events occurring in the cohort during 19 years of follow-up. The hospitalisations dataset did not contain names and addresses, and there was limited availability of unique identifiers in either dataset. We therefore sought to determine, in a subgroup of the cohort, the most accurate combinations of identifying variables with and without use of the national health insurance number, with the overall aim of establishing a stepwise deterministic algorithm for linking the two datasets. We also aimed to test the sensitivity of the linkage for correctly identifying that a true hospital admission event had occurred.

METHODS

Data Sources

The Victorian Admitted Episodes Dataset (VAED) is held by the Victorian Department of Health (DH), and includes information on all private and public hospital admissions in Victoria. While full names and addresses are not included in the dataset, approximately 81% of records in the VAED include Medicare card number, the national health insurance number allocated to all Australians. Medicare card numbers are unique to a family only and individual family members are identified by the Medicare suffix, comprising the first three letters of their given (first) name. The VAED is
episode of care based, and the DH has linked these episodes using identifiers such as hospital record number (if episodes occurred in the same hospital) and the first 8 digits of Medicare card number, as well as identifiers such as date of birth, gender, postal code, country of birth and the first three digits of first or middle name (if episodes occurred in different hospitals). The methods for this internal linkage have been described elsewhere (23). Data between 31 July 1996 and 31 December 2008 from this “internally linked VAED” were used for the current study. The VAED has previously been linked to a number of different datasets and registries including a transport accident dataset (22), a cardiac rehabilitation dataset (24), and a cardiothoracic elective surgery information system (25).

The Melbourne Collaborative Cohort Study (MCCS) is a prospective study of 41,514 subjects aged 27 to 80 years, recruited between 1990 and 1994. Details of the design, recruitment, and study procedures have been published elsewhere (26). In brief, subjects were volunteers from metropolitan Melbourne, recruited using electoral rolls, community centres and churches. Between 2003 and 2008, attempts were made to re-interview all surviving participants; 28,240 participants were re-interviewed. At this interview, participants were asked to provide their Medicare card number and details on hospitalisations for cardiovascular and other diseases. This linkage study was undertaken on a random sample of 2000 participants who had been re-interviewed. The entire sample had Medicare details available. In addition, the sample was stratified to include 67% with a self-reported hospital admission for a cardiovascular event at re-interview to ensure at least this number of participants with a record in the VAED. CVD was chosen as the event of interest to allow us to test the sensitivity of linkage using a clearly defined outcome.

The pilot sample of 2000 also included 101 participants with a confirmed hospital admission for myocardial infarction (AMI) or stroke between 1 July 1996 and the time of their re-interview. These 101 participants were used to test the sensitivity of linkage, with hospitalisation and diagnosis confirmed as follows: At re-interview all participants were asked “Has a doctor or nurse ever told you that you had a heart attack or myocardial infarction or stroke?” Hospital name and year of admission were also asked. From those who responded in the affirmative and had not reported a prior history of CVD at study baseline, a random sample of 400 participants was selected. We excluded 193 of these for the following reasons: hospital name not given or no medical records identified for that participant (n=84), hospital medical records destroyed (n=67), interstate hospitalisation (n=21), and no CVD event identified in the medical record (n=21).
Data considered to relate to the self-reported CVD event was obtained from the medical records for 207 participants and coded by expert panels of neurologists and cardiologists. From the original sample of 400, we identified 124 participants with a confirmed admission for AMI or stroke occurring between baseline and the time of their re-interview. We excluded a further 23 of these with an event prior to 1 July 1996 (the commencement of the VAED), leaving a total of 101 participants with a confirmed AMI or stroke. These 101 were included in the pilot sample of 2000, flagged as “confirmed admission.”

The study protocol was approved by Human Research Ethics Committees at Monash University, The Cancer Council Victoria and the Victorian Department of Health. All subjects provided written informed consent, including for linkage to the VAED.

**Identifying variables used for linkage**

Identifiers common to both the study sample and the VAED were the national health insurance (Medicare) card number and suffix (i.e., a given or middle name abbreviated to first 3 letters), date of birth, postcode, sex and country of birth. For linkage purposes, the first 8 numbers of the Medicare card number only were used (Medicare8).

**Linkage Method**

Medicare card number is not fully available in either the VAED or the MCCS. We therefore undertook linkage using two algorithms in order to assess the effect of missing Medicare numbers. The first algorithm used combinations of Medicare and demographic details as linkage variables, and the second used demographic details only. For each algorithm, identifiers from each dataset were grouped into several combinations and matched in a stepwise deterministic strategy, using multiple iterations. Matches in each iteration were accepted only if the identifiers were identical between the two datasets. Records that matched in each iteration were removed from the source datasets for subsequent iterations.

**Agreement between linkage with and without Medicare number**

Agreement between the two linkage algorithms (i.e. Medicare card number plus demographic details versus demographic details only) was assessed based on the assumption that if the two
algorithms worked equally well, then each would link MCCS participants to the same record in the VAED.

*Sensitivity of linkage - Participants with confirmed admissions for AMI or stroke*

The sensitivity of the linkage process was assessed using the 101 MCCS participants with a ‘confirmed admission’ for AMI or stroke. Information was extracted from the VAED relating to any hospital admissions they may have had between 1 July 1996 and 31 December 2008. If a participant’s confirmed hospitalisation matched a record from the VAED by hospital name, dates of admission and discharge (within ten days), then the episode was considered to have been correctly identified by the linkage.

Sensitivity was calculated as the number of confirmed admissions that were correctly identified in the VAED divided by the total number of confirmed admissions.

**RESULTS**

*Availability of identifying variables*

Medicare details were available for 100% of the study sample, and 80% of the VAED records. There was 100% availability of demographic variables from both datasets.

*Linkage Method*

*Tables 1 and 2* describe the combinations of linkage variables used and the number of study participants matched during each iteration, both with and without use of Medicare card number. When the Medicare number was used, 1865 of the 2000 (93%) records were matched to a record in the VAED *(Table 1)*. The first 4 iterations of this algorithm used various combinations of Medicare number, suffix and sex only, and resulted in 1702 (85%) matches between the study sample and the VAED. When the Medicare number was not used, combinations of date of birth, sex, country of birth and postcode yielded matches for 1843 records (92%) *(Table 2)*.
### Table 1: Linkage using Medicare number and suffix (V1)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Linkage Variables used in V1</th>
<th>Records Linked</th>
<th>Records remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Medicare8† + Medsuf1‡</td>
<td>1,667</td>
<td>333</td>
</tr>
<tr>
<td>2</td>
<td>Medicare8+first and second letters of Medsuf1+Sex</td>
<td>7</td>
<td>326</td>
</tr>
<tr>
<td>3</td>
<td>Medicare8+ second and third letters of Medsuf1+Sex</td>
<td>2</td>
<td>324</td>
</tr>
<tr>
<td>4</td>
<td>Medicare8+first and third letters of Medsuf1+Sex</td>
<td>0</td>
<td>324</td>
</tr>
<tr>
<td>5</td>
<td>Medicare8+Yearbirth+Sex</td>
<td>26</td>
<td>298</td>
</tr>
<tr>
<td>6</td>
<td>Medsuf1+Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode1</td>
<td>32</td>
<td>266</td>
</tr>
<tr>
<td>7</td>
<td>Medsuf1+Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode2</td>
<td>4</td>
<td>262</td>
</tr>
<tr>
<td>8</td>
<td>Medsuf1+Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode3</td>
<td>2</td>
<td>260</td>
</tr>
<tr>
<td>9</td>
<td>Medsuf1+Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode4</td>
<td>1</td>
<td>259</td>
</tr>
<tr>
<td>10</td>
<td>Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode1</td>
<td>124</td>
<td>135</td>
</tr>
<tr>
<td><strong>Total Matched</strong></td>
<td></td>
<td>1865</td>
<td>135</td>
</tr>
</tbody>
</table>

† First 8 digits of the Medicare number; ‡ First 3 letters of the first name

### Table 2: Linkage not using Medicare number and suffix (V2)

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Linkage Variables used in V2</th>
<th>Records Linked</th>
<th>Records remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode1</td>
<td>1,599</td>
<td>410</td>
</tr>
<tr>
<td>2</td>
<td>Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode2</td>
<td>193</td>
<td>208</td>
</tr>
<tr>
<td>3</td>
<td>Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode3</td>
<td>39</td>
<td>169</td>
</tr>
<tr>
<td>4</td>
<td>Yearbirth+Monthbirth+Daybirth+Sex+Country of Birth+Postcode4</td>
<td>12</td>
<td>157</td>
</tr>
<tr>
<td><strong>Total Matched</strong></td>
<td></td>
<td>1843</td>
<td>157</td>
</tr>
</tbody>
</table>

### Agreement between linkage with and without Medicare number

The two linkage methods matched 1651 MCCS records to records with the same VAED linkage ID (the computer generated number given to groups of hospitalisations thought to belong to the same individual as a result of the linkage process) (**Figure 1**). Another 170 MCCS participants were matched to different VAED IDs, implying that the two linkage methods had matched these participants to different records in the VAED dataset. This represents an 8.5% disagreement fraction. When the Medicare card number was not used (demographic variables only), all 170 records in each dataset matched exactly for every linkage variable. When the Medicare number and demographic details were used, 128 of the 170 matched exactly for every linkage variable. Assuming that these 128 matches were correct (as they had matched completely on the most distinctive combination of variables available), it is likely that at least 128 of the 170 from the
linkage without Medicare numbers were ‘false positive’ matches, due to there being more than one person with a VAED record from the same postcode, born on the same date, and of the same sex and country of birth. Of the 42 remaining records identified, all matched on Medicare number and suffix. Most differences were seen in date of birth, which varied between 2-10 days and up to 40 years, suggesting possible errors in data entry or that the MCCS participant had been matched to another family member whose given name began with the same three letters.

There were 44 MCCS participants who were matched only in V1 (Figure 1). All of these matched completely on Medicare details, but all had at least one demographic variable that was unmatched; the most likely explanation for not being matched in V2. There were 22 MCCS participants who were matched only in V2. This was possibly because their Medicare card number had either been entered into the VAED or MCCS datasets incorrectly, or they did not have a Medicare card number in the VAED.

Figure 1: Agreement between linkage with and without Medicare number

![Diagram showing agreement between linkage with and without Medicare number]

**Linkage without Medicare – decreasing the number of ‘mismatches’**

We sought to decrease the 8.5% disagreement fraction between linkage with and without Medicare details by adding the Medicare suffix to the linkage algorithm containing demographic variables only. This field is available for approximately 80% of records in the VAED and all MCCS participants. Applying this algorithm, 1633 out of 2000 (82%) records were matched. When
compared with the linkage involving Medicare card number and demographic variables, there were 1620 participants who had the same VAED ID. Only 13 records, compared with 170 shown in Figure 1, had different VAED IDs on the two linkages. There were fewer matches than for the linkage based on demographic variables only (1633 compared to 1843), suggesting that adding the first three letters of the given name reduced the number of “false positives”.

**Final algorithm**

Based on the above findings, we developed a final algorithm that will be used to link the entire study cohort of 41 514 participants (Table 3). The algorithm is grouped into three stages. The first stage (Medicare card number and Medicare suffix) uses combinations of the national health insurance card number and first 3 letters of the first name plus all demographic details. The first iteration in this stage is assumed to provide the most correct match possible. Subsequent iterations of this first stage use variations of the Medicare suffix and also drop one variable at a time to allow for errors in data entry, with a maximum of variation in two variables allowed at one time. The second stage (Medicare card number) is used for matching records with variations in first names, for example where nicknames or middle names have been used. The final stage (Medicare suffix only) aims to match those records that do not have Medicare card number available in one or more dataset. Again, this group allows for variation in no more than two variables at a time. Running this algorithm resulted in 1740 (87%) of 2000 records linked.

**Testing sensitivity - MCCS participants with 'confirmed admission’**

In linkage undertaken using the final algorithm, 98 of 101 (97%) participants with a confirmed hospital admission for AMI or stroke were linked to a record in the VAED (Table 4). Of the 98 who were linked, the data we had obtained from the hospital medical record did not match the VAED data for admission and discharge dates and hospital name in 4 cases. Overall, this represents a sensitivity of 93% (94/101; 95% confidence interval 86% to 97%) for the VAED to correctly identify that a hospital admission for AMI or stroke has occurred when linkage was undertaken using the final algorithm.
Table 3: Suggested linkage cycles/iterations

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Medicare8 group</th>
<th>Records linked</th>
<th>Records remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>Medicare8+Medsuf1+Yearbirth+Monthbirth+Daybirth+Sex+COB+Postcodes1-4</td>
<td>1594</td>
<td>406</td>
</tr>
<tr>
<td>5-16</td>
<td>Repeat steps 1-4 using 3 different variations of Medsuf1</td>
<td>8</td>
<td>398</td>
</tr>
<tr>
<td>17</td>
<td>Repeat step 1 dropping postcode</td>
<td>23</td>
<td>375</td>
</tr>
<tr>
<td>18-21</td>
<td>Repeat steps 1-4 dropping COB</td>
<td>7</td>
<td>368</td>
</tr>
<tr>
<td>22-25</td>
<td>Repeat steps 1-4 dropping sex</td>
<td>1</td>
<td>367</td>
</tr>
<tr>
<td>26-29</td>
<td>Repeat steps 1-4 dropping day of birth</td>
<td>19</td>
<td>348</td>
</tr>
<tr>
<td>30-33</td>
<td>Repeat steps 1-4 dropping month of birth</td>
<td>2</td>
<td>346</td>
</tr>
<tr>
<td>34-37</td>
<td>Repeat steps 1-4 dropping year of birth</td>
<td>17</td>
<td>329</td>
</tr>
<tr>
<td>38-100</td>
<td>Repeat steps 5-16, dropping 1 variable</td>
<td>0</td>
<td>329</td>
</tr>
<tr>
<td>101-141</td>
<td>Repeat steps 1-4 dropping two demographic variables</td>
<td>4</td>
<td>325</td>
</tr>
</tbody>
</table>

Medicare8 group

| 142-145 | Medicare8+Yearbirth+Monthbirth+Daybirth+Sex+COB+Postcodes1-4 | 24 | 301 |
| 146-166 | Repeat steps 142-145 dropping 1 variable | 2 | 299 |

Medsuf1 Group

| 167-170 | Medsuf1+Yearbirth+Monthbirth+Daybirth+Sex+COB+Postcodes1-4 | 39 | 260 |

Total records linked = 1740/2000 = 87%

1 First 8 digits of the Medicare number; 2 First 3 letters of the first name

Table 4: Number of VAED matches to MCCS participants with confirmed CVD event

<table>
<thead>
<tr>
<th></th>
<th>Linkage using Medicare details (V1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of true CVD events</td>
<td>101</td>
</tr>
<tr>
<td>Number of admissions identified by VAED using NEWID</td>
<td>98</td>
</tr>
<tr>
<td>Number of CVD events correctly identified by VAED using NEWID</td>
<td>94</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>94/101 = 93%</td>
</tr>
<tr>
<td>Number of incorrect matches in VAED</td>
<td>4</td>
</tr>
<tr>
<td>Non-matched rate</td>
<td>4/101 = 4%</td>
</tr>
</tbody>
</table>

DISCUSSION

Overall findings

This study aimed to develop the most appropriate method for linking a large cohort study with a state-wide hospital admissions dataset, with limited availability of unique identifying variables in either dataset. We found that linkage using demographic variables only had the potential to create “false positive” links, which were reduced by adding the first three letters of the given
(first) name to the linkage variables. The number of records linked was highest when using combinations of the national health insurance number, the first three letters of the first name, date of birth, sex, country of birth and postcode. Accordingly, we have developed a stepwise algorithm using combinations of identifying variables that will provide the greatest accuracy in deterministic linkage of cohorts such as the MCCS to administrative health datasets. Sensitivity of this linkage algorithm to correctly identify that a hospital admission had occurred was 93%.

**Effect of errors in data entry**

While health record numbers or other unique identifier details in combination with demographic details are likely to provide the most accurate linkage for health-related datasets, they may be subject to data entry errors. We found for example, that data entry errors for day of birth ranged from between 2 to 10 days. There are few validation studies of deterministic record linkage of anonymous or de-identified datasets. One study that linked the New York State AIDS registry and a hospital discharge file using date of birth, sex, admission dates and hospital record number found 82-85% accuracy, assessed by medical record review and manual verification (27). In that study, errors in data entry accounted for most of the missed links. A study conducted in Indiana linking hospital admissions registries with a death registry using social security number (SSN) also found that errors in SSN were significant (20). In our final linkage cycle, the impact of data entry error is reduced by varying the combinations of variables used.

**Effect of missing national health insurance numbers and other data.**

We found that performing linkage in which the national health insurance card number and abbreviated given name were “missing” from the study data increased the probability of matching to the wrong record in the hospital administrative dataset. This may lead to misclassification of health outcomes, likely to be non-differential in nature as the reason for misclassification will generally be independent of the exposure. While this non-differential disease misclassification may have implications for subsequent analyses using health outcomes data obtained from linkage, it is most likely to bias associations between exposures and outcomes towards the null (28). Our final linkage algorithm allows for missing health insurance numbers in either dataset by including combinations of demographic variables and abbreviated first name only. It does not allow for the fact that a small number of records in the administrative dataset do not include an abbreviated first name and will consequently remain unlinked.
Use of names and other identifying variables in linkage

Deterministic record linkage uses an "exact" approach to matching, and is generally undertaken using unique health identifiers, often combined with demographic details and names. Grannis found that the most accurate combination of identifiers for deterministic linkage of hospital admissions and death registries was social security number, compressed first name, month of birth and gender(20). We also found that a similar combination was the most suitable for linkage of two health-related datasets. It is probable that linkage accuracy would be enhanced were full names also available for all datasets being linked, particularly when health or other identity numbers are assigned to families rather than individuals, as is the case with Australian Medicare card numbers. Others have also found that use of names significantly increases accuracy of deterministic linkage (15, 29).

Limitations

This study did not assess specificity, which is the ability of linkage to show that participants with no hospitalisations are correctly identified as such, or that participants are not linked to the wrong record. We were also unable to assess the validity of the 'linked VAED'. This dataset had previously been internally linked using Medicare card numbers, hospital record numbers or demographic data. Of the 101 'confirmed admissions', 3 did not link to a record in the VAED, indicating either errors in data entry or errors in the 'linked VAED'. However, the 'linked VAED' has previously been shown to be accurate (23), and sensitivity of linkage to correctly identify a hospital admission was relatively high at 96%, indicating that it is accurate for at least hospital name and dates of admission. A further limitation is that the Medicare card number may alter with change in family circumstances, such as divorce or marriage. The number of participants in the sample in whom this occurred is unknown, although is likely to be small given the high number of matches using Medicare number.

Our study was conducted in Melbourne, Australia using the hospital admissions dataset for the state of Victoria. It may therefore have limited applicability for other settings, although the methods will be informative. Our findings reinforce the importance of accurate data to obtain the best results from deterministic record linkage, and demonstrate that linkage is still achievable in the setting of limited availability of unique identifiers.
**Implications**

This pilot study has significantly increased the potential of a cohort study to determine health outcomes related to hospitalisation, even with limited availability of unique identifiers. The methods described may therefore be applicable to other settings in which linkage is undertaken using limited identifiers.

**CONCLUSION**

Our findings suggest that record linkage with hospital admissions datasets that have limited identifiers offers an opportunity to identify health outcomes in a cohort study. There are specific issues which affect the quality of linkage, and may have implications for use of data obtained from linkage. Including names and addresses in administrative datasets used for record linkage would significantly improve the quality of such a valuable research tool.

**Competing interests**

None declared

**Authors' contributions**

AB drafted the manuscript, participated in the study design, manually reviewed outcomes of linkage, undertook medical record review, and performed statistical analysis. AT participated in the study design and critically reviewed the manuscript. HK participated in the study design, assisted with interpretation of results and helped to draft the manuscript. VS assisted with the development of the linkage protocol, assisted with interpretation of results and helped to draft the manuscript. DE is an original investigator on the MCCS, participated in the study design and critically reviewed the manuscript. LS assisted with the development of the linkage protocol and undertook the linkage. RW assisted with statistical analysis and critically reviewed the manuscript. GT participated in the study design and critically reviewed the manuscript. GG DE is an original investigator on the MCCS and critically reviewed the manuscript. AP participated in the study design, assisted with interpretation of results and helped to draft the manuscript. All authors read and approved the final manuscript.
**Acknowledgements**

This study was made possible by the contribution of many people, including the original MCCS investigators and the team who recruited the participants. We would like to express our gratitude to the many thousands of Melbourne residents who continue to participate in the MCCS.

**Funding Sources**

This work was supported by the National Health and Medical Research Council (NHMRC) (ID No. 209057, 334032, 396414). Further infrastructure support was provided by The Cancer Council Victoria and Monash University. AB is a PhD scholar funded by the NHMRC (ID No. 465352) GT is a Senior Research Fellow funded by the NHMRC (ID No. 390109). AP is funded by a VicHealth Fellowship. Cohort recruitment was funded by VicHealth and The Cancer Council Victoria.
REFERENCES

29. Li B, Quan H, Fong A, Lu M. Assessing record linkage between health care and vital statistics databases using deterministic methods. BMC Health Serv Res. 2006;6(48).
Monash University

Declaration for Thesis Chapter 6

Declaration by candidate

In the case of chapter 6, the nature and extent of my contribution to the work 'Unlocking the value of cohort studies for analysis of area disadvantage and health: a method for retrospective geocoding in the Melbourne Collaborative Cohort Study' was the following:

<table>
<thead>
<tr>
<th>Nature of contribution</th>
<th>Extent of contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical literature review and development of the research questions and study protocol, ethics applications for the project, data preparation, analysis design, statistical analysis, interpretation of results including manual review of geocoded addresses, writing of manuscript, submission to journal, accepts overall responsibility for the publication.</td>
<td>76</td>
</tr>
</tbody>
</table>

The following co-authors contributed to the work.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Shobhit Chandra</td>
<td>Facilitated development of research questions and protocol for geocoding, interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>2 Margaret Reynolds</td>
<td>Facilitated development of protocol for geocoding, undertook geocoding process, interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>3 Anna Peeters</td>
<td>Assisted with interpretation of results, commented on manuscript drafts</td>
</tr>
<tr>
<td>4 Andrew M. Tonkin</td>
<td>Grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>5 Rory Wolfe</td>
<td>Commented on data, commented on manuscript drafts</td>
</tr>
<tr>
<td>6 Graham G Giles</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research, commented on manuscript drafts</td>
</tr>
<tr>
<td>7 Dallas R English</td>
<td>Design and conduct of Melbourne Collaborative Cohort Study, grants used to fund research</td>
</tr>
<tr>
<td>8 Gavin Turrell</td>
<td>Facilitated development of research questions, commented on data, design of analysis, interpretation of results, commented on manuscript drafts</td>
</tr>
</tbody>
</table>

Candidate's Signature

Date 2012/12/15
Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.

(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;

(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;

(4) there are no other authors of the publication according to these criteria;

(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and

(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

| Location(s) | Department of Epidemiology & Preventive Medicine, Alfred Campus, Monash University |

[Please note that the location(s) must be institutional in nature, and should be indicated here as a department, centre or institute, with specific campus identification where relevant.]

<table>
<thead>
<tr>
<th>Signature 1</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature 2</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3/12/10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature 3</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15/12/2010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature 4</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13/12/2010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature 5</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25/1/2010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Signature 6</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5/11/2010</td>
</tr>
</tbody>
</table>
Unlocking the value of cohort studies for analysis of area disadvantage and health: a method for retrospective geocoding in the Melbourne Collaborative Cohort Study

Alison Beauchamp1* BHSc MPH; Shobhit Chandra2 BSc, MSc, PhD; Margaret Reynolds2 BA, MA; Anna Peeters1 BSc(Hons) PhD; Andrew Tonkin1 MBBS MD FRACP FCSANZ; Rory Wolfe1 BSc PhD; Graham G Giles1,3,4 BSc MSc PhD; Dallas R English 3,4 BSc MSc PhD; Gavin Turrell5 BA PhD.

1. Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia,
2. Department of Geography and Environmental Science, Melbourne, Australia
3. Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia
4. Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, University of Melbourne, Australia
5. School of Public Health, Queensland University of Technology, Brisbane, Australia

* Corresponding author

Email addresses:
AB: Alison.Beauchamp@monash.edu
SC: shobhit.chandra@monash.edu
MR: margaret.reynolds@monash.edu
AP: Anna.Peeters@monash.edu
AT: Andrew.Tonkin@monash.edu
RW: Rory.Wolfe@monash.edu
GG: Graham.Giles@cancervic.org.au
DE: Dallas.English@cancervic.org.au
GT: g.turrell@qut.edu.au
ABSTRACT

Background The effect of neighbourhood disadvantage on health outcomes is being increasingly recognised. Longitudinal cohort studies provide an opportunity to examine this relationship prospectively using geographic information systems (GIS) technology. However many cohort studies that were established decades ago did not collect geographic data on participant addresses at study commencement. We describe a method by which baseline (1990-1994) addresses of participants in a large cohort study (n=41,514) were retrospectively geocoded in order to examine the relationship between neighbourhood disadvantage at study baseline and subsequent fatal cardiovascular disease (CVD).

Results Over a series of four geocoding passes, 38,863 (94%) study addresses were matched to an address point database using GIS software. Addresses that were not matched by software were manually reviewed. In total, 41,388 (99.7%) addresses were successfully geocoded within the 1991 census areas. Each geocoded participant address was linked to Australian Bureau of Statistics 1991 Census areas including collectors’ districts, postal areas and statistical local areas. A measure of area disadvantage was then assigned to the Census areas containing study participants and these areas ordered into quintiles. Verified CVD deaths occurring after baseline were described according to quintile of area disadvantage using multilevel logistic regression. After adjustment for age, education, and ethnicity, women living in the most disadvantaged areas had higher rates of fatal cardiovascular events than their counterparts from the most advantaged areas (odds ratios 4.66, 95% credible intervals 2.07-11.81). No association was observed between neighbourhood disadvantage and fatal CVD for men.

Conclusion We demonstrate that retrospective geocoding of residential addresses for participants in large cohort studies can be undertaken with a high degree of completeness. To ‘unlock’ the hidden value of their data, custodians of health-related data should standardise address formats to be compatible with existing address-point databases.
BACKGROUND

A relationship between geographic environment and health is well-established, and many prospective cohort studies have demonstrated an important contribution of area characteristics to mortality risk.[1] The increasing use of geographic information systems (GIS) in epidemiology allows for more detailed exploration of the pathways by which location and its surrounding environment can influence health outcomes. GIS is a computer-based mapping and spatial analysis technology designed to collect and analyse data on environmental factors that may contribute to health outcomes. [2-4]

A first step in the use of GIS in epidemiological studies is that of geocoding, also known as address matching. This is the assigning of a spatially referenced point marked by geographical coordinates to a given residential address, allowing linkage of that address to spatial areas such as census boundaries. [5-6] This process can be undertaken by semi-automation, using GIS software to match addresses to a previously geocoded address-point database. Spatial areas can then be overlayed with map layers that assign specific attributes to the area, including environmental or census data. Within epidemiological studies, geocoding thus allows for analyses of many different factors such as neighbourhood-level characteristics, access to health care resources, or exposure to environmental toxins. [3-4]

Awareness of the importance of area context upon health has recently increased significantly, in part due to the use of newly developed analytical techniques such as multilevel modelling. [1] Many multilevel studies have now demonstrated that the health-related effects of poorer neighbourhoods are independent of individual-level measures of disadvantage. [7-8] This increasing awareness of the importance of area effects might prompt researchers to consider the value of ‘retrospectively geocoding’ existing cohort or population-based studies. [9] However, for studies that were established some years ago, such retrospective geocoding raises several problems. The first is the issue of matching a current address point database to study addresses that may be decades old. Over time, residential addresses may have been subdivided, suburb names changed or areas rezoned from residential to commercial land use. Second, it is important to ensure that the attributes assigned to the spatial areas of interest (such as census boundaries) are from the same time period as the study baseline. For example, when assigning census-based measures of socioeconomic disadvantage to addresses from a study that commenced in 1990, it is essential to use data from the Census period closest to that year.
Commonly however, such data from earlier time periods has either not been archived or is not widely available in the public domain.

The methodology of retrospective geocoding in established cohort studies has seldom been described, [3, 10] particularly in an Australian setting. We describe a method by which residential addresses from a large cohort study were retrospectively geocoded in order to obtain an area-based measure of socioeconomic disadvantage for each address at baseline. The overall aim of this geocoding project was to examine the relationship between neighbourhood disadvantage at study baseline and subsequent fatal cardiovascular events.

**METHODS**

**Data sources**

The Melbourne Collaborative Cohort Study

The Melbourne Collaborative Cohort Study (MCCS) is a prospective study of 41,514 volunteers who were aged 27 to 80 years when recruited in 1990 to 1994 by the Cancer Council Victoria. Twenty-four per cent of subjects were southern European migrants, deliberately over-sampled to extend the range of lifestyle factors. The majority of subjects lived in metropolitan Melbourne. Details of the design, recruitment, and study procedures have been published elsewhere. [11] All subjects provided written informed consent, and the study was approved by the Cancer Council Victoria’s Human Research Ethics Committee. Baseline examination included face-to-face interviews and questionnaires conducted in the subject’s preferred language, and collection of anthropometric and clinical measures. Information on participants’ mailing or residential address was also collected at baseline interview. Educational attainment was ascertained from the question ‘What is the highest level of education you completed?’ Education levels were categorised as: 1) primary only (comprising no school, or primary school only); 2) lower secondary (comprising some secondary education only); 3) higher secondary (comprising completed secondary education, trade certificate or some study towards a tertiary degree), and; 4) completed tertiary education (comprising degree, diploma or higher).

Fatal cardiovascular disease (CVD) events occurring in the cohort were verified through medical record review and adjudication. In brief, CVD-related deaths occurring between participant baseline attendance date and 31 December 2002 were identified through linkage with the
Victorian Registry of Births, Deaths and Marriages and the Australian National Death Index.

Participant’s medical records and autopsy reports were reviewed and categorised by panels of expert cardiologists and neurologists. Fatal CVD events were classified as related to coronary heart disease (CHD), stroke or “other” CVD cause, comprising indeterminate fatal CVD event, non-coronary cardiac death, other vascular death and heart failure.

Data used for geocoding

Geocoding is undertaken within a GIS and matches addresses to a previously geocoded database of address points, such as those held by government agencies. Using a GIS to perform this matching means that addresses can be automatically linked to spatial areas by their geographical coordinates. [3] In the state of Victoria, Vicmap Address is the authoritative geocoded database of property address points, containing over three million addresses. [2] The majority of address details in Vicmap Address have been assigned by local government, who remain the custodians of the data. Principal address details include the unit/house number, road names and locality. Post office boxes are specifically excluded. Vicmap Address is considered to have both a high level of completeness (over 94% in 2009), and a high level of spatial accuracy (90% of features are within 1mm at plot scale of their true position). [2] Data are stored in a vector format with geocoded addresses represented as features in a geographical coordinate system.

Data for measurement of neighbourhood disadvantage

In Australia, measurement of neighbourhood disadvantage is most commonly undertaken using the Socioeconomic Indexes for Areas (SEIFA), compiled by the Australian Bureau of Statistics (ABS). These indexes are derived from population census data and reflect the overall socioeconomic profile of an area based upon attributes such as income, educational attainment, levels of public sector housing, or unemployment. There are four SEIFA indexes including the Index of Relative Disadvantage (IRSD), which was used in this analysis. Each SEIFA index assigns a score to an area that shows the disadvantage or advantage of that area relative to the Australian average. SEIFA scores can be applied to small areas of approximately 250-500 dwellings through to larger areas, such as Statistical Local Areas (SLA’s). The smallest area at the time period relevant to this project is called a Census Collection District (CD). The boundaries of each CD are determined by the ABS at every census and are based on the number of houses that each census collector can visit. There were 4 977 CDs in metropolitan Melbourne at the 1991 Census. While some report that small areas such as CDs are likely to provide a more
sensitive analysis than larger areas,[12] they may not always be appropriate to use in analyses because of small numbers of participants. Therefore we also referenced to two larger areas; Postal Areas (POA) and SLA. The latter corresponds to local government and council boundaries, which provide services and infrastructure at the local and regional levels. [13]

Geocoding was undertaken at the Centre for GIS, School of Geography and Environmental Science (SGES) Monash University, who hold spatial data sets from the relevant time period in Melbourne. Ethics approval for geocoding of participant’s addresses was obtained from The Cancer Council Victoria’s Human Research Ethics Committee and Monash University.

**Process**

**Step 1: Preparation of addresses**

The first step was to clean the baseline study addresses so that they matched the *Vicmap Address* format (Figure 1). This included searching for anomalous postcodes and spelling mistakes, replacing incomplete addresses using the original attendance forms, and ensuring abbreviations were removed. Addresses were then entered into a spreadsheet with each address component in a separate field. The dataset contained the participant’s full address at baseline and their MCCS identification number only.

**Step 2: Geocoding of study addresses**

We used GIS software (MapInfo version 9) to match study addresses to the *Vicmap Address* 2009 database in a series of four passes, with each pass reducing the number of addresses left unmatched. Each pass used a different *Vicmap Address* format, with the first comprising property number, road name/type/suffix, locality and postcode. Passes 2-4 dropped suburb, postcode or property number, respectively (Table 1). Passes 1-3 were matched to the exact address point, while pass 4 was matched to the centre of the road.

Addresses that did not match to *Vicmap Address* on the first four passes were manually checked against current street directories and internet search engines to ascertain if they had undergone any transformations such as unit subdivision or retail reallocation. If a street address could not be located, the spelling or grammatical structure of the street name was reviewed, including use of apostrophes, or the numerous representations of the saint prefix such as St, St. or Saint. Data managers also revisited the original attendance forms for a new interpretation of street name spelling, and street type; for example, ‘crt’ written as ‘court’ rather than ‘crescent’. These
reviewed addresses were matched again to *Vicmap Address* using a combination of property number, road name/type/suffix, locality and postcode as described above. Addresses that remained un-geocoded following this step were manually matched to *Vicmap Address* by visual checking of the address’ position within digital spatial layers, such as cadastral boundaries (property boundaries), roads and suburbs. Post office boxes were assigned to the centroid of the digital postcode polygon.

**Step 3: Assigning study addresses to areas**

Each study address was then referenced to the CD, POA and SLA in which it belonged at the time of the 1991 ABS Census (study baseline).

**Step 4: Linking SEIFA scores to areas**

Every CD, POA and SLA identified as containing one or more MCCS participant was then linked to its 1991 SEIFA score, available from the ABS. CDs were ranked in ascending order and grouped into quintiles of SEIFA score. These quintiles were assigned to study addresses to give a CD-level measure of socioeconomic disadvantage for each individual participant at baseline. This step was repeated for POAs and SLAs.

**Statistical analysis**

Analyses were performed using Stata 10.1 (Stata Corp, College Station, Texas, USA) and MLwiN version V2.17. Multilevel logistic regression was used to assess whether neighbourhood disadvantage (measured using IRSD quintiles) was associated with fatal CVD events in the MCCS after controlling for the individual-level variables of age and ethnicity, and educational attainment. Tests of interaction were significant for sex and IRSD; analysis is therefore presented separately for men and women. A three-stage modelling approach was used. Model 1 included a random intercept only and did not condition on any other factor; Model 2 controlled for within-CD variation in age, education and ethnicity; and Model 3 further adjusted for CD socioeconomic disadvantage. Quintile 5 (i.e. the most advantaged CDs) was used as the reference category and the results are reported as odds ratios (OR) and their 95% credible intervals (CrI). When estimating the fixed and random parameters for these models we used Marko Chain Monte Carlo simulation: this procedure was implemented using the Metropolis-Hastings algorithm via MLwiN software with standard non-informative prior distributions on all parameters. To achieve convergence of the simulated chains for the variance parameters (assessed using the Raftery-
Lewis and Brooks-Draper diagnostics) the Metropolis-Hastings algorithm was implemented for 500,000 iterations. [14]

**RESULTS**

*Number of addresses geocoded*

The numbers of participant addresses that were geocoded in each pass are shown in Table 1. Using all address components, the first pass geocoded 25,620 of 41,514 addresses (62%). The second pass, using all address components excluding locality, geocoded 8,204 of 15,894 addresses (52%). Passes three and four geocoded 31% and 50% addresses, respectively. At the end of these four passes, a total of 38,863 (94%) study addresses had been geocoded to the street address level.

A total of 2,651 addresses that did not match on passes 1-4 were manually reviewed by the data managers at Cancer Council Victoria. Once reviewed, these addresses were matched again to *Vicmap Address* using the combination of property number, road name/type/suffix, locality and postcode, with a further 1,904 addresses matched at the end of this step, leaving 747 (2%) unmatched.

The remaining 747 addresses were matched to *Vicmap Address* by visually checking their position within digital spatial layers. Study addresses had not matched for a number of reasons including property subdivision where residences had been demolished and the land parcel amalgamated with others to form large blocks of units, while others had changed to shops, factories or offices. Of the 747 addresses that were manually matched, 621 addresses subsequently geocoded while 126 addresses did not (0.3% of the entire sample).

Of the 126 addresses that were not able to be geocoded, 36 were rural addresses with town name only, 45 were incomplete, and 11 were not residential addresses. We were unable to ascertain reasons for the remaining 34 non-geocoded addresses. The two most likely explanations were that either the address was correct but not present in *Vicmap Address*, or the address itself was incorrect.

There were 183 addresses (0.4%) of the entire sample that were post office boxes only and were assigned to the centre of the locality polygon. These participants were subsequently “flagged” within the MCCS dataset as having only a post office box address.

*Number of census areas identified as containing a study address*
In total, 3400 unique CDs were identified as containing an MCCS participant address (Table 2). The number of addresses contained in each CD ranged from 1-101 (median 21, inter-quartile range 13,34). Of these, the ABS defined 3334 CDs from the 1991 Census as metropolitan areas with the remaining 66 (representing 341 addresses) defined as rural areas. Table 2 also describes details of POAs and SLAs identified as containing a study address.

Linking SEIFA scores to areas

As described above, CDs were ranked into quintiles of SEIFA score, and the quintile within which each address fell was then assigned to each individual participant. For confidentiality reasons, the ABS excludes CDs with very small numbers of residents from its data on SEIFA scores. Accordingly, there were 8 CDs (n=63 participants) from the 1991 Census that were unable to be assigned a SEIFA score on this basis.

Fatal cardiovascular events by area

For this analysis, from the original 41 514 subjects, we excluded 126 with addresses that were unable to be geocoded, 186 with Post Office boxes only, 63 participants with CD numbers not matched to the 1991 ABS data, and 341 who lived in rural areas. A further 285 participants who were aged 70yrs and over at baseline, and 2507 self-reporting a history of CVD at baseline were excluded, leaving a total of 38 206 available for analysis (23 051 women and 15 155 men).

There were 394 adjudicated CVD deaths (252 men, and 142 women) during a mean follow-up time of 9.4 years per person. Table 3 shows the association between neighbourhood disadvantage, educational attainment and CVD mortality. For men, no significant association was seen between neighbourhood and CVD mortality. The odds ratio for having a fatal CVD event was higher for men with primary education compared to those with tertiary education (OR 1.82, 95% CrI 1.12, 3.01), and this association remained after adjusting for neighbourhood disadvantage (model 3). Among women, there were no significant associations between education and fatal CVD. However, women living in the most disadvantaged neighbourhood were more likely to have a fatal CVD event than those in the least disadvantaged neighbourhood (OR 4.66, 95% CrI 2.07, 11.81).
DISCUSSION

Our study demonstrates that retrospective geocoding of baseline addresses in a large cohort study can be successfully undertaken with addresses that are up to 20 years old. Geocoded addresses can be referenced to a time-appropriate census area, and assigned an area-based socioeconomic measure that can be used in analysis of mortality and neighbourhood disadvantage. Within established cohort studies, `retrospective’ geocoding of baseline addresses will therefore optimise the ability of the study to provide more comprehensive data about the socioeconomic attributes of an area and thus allow for a greater depth of analysis.

Few studies have provided a detailed description of the methodology for retrospective geocoding of addresses in cohort studies. [3] In particular, other cohort studies of CVD outcomes that have used GIS have often not reported their methods for geocoding. [15-18] Therefore the reliability of their geocoded data cannot be assessed. Describing the steps involved in geocoding allows researchers to consider a similar process for a variety of applications. For example, a case-control study from Texas reported retrospective geocoding of addresses from state-held vital records and a birth defects registry from 1997-2000, [19] and in another United States-based cohort study, childhood and early adulthood addresses of 12 681 older adults were geocoded in order to obtain measures of socioeconomic status over the lifecourse. [20]

In our study, 99.3% of 41514 addresses were successfully geocoded using a combination of automated and manual review. One cohort study, using a similar method of retrospective geocoding for a smaller number of addresses (n=5 301) and with a more recent study baseline than this present study (2000-2004)[3] reported a matching proportion close to ours (98.8%). The high geocoding proportion in our study was most likely due to the accuracy of addresses held by the MCCS. This suggests that cleaning and formatting of addresses is the most important stage of the process, although it proved to be the most time-consuming because it required returning to the original paper forms for many cases. Ensuring that addresses in a cohort study or other health-related dataset such as a clinical registry are complete, accurate and in a format that is compatible with an official geocoded address point database will mean that the process of geocoding is considerably simplified. However, it may be possible to increase geocoding success when addresses are incomplete or inaccurate by using a variety of other methods. One population-based U.S. study retrospectively geocoded 97% of 14 804 addresses by matching to
an address point database and internet mapping engines, plus contacting participants to check the accuracy of addresses. [21]

Rather than excluding participants with post office box addresses, we chose to geocode them to their postcode centroid. This might cause bias if analysis was undertaken at the CD level. A study from California traced the street addresses for 4 537 women from a breast cancer registry who held post office box addresses. A comparison of geocodes from the street addresses with those assigned to the post office boxes (which were geocoded to the Zip code centroid) found incorrect geocodes for 81% of box holders.[22] In our study, 183 participants had post office boxes, and these were flagged within the final dataset to allow for their removal from analysis undertaken at the CD level.

**CVD mortality**

We found that for women, neighbourhood disadvantage was associated with a greater risk of CVD mortality after adjusting for individual level SES, suggesting an independent effect of area upon health. We did not find conclusive evidence of an association between CVD and neighbourhood for men. Other multilevel studies have also shown that area effects on cardiovascular health are stronger for women than for men. [16, 18, 23-24] Possible reasons for this difference include the effect of neighbourhood attributes upon risk factors for CVD, particularly obesity or physical inactivity. For example, poor neighbourhood safety may be a deterrent for women to exercise.[25] Alternatively, it has been suggested that women are more susceptible to the psychosocial stressors associated with living in a disadvantaged neighbourhood. [24] These findings indicate that policies to reduce CVD in women should address area-related disadvantage where appropriate, to maximise the likelihood of their success.

**Limitations**

We were unable to validate the quality of the geocoding. While 97.3% of the addresses matched to *Vicmap Address*, this is a measure of the agreement between the two databases and not of the actual accuracy of geocoding. The value of the process is based both on the proportion of addresses that can be geocoded automatically and also the spatial accuracy of the address points. [26] *Vicmap Address* is the authoritative geocoded database of property address points in the state of Victoria, and has both a high level of completeness and spatial accuracy. While the quality of Vicmap Address geocoding for rural addresses is lower than that for metropolitan addresses, [2] only a small number of MCCS participants lived in rural areas at baseline.
Our study was based in Melbourne, and used a local address point database to geocode addresses; therefore the findings may have limited generalisability. The methodology we have described would be applicable to others who seek to undertake retrospective geocoding. The majority of health-related studies that have described their process of geocoding in detail originate from North America and this present study adds to this literature by providing a perspective from a country outside the United States.

**Conclusion**

Geocoding offers a powerful tool for epidemiologists to examine spatial relationships between area-level disadvantage and health. This description of the steps involved in retrospectively geocoding addresses for an established cohort study provides a basis for others to consider undertaking a similar process. As a first step to ‘unlocking’ the hidden value of their data, researchers and other custodians of health-related data should adopt standardised address formats that are compatible with existing spatial databases.

**List of abbreviations**

GIS = Geographic information systems

MCCS = Melbourne Collaborative Cohort Study

CVD = Cardiovascular disease

CHD = Coronary heart disease

SEIFA = Socioeconomic Index for Areas

ABS = Australian Bureau of Statistics

IRSD = Index of Relative Disadvantage

SLA = Statistical Local Areas

CD = Census Collection District

POA = Postal Areas

SGES = School of Geography and Environmental Science

OR = Odds ratios
CrI = 95% credible intervals

**Competing interests**

None declared

**Authors' contributions**

AB drafted the manuscript, participated in the study design, manually reviewed outcomes of geocoding, and performed statistical analysis. SC participated in the study design, undertook geocoding, assisted with interpretation of results and helped draft the manuscript. MR participated in the study design, undertook geocoding, assisted with interpretation of results and helped draft the manuscript. AP participated in the study design, assisted with interpretation of results and helped to draft the manuscript. AT participated in the study design and critically reviewed the manuscript. RW assisted with statistical analysis and critically reviewed the manuscript. GG DE is an original investigator on the MCCS and critically reviewed the manuscript. DE is an original investigator on the MCCS and critically reviewed the manuscript. GT participated in the study design, assisted with statistical analysis and interpretation of results and helped draft the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

This study was made possible by the contribution of many people, including the original MCCS investigators and the team who recruited the participants. We would like to express our gratitude to the many thousands of Melbourne residents who continue to participate in the MCCS.

**Funding Sources**

This work was supported by the National Health and Medical Research Council (NHMRC) (ID No. 209057, 334032, 396414). Further infrastructure support was provided by The Cancer Council Victoria and Monash University. AB is a PhD scholar funded by the NHMRC (ID No. 465352) GT is a Senior Research Fellow funded by the NHMRC (ID No. 390109). AP is funded by a VicHealth Fellowship. Cohort recruitment was funded by VicHealth and The Cancer Council Victoria.
### Table 1  
**Number of study addresses geocoded on each pass**

<table>
<thead>
<tr>
<th>Pass</th>
<th>Vicmap Address components used</th>
<th>Number (%) geocoded</th>
<th>Number (%) not geocoded</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number, road name/type/suffix, locality and postcode</td>
<td>25620 (62%)</td>
<td>15894 (38%)</td>
</tr>
<tr>
<td>2</td>
<td>Number, road name/type/suffix and postcode</td>
<td>8204 (52%)</td>
<td>7690 (48%)</td>
</tr>
<tr>
<td>3</td>
<td>Number, road name/type/suffix and locality</td>
<td>2376 (31%)</td>
<td>5314 (69%)</td>
</tr>
<tr>
<td>4</td>
<td>Road name/type/suffix and locality</td>
<td>2663 (50%)</td>
<td>2651 (50%)</td>
</tr>
</tbody>
</table>

### Table 2  
**Census areas (1991) in the Melbourne Collaborative Cohort Study**

<table>
<thead>
<tr>
<th>Census Collection District</th>
<th>Number of areas containing MCCS participant address</th>
<th>Participants in each area Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postal Area</td>
<td>138</td>
<td>504 (298, 738)</td>
</tr>
<tr>
<td>Statistical Local Area</td>
<td>80</td>
<td>1498 (926, 2540)</td>
</tr>
</tbody>
</table>
Table 3  Area and individual level effects on fatal cardiovascular events occurring in 38 206 men and women in the Melbourne Collaborative Cohort Study

<table>
<thead>
<tr>
<th></th>
<th>Men (n=15 155)</th>
<th>Women (n=23 051)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1 (Null model)</td>
<td>Model 2 (plus education)*</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.10</td>
<td>-10.22</td>
</tr>
<tr>
<td></td>
<td>-5.53</td>
<td>-16.10</td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[(Odds ratios (95% CrI)]</td>
<td></td>
</tr>
<tr>
<td>IRSD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>1.27 (0.82, 2.00)</td>
<td>4.66 (2.07, 11.81)</td>
</tr>
<tr>
<td>Q2</td>
<td>1.36 (0.89, 2.10)</td>
<td>4.95 (2.24, 12.40)</td>
</tr>
<tr>
<td>Q3</td>
<td>1.09 (0.70, 1.72)</td>
<td>3.84 (1.69, 9.85)</td>
</tr>
<tr>
<td>Q4</td>
<td>1.06 (0.67, 1.69)</td>
<td>3.53 (1.53, 9.09)</td>
</tr>
<tr>
<td>Q5</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>1.82 (1.12, 3.01)</td>
<td>1.71 (1.04, 2.86)</td>
</tr>
<tr>
<td>Some secondary</td>
<td>1.41 (0.95, 2.12)</td>
<td>1.34 (0.89, 2.04)</td>
</tr>
<tr>
<td>Completed secondary</td>
<td>1.39 (0.92, 2.14)</td>
<td>1.35 (0.88, 2.08)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neighbourhood level variance, (standard error)</td>
<td>0.036 (0.41)</td>
<td>0.053 (0.083)</td>
</tr>
<tr>
<td></td>
<td>0.922 (0.560)</td>
<td>0.750 (0.514)</td>
</tr>
</tbody>
</table>

*plus age and ethnicity (not shown)

Abbreviations: CrI=Credible Interval, IRSD=Index of Relative Socioeconomic Disadvantage
Figure 1: Steps used in geocoding of MCCS participant addresses

Step 1: Study addresses prepared for geocoding by correction of missing or incorrect data
   Each address component separated into individual fields

Step 2: Geocoding - MapInfo used to match study addresses to Vicmap Address in four passes.

Geocoded - Yes  Geocoded - No

Study addresses manually reviewed using street directories and original attendance forms

Reviewed study addresses matched to Vicmap Addresses using MapInfo GIS

Geocoded - Yes  Geocoded - No

Manually matched to Vicmap Addresses using visual checking from digital spatial layers

Geocoded - Yes  Geocoded - No

Step 3: Study addresses assigned to their 1991 Census Collection

Step 4: SEIFA scores from 1991 Census assigned to all Census Collection Districts, Postal Areas and Statistical Local Areas containing a study address
REFERENCES


The above 2 analyses describe methodologies by which existing cohort studies can be successfully modified so as to be more useful for health inequalities research. Passive follow up of participants in surveys and cohort studies can be undertaken through record linkage, while retrospective geocoding of baseline participant addresses allows for analysis of relationships between earlier environmental exposures and subsequent health outcomes.

As discussed in chapter 2, there are three main areas of research into health inequalities; the role of traditional pathways including risk factors and access to health care, life course research, and the effect of neighbourhood on health outcomes. In chapters 4 and 5 we present analyses exploring the relationships between risk factors, CVD and SES, and in the current chapter we present an analysis of neighbourhood disadvantage and fatal CVD in the MCCS. The next chapter aims to explore the role of preventive health strategies and secondary health care upon the social gradient in CVD.
How can we reduce health inequalities?

7.1 Chapter overview

This chapter aims to identify future directions by which health policy and services may play a part in reducing health inequalities. In Figure 3 (chapter 2), access to preventive health care is seen as a midstream or intermediate pathway by which SES impacts upon health outcomes. Access can be defined as both the ability to physically access health services (against barriers such as payment, geographic availability or transport difficulties), and also to avail oneself of the services (against barriers such as health literacy, cultural or language difficulties, or awareness of services). Lower SES groups have been shown to be less likely to access preventive health services for a variety of reasons, and many studies have demonstrated that the availability of secondary prevention therapies is also socially patterned. [10, 150, 154, 220] Thus it may be that health care is a possible pathway between the social determinants of health and inequalities in CVD outcomes.

The paper presented in this chapter uses a disease continuum framework to examine current ‘best-practice’ CVD preventive and treatment interventions from a socioeconomic perspective. This review describes the possible impact of these interventions on the SES-CVD gradient and makes recommendations as to the key areas most likely to impact upon the gradient.
Monash University

Declaration for Thesis Chapter 7

Declaration by candidate

In the case of chapter 7, the nature and extent of my contribution to the work "Best practice for prevention and treatment of cardiovascular disease through an equity lens: a review" was the following:

<table>
<thead>
<tr>
<th>Nature of contribution</th>
<th>Extent of contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of the research question, literature search and critical review of the</td>
<td>80</td>
</tr>
<tr>
<td>literature, interpretation of findings, writing of manuscript, submission to journal,</td>
<td></td>
</tr>
<tr>
<td>accepts overall responsibility for the publication.</td>
<td></td>
</tr>
</tbody>
</table>

The following co-authors contributed to the work.

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna Peeters</td>
<td>Assisted with interpretation of findings, commented on manuscript drafts</td>
</tr>
<tr>
<td>Andrew M. Tonkin</td>
<td>Assisted with development of the research question, commented on manuscript drafts</td>
</tr>
<tr>
<td>Gavin Turrell</td>
<td>Assisted with interpretation of findings, commented on manuscript drafts</td>
</tr>
</tbody>
</table>

Candidate’s Signature

Date 20/12/15
Declaration by co-authors

The undersigned hereby certify that:

(1) the above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.
(2) they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
(3) they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
(4) there are no other authors of the publication according to these criteria;
(5) potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit; and
(6) the original data are stored at the following location(s) and will be held for at least five years from the date indicated below:

Location(s)  Department of Epidemiology & Preventive Medicine, Alfred Campus, Monash University

[Please note that the location(s) must be institutional in nature, and should be indicated here as a department, centre or institute, with specific campus identification where relevant.]

<table>
<thead>
<tr>
<th>Signature 1</th>
<th>Signature 2</th>
<th>Signature 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>16/12/2010</td>
<td>15/12/2010</td>
</tr>
</tbody>
</table>
Original Scientific Paper

Best practice for prevention and treatment of cardiovascular disease through an equity lens: a review
Alison Beauchamp\textsuperscript{a}, Anna Peeters\textsuperscript{a}, Andrew Tonkin\textsuperscript{a} and Gavin Turrell\textsuperscript{b}

\textsuperscript{a}Department of Epidemiology and Preventive Medicine, Monash University, Melbourne and \textsuperscript{b}School of Public Health, Queensland University of Technology, Brisbane, Australia

Received 6 November 2009 Accepted 13 March 2010

Background Despite declining rates of cardiovascular disease (CVD) mortality in developed countries, lower socioeconomic groups continue to experience a greater burden of the disease. There are now many evidence-based treatments and prevention strategies for the management of CVD and it is essential that their impact on the more disadvantaged group is understood if socioeconomic inequalities in CVD are to be reduced.

Aims To determine whether key interventions for CVD prevention and treatment are effective among lower socioeconomic groups, to describe barriers to their effectiveness and the potential or actual impact of these interventions on the socioeconomic gradient in CVD.

Methods Interventions were selected from four stages of the CVD continuum. These included smoking reduction strategies, absolute risk assessment, cardiac rehabilitation, secondary prevention medications, and heart failure self-management programmes. Electronic searches were conducted using terms for each intervention combined with terms for socioeconomic status (SES).

Results Only limited evidence was found for the effectiveness of the selected interventions among lower SES groups and there was little exploration of socioeconomic-related barriers to their uptake. Some broad themes and key messages were identified. In the majority of findings examined, it was clear that the underlying material, social and environmental factors associated with disadvantage are a significant barrier to the effectiveness of interventions.

Conclusion Opportunities to reduce socioeconomic inequalities occur at all stages of the CVD continuum. Despite this, current treatment and prevention strategies may be contributing to the widening socioeconomic-CVD gradient. Further research into the impact of best-practice interventions for CVD upon lower SES groups is required. Eur J Cardiovasc Prev Rehabil 17:599–606 © 2010 The European Society of Cardiology

Keywords: cardiovascular disease, clinical guidelines, prevention, public health policy, smoking, socioeconomic status

Background Cardiovascular disease (CVD) is the leading health problem for most countries of the world, accounting for 30% of the global burden of disease in 2005 [1]. Socioeconomic inequalities in CVD are well established, and many studies report a gradient in the disease from the most to the least disadvantaged [2,3]. Despite recent declines in age-adjusted CVD mortality rates in many developed countries, evidence suggests that this gradient is widening so that lower socioeconomic groups carry an increasingly disproportionate burden of disease [2].

The evidence base to support interventions to prevent or treat CVD is arguably as robust as in any area of health. These interventions are implemented across all stages of the disease continuum from those who are well through to those with chronic heart failure, and include population-based measures, those based on systems of care, and medical therapies (Fig. 1). Many of these interventions underpin current global strategies to stem the burden associated with CVD, and are described in numerous national and international position statements and guidelines as being ‘gold-standard’ or best practice.
It is known that some population-based prevention strategies are more accessible to the better educated [4] and that there are disparities in access to health care [5]. However, the combined and cumulative effect of these inequalities across the CVD continuum is largely unknown. It is vital that the impact of CVD interventions upon lower socioeconomic status (SES) groups is fully understood. Unless they are at least equally effective among disadvantaged groups they will not reduce inequalities, and may indeed contribute to the widening socioeconomic gradient in CVD.

In this review, key interventions from each stage of the CVD continuum recommended by international policy or clinical guidelines are examined. A specific aim was to review evidence for the effectiveness of each of these interventions among lower SES groups. We also aimed to determine whether socioeconomic barriers to their effectiveness have been identified, and to describe the potential or actual impact of these interventions on the social gradient in CVD. The specific interventions reviewed are as follows:

1. **Smoking reduction strategies among the well population:**
   - Cigarette smoking is a major risk factor for CVD, and a number of interventions to reduce its prevalence are recommended, including: (a) raising the price of tobacco through increased taxation, considered an effective deterrent to smoking at a population level [6–8]; and (b) nicotine replacement therapy (NRT), shown to be moderately effective in supporting smoking cessation within the general population [9].
   - As the cost of NRT is a potential barrier to its use, particularly among lower-income groups [10], only studies of free or subsidised NRT were reviewed.

2. **Absolute risk assessment to identify those who are asymptomatic but at most risk:**
   - Most guidelines for the prevention of CVD currently recommend the use of absolute (or “global”) risk equations to identify asymptomatic people at high risk of CVD events. These equations are based on levels of the range of important cardiovascular risk factors, rather than a single elevated risk factor in isolation [11].

3. **Secondary prevention medications and cardiac rehabilitation:**
   - We reviewed two evidence-based interventions known to be highly effective at reducing morbidity and mortality after acute myocardial infarction (AMI): (a) secondary prevention pharmacotherapy, the combination of aspirin, β-blockers, angiotensin-converting enzyme inhibitors and HMGCoA reductase inhibitors (statins), estimated together to potentially reduce the relative risk of subsequent major events after AMI by up to 75% [12]; and (b) cardiac rehabilitation (CR), shown to reduce all-cause mortality by up to 47% after a cardiac event [13].

4. **Heart failure self-management programmes:**
   - In the end stages, interventions target those with chronic disease such as heart failure. The aim of treatment is to reduce symptoms as well as hospitalization rates and mortality. One important intervention is heart failure self-management programmes. These are
Study selection and inclusion criteria

Titles and abstracts were examined for relevance by two independent reviewers. Potentially relevant studies were assessed against the inclusion criteria, and disagreement resolved through discussion. Studies were included if they used quantitative outcomes to examine the effectiveness of the particular intervention among groups or individuals according to SES. These outcomes were (a) for the two smoking interventions, changes in rates of smoking prevalence or consumption; (b) for absolute risk equations, their predictive performance or changes in the proportion of people assessed at being at high risk of CVD; (c) for secondary prevention medications, cardiac rehabilitation and heart failure programmes, outcomes included changes in mortality rates, further CVD events or hospital readmissions, changes in cardiovascular risk factors, or behavioural modification.

Studies were also examined for any description or exploration of barriers to the uptake or effectiveness of the particular interventions among lower SES groups or individuals.

Exclusion criteria

As CVD largely occurs in adults, we excluded studies of interventions among children and adolescents, in whom other specific factors may be operative. We also excluded studies of sex or ethnicity-related inequalities, unless participants were specifically described as being of lower SES.

Data extraction and synthesis

The studies included in the review had heterogeneous design features and variable outcome measures. As such, it was not possible to conduct quantitative analyses and a narrative synthesis of the data is presented.

Results

We screened 517 article abstracts, identified 225 potentially eligible papers and included a final total of 49 studies. Further details of included studies are shown in Appendix I, Supplemental digital content 1 (http://links.lww.com/EJCPR/A4).

Interventions aimed to reduce smoking in the well population – increasing cigarette prices through raised taxation, and subsidized NRT

Evidence for effectiveness among lower SES groups

Fifteen studies were identified that examined the effectiveness of increasing tobacco taxation as a means of reducing smoking among lower SES groups. Most studies had a cross-sectional design, with the majority reporting that disadvantaged groups in both developed and developing countries are responsive to changes in the price of cigarettes [7,14–24]. However, evidence from three studies conducted in New Zealand, the United States and Europe suggested that increasing tobacco taxes as a smoking cessation strategy had only a limited effect on lower SES groups [4,6,25], and in Vietnam, smokers with higher education were more likely than those with lower education to quit smoking as cigarette prices increased [26].

A further 14 studies that examined the effect of NRT by SES were identified, including five randomized controlled trials [10,27–30]. All studies were conducted within developed countries, primarily the United States. Most studies combined NRT with telephone-based support, and this strategy seemed to be effective at reducing smoking prevalence for up to 12 months in lower socioeconomic groups both at a community level [9,30–32], or when disadvantaged groups were specifically targeted [10,27,33–38].

Barriers to effectiveness among lower SES groups

The ready availability of cheaper tobacco, single cigarettes, or black-market sources are described as barriers to the effectiveness of increased tobacco taxation among lower socioeconomic groups in the United States and northern Europe [7,24]. Low-income smokers in New York City preferred to buy smuggled cigarettes rather than stop smoking when tobacco prices increased, because smoking was seen as the cultural norm [7].

In all of the studies reviewed above, NRT was either free or subsidized, addressing an important barrier to its uptake among lower-income groups. Despite this, one study found that Medicaid clients in New York had only limited knowledge of the availability of subsidized NRT [21]. Other studies found that low-income smokers perceived NRT as ineffective, primarily because of underlying
factors that encouraged long-term smoking, such as peer group acceptability, stress or living in a disadvantaged area [34,38].

**Potential impact on socioeconomic inequalities in smoking**

Evidence shows varying effects of increasing tobacco taxes among lower socioeconomic groups in both developed and developing countries. These differences may be because of the stage of the smoking epidemic in that country, or the effect of other tobacco reduction interventions [25], and suggest that disadvantaged people may be less responsive to increased taxation in certain settings compared with others. In addition, lower SES groups are described as unfairly burdened by increased taxation, because the limited availability of smoking cessation support and their greater nicotine dependency means they will continue to smoke even if the cost is increased [4,7,8]. The direct impact of tobacco taxation on socioeconomic inequalities in smoking is therefore unclear.

Furthermore, although the available evidence suggests that subsidized NRT is effective among lower SES groups in the short term, longer-term smoking cessation can be difficult to sustain because of the many underlying factors associated with socioeconomic disadvantage. It is likely that subsidized NRT will be most effective when part of a comprehensive strategy that includes other approaches to support continued cessation. In England, educational differences in smoking rates are decreasing [8], and this may be because broader policies such as price increases are supported by measures specifically directed to the disadvantaged, including the provision of subsidized NRT and cessation services in deprived areas [8]. Use of revenue raised from tobacco taxes to fund such targeted cessation strategies offers a further opportunity to decrease smoking inequalities.

**Interventions targeting those at high risk of CVD – the measurement of absolute risk**

**Evidence for the effectiveness among lower SES groups**

The most commonly used absolute risk prediction equations are those derived from the Framingham cohort study, conducted in Massachusetts, USA. Although these equations are known to perform differently in different populations and ethnic groups, they can be ‘recalibrated’ using local incidence and risk factor distribution data.

Three studies from the United Kingdom (UK) and Australia were identified that examined the performance of the Framingham equations among lower SES groups, with overall findings that their use leads to underprediction of risk in those who are socially disadvantaged [39–41]. New risk equations from the UK that include area-based SES measures have recently been developed, and it was found that the proportion of lower SES individuals appropriately identified as being at high risk was significantly increased when using these new equations in comparison with the Framingham equations [11,42,43]. Adding a measure of lifetime material disadvantage to the Framingham equation in a Scottish cohort did not significantly improve its ability to discriminate between those at higher and lower risk of CVD [44].

**Barriers to effectiveness among lower SES groups**

Few barriers to the effectiveness of absolute risk equations among lower SES groups have been reported, although it is known that such people are less likely to visit their family doctor for preventive reasons. This limits the opportunities for CVD risk factor screening. Even when those at higher risk are identified, compliance with medications or lifestyle change is difficult because of the many preexisting financial and social barriers to health associated with disadvantage [11]. The Framingham equations in particular require diagnostic tests which are not available in all situations, including the poorer rural areas of lower-income countries. Accordingly, simple risk equations have been developed that require only history, blood pressure and urinalysis for use in these settings [45].

**Potential impact on socioeconomic inequalities in CVD**

Lower socioeconomic groups have a greater risk of developing CVD than those with higher SES [39], and the use of absolute risk equations in these populations is an opportunity to effectively identify and treat this risk. However, the Framingham equations, commonly used at present, do not take into account the excess risk associated with disadvantage. New risk equations that include SES are a means of identifying population subgroups in whom the need for preventive treatment is greater and as such, could be considered to be more equitable [11]. It is also suggested that lowering the treatment threshold for disadvantaged individuals could help compensate for the extra risk conferred by lower SES [40].

**Interventions targeting those with clinical manifestations of disease – combination pharmacotherapy for secondary prevention of CVD, and cardiac rehabilitation**

**Evidence for effectiveness among lower SES groups**

Three studies were identified that examine the effectiveness of combination pharmacotherapy within different socioeconomic groups. Two European studies found that patients from lower SES classes achieved comparable risk factor control to those in higher classes [46,47], and in Quebec, an increase in co-payment for cardiac medication among elderly patients did not adversely affect health outcomes, regardless of SES [48].

Four studies that examined the effectiveness of cardiac rehabilitation by socioeconomic groups were identified [13,49–51]. Overall, the evidence points to the potential
for CR to be effective at modifying major risk factors across all socioeconomic strata [49,50]. A study from the United States found that medically indigent participants benefited as much from attendance at CR as more affluent patients, with similar improvements in exercise tolerance, and dietary and smoking behaviours [50]. This study was conducted in a small centre-based programme with high staff:patient ratios and prearranged transportation.

**Barriers to effectiveness among lower SES groups**

On account of their higher risk of recurrent CVD events, prescribing rates for evidence-based medications should at least be similar and probably higher among disadvantaged populations. However, many studies have shown that even in countries with subsidized medication there is either a negative association [5,52,53], or no association [46,48] between socioeconomic disadvantage and prescribing of CVD medications. There is limited discussion of the reasons underlying these findings, although lower-income groups may have higher rates of noncompliance because of financial constraints [5]. Disadvantaged patients may also have more comorbidities, thereby limiting the medications that can be prescribed [5] (for example, β-blockers are relatively contraindicated in those with reversible airways disease).

Poor referral and attendance rates at CR have been associated earlier with socioeconomic factors such as neighbourhood deprivation and unemployment [13,54,55]. Specific barriers for lower-income groups include programme cost and lack of transport [50,55]. High levels of depression, common among socially disadvantaged adults [13], are also shown to be associated with nonattendance at CR [55]. Attendance is also influenced by the strength of recommendation from the physician, who may be more reluctant to refer lower SES patients because of scepticism about their ability to make lifestyle changes [50,55].

**Potential impact on socioeconomic inequalities in CVD**

The available evidence suggests that although CR and combination pharmacotherapy are effective among lower SES groups, access-related barriers to both are significant. Socioeconomic inequalities in rates of coronary revascularization have also been described [5,56], indicating an overall need for more research into ways to improve uptake of secondary prevention services and therapies among disadvantaged populations. The use of innovative approaches to address barriers to access is warranted. For example, prearranged collection or home-based programmes can address difficulties with transportation to CR programmes [50].

The polypill has also been proposed as a measure by which co-formulation of low-cost generic compounds might allow more equitable access to proven medications for primary and secondary prevention of CVD. This is of particular relevance to underdeveloped countries in which availability of medications is limited in poorer rural areas.

Modelling of findings from a recent randomized controlled trial in India suggested that the use of a polypill for primary prevention could potentially reduce cardiovascular heart disease and stroke by 62 and 48%, respectively [12].

**Interventions targeting patients with chronic heart failure – Heart Failure Self-Management programmes**

Five studies were identified that examined the effect of Heart Failure Self-Management (HFSM) programmes among disadvantaged populations in the United States [57–61]. When delivered within lower SES groups, these programmes were effective in reducing readmission and mortality rates and improving functional ability. One randomized controlled trial that compared lower with higher educated participants attending the same programme found no education-related differences in subsequent cardiac events or rehospitalisation [59].

**Barriers to effectiveness among lower SES groups**

Self-management of heart failure is complex, and people with lower levels of education and health literacy may be less able to effectively manage their disease than those who are better educated [57]. Financial barriers have also been identified, including costs of medication and visits to health professionals [59].

**Potential impact on socioeconomic inequalities in heart failure outcomes**

Evidence from the United States shows that HFSM programmes are effective at reducing both mortality and hospital readmission rates among lower SES groups with CHF. It may be that the intensive and regular contact with health professionals reduces the social isolation commonly experienced by disadvantaged people, and encourages their adherence to medication and exercise regimens [61]. In addition, the multidisciplinary team approach of HFSM programmes allows for a variety of treatment modalities to be used, which can be tailored to each patient’s lifestyle and resources. Programmes that are creative in their approach (such as those that use teaching materials designed for those with lower education) are also likely to be successful [57].

**Discussion**

This review has examined the effectiveness of key evidence-based interventions for the prevention and treatment of CVD within lower SES groups, identified barriers to their utilization and effectiveness, and highlighted their potential or actual impact upon the socioeconomic gradient in CVD. Overall, only limited evidence was found for the effectiveness of the interventions examined and there was little exploration of SES-related barriers to their uptake. This has significant
implications for public health policy. If these important interventions are not effective among the disadvantaged then these groups will continue to carry a disproportionate burden of CVD, and current best practice itself may be contributing to the socio-economic gradient in CVD. Indeed, because CVD imposes a major economic burden in many countries, and much of this relates to hospitalization or residential care, it can be argued that it makes sound economic sense to implement strategies particularly among those who carry the greatest burden of disease.

What might work at decreasing the SES/CVD gradient?

Key messages

Key messages that are pertinent for future efforts to reduce CVD among disadvantaged groups are shown in Table 1.

Our findings identify opportunities for intervention across all stages of the disease continuum, and illustrate the importance of both primary and secondary prevention strategies in reducing the burden of CVD among lower SES groups. It is important to note that the relative contribution of secondary prevention strategies to the poorer prognosis for lower SES groups is unclear. Studies from India and Finland suggest that disparities in evidence-based treatments during and after hospital admission for AMI account for most of the socioeconomic gradient in CVD outcomes [52,56]. However, others have shown that these same treatments explain less of the social gradient than do clinical status and CVD risk factor profile on admission [53]. These findings imply that reducing social disparities in CVD will require more than just improved access to treatment [53], and point to the vital importance of effective primary prevention strategies at earlier stages of the disease continuum so that the cumulative burden of CVD among disadvantaged individuals is reduced. One recommended approach to primary prevention is that of combining community-based or population-based interventions with those directed specifically to ‘high-risk’ individuals [62]. If lower SES individuals are seen as ‘high-risk’ because of the extra risk conferred by socioeconomic disadvantage, then this combination approach offers an opportunity to reduce the gradient seen in major CVD risk factors. This strategy has been effective at reducing smoking prevalence in disadvantaged groups in the UK, and could equally be applied to other risk factors. For example, community measures to increase physical activity levels through introducing local walking trails could be made more attractive by enhancing the safety of such trails within disadvantaged neighbourhoods. Although outside the scope of this review, it is also recognized that early-life disadvantage plays an important role in the development of the social gradient in CVD. Reducing risk factors and environmental exposures in children through population or community-based measures will therefore likely have an impact on future CVD burden [62].

Limitations

Although a limited number of interventions have been examined, those selected are known to have a significant impact on CVD morbidity and mortality. Therefore, we consider that our findings are important in advancing knowledge and understanding about strategies to reduce CVD among disadvantaged groups.

The available evidence was limited, and studies of poorer quality and small sample size were not excluded. This may restrict the generalizability of our conclusions. We also acknowledge that the focus of this review is on developed countries. This is primarily because evidence from the developing world remains limited, despite these countries carrying the greater burden of CVD [62]. Furthermore, although not all underdeveloped countries show an inverse socioeconomic gradient in CVD, this may change as they become more industrialized. For example, in India, there has been a transition over the last decade from a positive association between social advantage and CVD to an inverse one. This shift is thought to be because of several factors including increased urbanization and greater uptake of health protective behaviours among the higher social classes [63]. Therefore, although our findings may not have immediate relevance for these countries, they are likely to be important in the future.

Exclusion of ethnicity-related inequalities (unless specifically described as lower SES) may have also underestimated the full extent of inequalities in CVD, particularly in the context of increasing multietnicity within cities such as New York and London. The issue of ethnicity-related differences in CVD is highly complex,
and complicated by cultural and genetic factors [64]. In particular, migrant groups may be at greater risk because of interactions between genetic susceptibility and their new environment. However, not all migrants or minority ethnic groups are from a lower SES background. It has also been shown that a substantial portion of the 10-year risk of CVD that was associated with ethnicity in the United States could be attributed to SES or geography [64]. In view of this complexity and to limit the scope of the study, we elected to include only ethnic minorities if they were identified as having low SES.

Conclusion
There are significant socioeconomic disparities in the uptake and delivery of key prevention and treatment strategies at all stages of the CVD continuum, and a paucity of evidence for their effectiveness among lower socioeconomic groups. Increasing this evidence base will require greater research efforts and a concerted and coordinated approach across many sectors of society. Although there will be little impact on CVD inequalities until the underlying determinants of health inequalities are addressed, policy makers and clinicians must recognize the importance of SES as an independent risk factor for disease, and seek ways of incorporating it into the current best-practice management of CVD.

Acknowledgement
This work was supported by the National Health and Medical Research Council (ID No. 465352 to A.B. and ID No. 390109 to G.T.). A.P. was funded by a VicHealth Fellowship.

Conflicts of interest: none declared.

References
15 Lee JM, Hwang TC, Ye CY, Chen SH. The effect of cigarette price increase on the cigarette consumption in Taiwan: evidence from the National Health Interview Surveys on cigarette consumption. BMC Public Health 2004; 4:61.


Copeland L, Robertson R, Elton R. What happens when GPs proactively.

Ivers RG, Farrington M, Burns CB, D'Abbs P, Richmond R, Tipiloura.


Daly J, Sindone A, Thompson DR, Hancock K, Chang E, Davidson P. Barriers to participation in and adherence to cardiac rehabilitation programs: a critical literature review. Prog Cardiovasc Nurs 2002; 17:8–17.


CHAPTER 8

Discussion and conclusions

This thesis aimed to describe and explore pathways and mechanisms by which social inequalities in CVD occur, and to describe the extent to which healthcare systems and targeting risk factors might potentially modify health inequalities. Specific questions addressed include:

- What is the association between SES (measured using educational attainment) and CVD mortality in an Australian setting, and what proportion of this association can be explained by cardiovascular risk factors
- Will addressing ‘midstream’ factors such as cardiovascular risk factors and access to health systems have an impact upon the social gradient in CVD
- Does SES predict the incidence of cardiovascular risk factors

The thesis contributed to two of the main areas of research into health inequalities, namely traditional pathways research, and neighbourhoods and health research. The thesis has broad ramifications for health inequalities generally. Firstly it provided further evidence for the important role of risk factors in the development of inequalities in CVD, and arguably in other chronic diseases that may share common lifestyle and other risk factors (such as cancers, diabetes or lung disease). Using longitudinal data, this thesis also showed that gradients in some major risk factors are predicted by SES (measured using educational attainment), thus providing further evidence for the importance of addressing inequalities in the social determinants of health earlier in life. A further finding was that current recommended and commonly used prevention and treatment strategies may be increasing health inequalities, and this is likely to be the same for other diseases. Thus the findings from this thesis not only add to the important area of CVD inequalities, but are also applicable to health inequalities more generally. Specific findings are discussed below.
8.1 Overview of major findings

8.1.1 Educational inequalities in CVD and in major risk factors

The analysis presented in chapter 4 described an inverse educational gradient in fatal CVD events. This has previously been seen in many other studies, although much evidence for socioeconomic gradients within Australia is based on the use of large scale area-based indicators such as Statistical Local Areas. [66, 70] Our study is the most recent to describe CVD inequalities using an individual-level SES indicator. The results from this thesis provide strong evidence for the current existence of SES-CVD gradients, reinforcing the need for further research into the causes of these gradients.

Chapters 4 and 5 also present contemporary evidence for continuing educational gradients in major risk factors for CVD, including smoking, overweight, diabetes and hypertension. As described in chapter 2, while it is known that these risk factor gradients contribute to gradients in CVD, the magnitude of their contribution is unclear. Chapter 4 aimed to describe the relative contribution of major risk factors to the excess risk of fatal CVD events that is associated with educational attainment. While other studies have examined this question using a similar methodology, there is considerable variation in the amount explained. This thesis contains one of the largest studies to examine this in a population of both men and women. The study is one of the few to examine grouped risk factors, and the first to do so using Australian data. It was found that the combination of eleven behavioural, social and physiological risk factors explained a significant proportion of the difference between the lowest and highest educated groups. As such, these findings add strong support to the importance of gradients in risk factors towards gradients in CVD.

The use of longitudinal data in this thesis provides further evidence for the temporal sequence between SES and social gradients both in CVD and in some risk factors. Educational attainment was shown to predict incident CVD in chapter 4; and in chapter 5 was shown to be associated with the incidence of diabetes in both men and women, and with incident overweight and obesity in women. This has also been seen in other studies, although few longitudinal studies have included more than one risk factor or have examined the relationship between educational
attainment and accumulation of incident risk factors over time. As such, this evidence reinforces that addressing socioeconomic inequalities in the underlying determinants of health will be required if we are to truly alleviate health inequalities.

This thesis also identified that gradients in some risk factors start even before conventional cut-off points for risk. In chapter 5, it was shown that among women who were not “at-risk”, educational gradients were still apparent in mean levels and prevalence of major risk factors. No other studies were identified in which risk factors measurements were described according to whether or not people were defined as being “not at risk”. The findings from this thesis thus imply that early and targeted prevention, screening and treatment of risk factors in lower socioeconomic groups is needed.

### 8.1.2 Individual risk factors

There is also some uncertainty about which of the individual risk factors are the most important for explaining socioeconomic differences in CVD mortality. For example, some studies identify smoking as the most important health behaviour, but others have found that physical activity is more important. The analysis presented in chapter 4 found that smoking and central adiposity were the two individual risk factors contributing most to the educational gradient in CVD seen in the MCCS, and these findings held regardless of whether or not southern European born migrants were included in analysis. Because trends in population levels of these and other risk factors are changing, such contemporary evidence concerning their impact upon health inequalities is vital. Specific findings of importance are highlighted below.

Socioeconomic gradients in smoking are well documented and have been shown to be associated with inequalities in CVD. As stated above, trends in smoking prevalence are changing, and overall declines have been seen in Australia. [195] The findings presented in chapter 5 of narrowing educational gradients in smoking within the AusDiab cohort are also encouraging. As described in chapter 7, these narrowing gradients in smoking may be related to population-based activities such as increased taxation on tobacco, shown to be effective among lower SES groups in Australia. [221] Despite this, the study in chapter 4 found that smoking remained the single most important contributor to socioeconomic gradients in CVD even in a setting of relatively low prevalence. This reinforces the importance of smoking as a contributing factor to CVD.
inequalities, and suggests we should not become complacent in our efforts to address smoking among lower socioeconomic groups.

Evidence suggests that smoking leads to CVD partly through its action on lipoproteins. However, it is unclear as to whether this association is independent of lifestyle factors, in particular diet and physical activity. If the association between smoking and lipids is mediated by these lifestyle risk factors, this would have particular relevance to lower SES groups who are not only more likely to smoke, but also to have a worse diet and be less physically active. The study in chapter 5 examined the association between smoking status and lipoprotein subclasses. This was the first study to examine this association using NMR-spectroscopy, which provides a very detailed profile of lipoprotein subclass concentration and size. The study found that female smokers had a more adverse lipid profile than non-smokers, and that this association occurred independently of risk factors including dietary intake, physical activity and weight. It may therefore be necessary to treat abnormal lipids in female smokers aggressively, using medication rather than relying on lifestyle modification which is unlikely to be effective. This study therefore adds to the body of knowledge on pathways between smoking and CVD, and has significant implications for lower socioeconomic women who may have higher rates of smoking.

Our findings provide contemporary evidence for the importance of central obesity in explaining CVD inequalities and also show that the incidence of overweight and obesity is greater among lower educated women. In view of the current and projected obesity epidemic in Australia and other developed countries, it is important that our knowledge of the impact of obesity on future SES-CVD gradients is up-to-date. [Backholer K, et al (2010) ‘Projected socioeconomic disparities in the prevalence of obesity amongst Australian adults’, submitted manuscript] Using contemporary data, central adiposity was identified in chapter 4 as the second-most important risk factor contributing to excess risk of CVD among the lowest educated group in the MCCS. In addition, the study in chapter 5 identified that women with lower education were more likely to develop overweight or obesity over a five-year period. Overweight and obesity have been shown to lead to major physiological risk factors including diabetes and hypertension. [222] Social gradients in obesity are currently not as pronounced as for other risk factors. [102] However, if mean population weight continues to increase at its present rate, it is likely that socioeconomic gradients in overweight and obesity will become more apparent. [Backholer K, et al] The findings
from chapters 4 and 5 therefore suggest that the contribution of overweight and obesity to future gradients in CVD should not be underestimated, particularly among women.

The findings from this thesis are of strong educational gradients in the prevalence of diabetes in both men and women in the MCCS and AusDiab study. In addition, the incidence of diabetes was higher in lower educated women in AusDiab. Our findings add support to other studies which have also shown that the incidence of diabetes is inversely associated with SES. [104-106] As noted above, increased weight has also been shown to be associated with a higher risk of type 2 diabetes in other studies. Diabetes is a significant risk factor for CVD, and these findings suggest that as gradients in diabetes continue and even widen, so will gradients in CVD.

8.1.3 The relative importance of risk factors differs by ethnicity

It was also shown in chapter 4 that the relative importance of risk factors in explaining educational gradients in CVD differs between those born in southern Europe, and those born in northern Europe/ Australia and New Zealand. It was found that when southern European born migrants were excluded from analyses, behavioural risk factors explained a greater proportion of the excess risk of fatal CVD in the lowest educated group than did physiological risk factors. When southern European migrants were included in the analyses, physiological risk factors explained the greatest proportion, (which may be due to the higher prevalence of diabetes among this sub-group). In addition, adjustment for social connectedness had a greater effect when southern European born migrants were excluded from analyses. Although low numbers of events in southern European born participants in the MCCS precluded further exploration of these differences, overall these findings suggest that the pathways between SES measured using educational attainment) and CVD may differ between ethnic groups. This has been seen by others, [223, 224] with some European studies demonstrating cross-national variation in the strength of the SES-CHD association between southern and northern European countries.[225, 226] This variation may be due to biological and genetic factors, [81] as well as cultural influences on lifestyle such as diet and smoking, [4] and implies that strategies to reduce CVD inequalities should be relevant to the context and population in which they are undertaken. [223, 224]
8.1.4 Best-practice interventions to prevent or treat CVD may be contributing to the social gradient in the disease

As described in chapter 2, many authors describe inequities in access to CVD prevention, screening and treatment strategies. While the contribution of such inequities to social gradients in CVD is unclear, in the interests of social justice and fiscal responsibility such inequities must be addressed. In particular, because the burden of CVD is much greater among lower SES groups (as shown in chapter 4), it is essential that any access-related barriers to interventions to prevent or treat the disease are identified. It is also essential that such strategies are examined for their effectiveness across the socioeconomic spectrum, because if they are less effective among lower SES groups, then this will only contribute to widening gradients in the disease. In chapter 7, a review found consistent evidence for the inequitable access and uptake of current best-practice interventions to prevent or treat CVD, and little evidence for their actual impact upon lower SES groups. These findings add considerably to the literature as no similar review has been identified, and as discussed in section 8.5, suggests that further research in this area is required. The findings from chapter 7 have significant implications for public health policy and clinical practice as they indicate that current best-practice interventions may be contributing to the socioeconomic gradient in CVD.

8.1.5 Potential opportunities to reduce socioeconomic inequalities in CVD occur at all stages of the disease continuum

The review in chapter 7 also identified opportunities for intervention across all stages of the disease continuum, and illustrated the importance of both primary and secondary prevention strategies in reducing the burden of CVD among lower SES groups. Both chapters 2 and 7 suggested that for primary prevention of CVD, a combination of population-based and targeted approaches may provide the best opportunity for improving uptake of prevention strategies. Chapter 7 also concluded that if screening for those at high-risk of CVD is undertaken it should account for the independent effect of SES, with allowances made for socioeconomic differences in compliance with medication or lifestyle modification. Furthermore, the review in chapter 7 found that both secondary prevention combination pharmacotherapy and cardiac rehabilitation are likely to be effective in lower SES groups following AMI, but that access to both is socially
patterned and barriers therefore need to be identified and addressed. Finally, it was identified that chronic disease self-management programs are likely to be effective in lower SES groups with heart failure because their intensive and multi-disciplinary nature means they can be tailored to an individual’s needs. Chapter 7 thus provides key points along the disease continuum at which public health policy makers could potentially intervene to reduce the SES-CVD gradient.

8.1.6 Existing Australian cohort studies can be adapted for use in studies of health inequalities through record linkage and geocoding of participant addresses

Australia’s research capacity into health inequalities is limited, as few cohort studies are designed specifically to study this area. It is known that record linkage offers a way to combine different datasets in order to increase the body of knowledge about individuals. [227] Examples of how this could be used in health inequalities research include linkage with educational, social welfare, or health administrative datasets in order to create a complete picture about a person across the lifecourse. In addition, in chapter 5, it was shown that lower education was associated with less likelihood of returning for follow-up when eligible. Record linkage offers a means by which passive follow up of cohort participants over time can be undertaken through linkage with hospital morbidity datasets. However, few studies have described the methodology for linkage of two datasets, particularly where there is no unique identifier, and this may deter researchers from undertaking the process. The study in chapter 6 describes in detail a methodology for linking a large and established cohort study with a hospital administrative dataset in order to determine CVD outcomes in the cohort, thus providing a valuable resource for social epidemiologists and other researchers. This is discussed further in section 8.5.

In addition, there has been a recent explosion of interest in the role of neighbourhood factors in explaining health inequalities. This is of particular importance when studying the underlying determinants of health, as individuals are linked to their environments in a multitude of ways. As outlined in chapter 2, multilevel modelling is a statistical technique that allows for the analysis of area-effects over and above individual-level effects, thus providing valuable information about the role of area and context in causing social gradients in health. Geographical information systems (GIS) offer the potential to explore environmental effects on disease outcomes within cohort
studies using sophisticated software. A first step to using GIS is to assign geographical coordinates to participant’s addresses using a process known as geocoding. Cohort studies that were established some time ago will now be showing health outcomes that will be valuable for current analysis, but if geocoding was not undertaken at baseline, then researchers are potentially missing out on a whole range of possible longitudinal analyses to explore associations between neighbourhood disadvantage and health. As stated several times throughout this thesis, cross-sectional analyses provide a useful starting point for describing health inequalities, but it is longitudinal datasets that will allow us to explore the underlying pathways in more detail. As such, a detailed description of a method for retrospective geocoding of an existing dataset as presented in chapter 6 is a valuable resource for social epidemiologists, particularly in an Australian setting as no other studies have described this methodology.

In summary, findings from this thesis reinforce the presence of strong educational gradients in fatal CVD and in major risk factors for the disease. They provide supporting evidence that these gradients in risk factors contribute significantly towards gradients in CVD, and also reinforce that educational attainment is a predictor both of incident risk factors and of incident CVD, and temporally precedes the onset of increased risk. Individual risk factors of greatest importance include smoking, central adiposity, and diabetes, and possible mechanisms by which gradients in these risk factors may contribute to current and future socioeconomic gradients in CVD were elucidated. The importance of considering the cultural and lifestyle differences between certain ethnic groups was also highlighted. This thesis also identified that, despite the greater burden of CVD among lower SES groups, there are known inequities in access to best-practice interventions for the disease. In addition, little is known about either the potential impact of these interventions upon the SES-CVD gradient, or about barriers to their uptake and effectiveness. As such, current best-practice interventions to prevent or treat CVD may be contributing to widening socioeconomic gradients in the disease. Finally, this thesis described in detail two methods by which existing cohort studies, clinical registries and other datasets can be adapted for use in health inequalities research, thus increasing Australia’s research capacity in this area of vital importance.
8.2 Strengths of the data sources used in this thesis

8.2.1 Prospective cohort study design

The longitudinal study design of both AusDiab and the MCCS allows exploration of the temporal sequencing of exposures and outcomes, a criterion for establishing causality. This is of particular importance in studies of health inequalities where reverse causality can be an important issue. The longitudinal study design also removes the influence of recall bias, which may be associated with SES, in particular among those with lower education and health literacy. Finally, the prospective cohort study design allows us to examine patterns of change in risk factors over time at an individual level, including the effects of socioeconomic disadvantage across the life course. As such, this study design is particularly well-suited to health inequalities research.

8.2.2 Sample sizes

AusDiab is the largest national population-based longitudinal study in Australia, comprising 6 400 participants who attended both baseline and follow-up. The MCCS included a very large sample size of 41 514 participants at baseline, and is one of the largest cohort studies in Australia to undertake research into health inequalities. The large sample size of both studies allows for stratification of data by socioeconomic and demographic characteristics, enabling a more detailed exploration of causal pathways and mechanisms. In addition, data from both studies are based on contemporary populations of men and women. Thus findings from analyses are current, robust, and relevant to both males and females.

8.2.3 Accurate ascertainment of CVD outcomes in the MCCS

All fatal CVD events occurring in the MCCS were verified using a rigorous process of data collection and adjudication. Mortality follow-up was almost complete, with 94% of all deaths occurring in the first 12 years of the study available for verification. The verification process involved deaths adjudicated by panels of expert clinicians using internationally accepted definitions. Agreement from at least two adjudicators was required and adjudicators were blinded to outcomes assigned by their colleagues.
8.2.4 Accurate and comprehensive measurement of clinical indicators at baseline

Both the MCCS and AusDiab studies collected comprehensive measures of behavioural and physiological risk factors at baseline. While behavioural risk factors were mostly self-reported, questionnaires for smoking and physical activity were validated, and in the MCCS were in the participant's preferred language. Detailed information was collected on dietary intake for both studies using the food frequency questionnaire, and this allowed for inclusion of a number of dietary variables in analysis which was particularly relevant for the analysis of lipoproteins and smoking, and for the analysis of the role of behavioural risk factors. Physiological risk factors were accurately measured using standard methods. A particular strength of the AusDiab study was the use of an OGTT, the gold standard for assessing glucose tolerance and hence diagnosing diabetes.

8.3 Limitations of the data sources used in this thesis

Each of the relevant limitations is discussed in detail in the papers presented here. The purpose of this summary is to provide an overview of the main limitations. It is important to note that one purpose of this thesis is to determine the feasibility of using existing Australian datasets to contribute to health inequalities research; some limitations of the data are therefore an inherent part of the thesis.

8.3.1 Response rates and loss to follow up in AusDiab

The baseline response rate for AusDiab was approximately 37%, and only 59% of those eligible (n=6 400) returned for follow up in 2005. This affects the generalisability of our findings, as the sample may not be representative of the Australian population. There were also some demographic and educational differences between responders and non-responders at baseline, possibly leading to selection bias. For those who were eligible for follow-up, lower educational attainment was associated with non-attendance, and so our findings are likely to underestimate the
association between SES and incidence of CVD risk in AusDiab. However, mean levels of baseline risk factors were generally similar between those who did and did not attend follow-up. While non-attendees were more likely to be smokers, this was the case for all educational groups, and therefore would be unlikely to affect our findings concerning the relationship between education and smoking.

8.3.2 Limited measures of SES

Education was the sole SES indicator used in this thesis, resulting in limited exploration of the different pathways between SES and CVD. The use of single indicators of SES in analyses of health outcomes may result in residual confounding, as one indicator alone cannot capture all aspects of SES. [33] This residual confounding may lead to overestimation of the association between SES and health. However, use of more than one SES indicator is also likely to be biased because the strong correlation and subsequent shared variance between SES indicators will introduce bias into the analyses. [33]

The usefulness of education as an indicator can depend on the characteristics of the cohort being studied. [33] In particular, education may have different meanings among older compared to younger cohorts. Use of education level as an indicator is generally suited to the age range of the MCCS participants (as the majority of those attending a tertiary institution would be expected to have done so by age 40 years); however, in the AusDiab cohort, the minimum baseline age of participants was 25 years, and not all will have achieved their maximum educational attainment. In addition, a further limitation of education occurs if migrants were brought up in a country in which level of education has different implications to those of the host country. [21] This is particularly relevant in the MCCS, in which the association between education level and CVD is much less apparent for Southern European migrants (who were deliberately over sampled in the MCCS), as it is for other ethnic groups in the cohort.

However, education was considered to be the most reliable and appropriate measure of SES in both studies for a variety of reasons. As illustrated in chapter 2, education as an SES indicator is generally less affected by reverse causality than other indicators, and hence is more reflective of early life circumstances than occupation or income. [23] This is an important consideration in
studies of CVD in which the disease process is thought to begin in childhood or even earlier. Education has also been shown to be associated with mortality in both men and women, whereas the associations with occupation, income or social class have been shown to differ by sex. [228] Furthermore, educational inequalities in CVD mortality have been seen in southern and northern Europe [4] and in Australia/ New Zealand; [8] all of which are regions represented in the cohort, and finally, education level is easily measured and participants may be more responsive to this question than to those regarding income and wealth, [23] meaning there is less tendency to misclassification.

8.3.3 Control for confounding factors

As mentioned above, residual confounding may result from use of a single indicator for SES. Further, residual confounding by other unmeasured factors is a potential limitation. This is of particular importance in studies of health inequalities because the complex nature of interactions between people and their environment makes it difficult to control for all known confounders.

Confounding by ethnicity in the MCCS is a further potential limitation. Within the MCCS, there was a significantly greater proportion of southern European born migrants with primary education only compared to those born in northern Europe or Australia and New Zealand. It is possible that some of this difference could be caused by differential misclassification resulting from language barriers, although questions about education were in the subjects’ preferred language. However, as described above, educational attainment may have different socioeconomic meanings within the countries represented in this cohort. For example, having only primary education may not be uncommon in southern Europe particularly among older cohorts, and education may thus not be a sensitive indicator of SES in this population.

8.3.4 Measurement error and misclassification

Self-report of behavioural risk factors by participants may lead to measurement error, in turn leading to under or over-estimation of the contribution of risk factors to gradients in CVD. Baseline CVD in both the MCCS and AusDiab was self-reported and if associated with lower education, may have resulted in overestimation of the effect of SES on incident CVD if
participants did, in fact, have prior CVD. Other studies have shown that accuracy of self-reported CHD is moderately high at 84% for ischaemic heart disease, [229] 94% for AMI and 80% for stroke, [230] although findings from these studies were not reported by SES.

8.3.5 Associations between SES, risk factors and stroke

In contrast to most findings, some European studies have shown that the same risk factors that explain inequalities in CHD have a lesser role in explaining inequalities in stroke. [225, 231] One study found that diabetes and physical functioning contributed more to inequalities in stroke than did conventional risk factors such as smoking.[225] This may be related to the type of stroke (ischaemic or haemorrhagic) with one author finding that haemorrhagic stroke is less likely to demonstrate a socioeconomic gradient in some European countries. [232] Fatal CVD outcomes in the MCCS did not differentiate between ischaemic and haemorrhagic stroke; of the 99 fatal stroke events, 65 were adjudicated as being haemorrhagic, which may affect the strength of association seen between fatal CVD and educational attainment in the MCCS.

8.4 Implications of thesis findings

These findings have significant implications in three main areas; social epidemiology research, health policy, and clinical practice. First, the widening social gradients in risk factors that were identified in this thesis mean that public health policy must immediately be directed towards narrowing these risk factors if future gradients in CVD are to be reduced. In the long-term, this will require a multi-sectoral, multi-level approach towards reducing inequities in the social determinants of health; however in the shorter-term, addressing risk factors among lower SES groups through a combination of targeted and population-based approaches may be effective. While evidence for ‘what works’ at reducing health inequalities may be limited at this stage, enough is known to be able to implement effective strategies; in particular, a greater emphasis on ensuring that current best-practice interventions are accessible to all is required.

Second, for clinical practitioners, the findings from this thesis clearly support the importance of incorporating SES into treatment decisions, particularly those that require lifestyle modification.
Clinicians must recognise the many barriers to uptake of treatments, medication compliance or behaviour change. At both a primary-care and a health systems level it is essential that socioeconomic barriers to service access are identified and addressed.

Finally, this thesis has implications for social epidemiologists and other researchers into health inequalities. As identified in chapter 2, inequalities in CVD have been well described and evidence for the importance of risk factors in contributing to these gradients is growing. What remains unclear however are the mechanisms by which risk factor gradients start, how they might lead to CVD inequalities, and the importance of cumulative risk over the lifespan. Longitudinal studies have much to offer in this area of research, and this thesis suggests two methods by which existing datasets can be utilised so as to provide a valuable resource for social epidemiologists to explore the causes of inequalities in CVD. Further areas for research are described more fully in section 8.5.

8.5 Areas for further research

Studies have found that southern European migrants have lower rates of CVD despite having a higher prevalence of diabetes and other risk factors, and despite having less formal education as identified in chapter 4. [4] An earlier study conducted in the MCCS reported that the Mediterranean diet, predominant among southern Europeans was cardioprotective. [233] Social factors may also play a part, as chapter 4 identified that social connectedness had a greater impact upon the educational gradient in CVD within the MCCS when southern Europeans were excluded from analysis, implying that the social factors adjusted for may not be important contributors to CVD inequalities in this ethnic group. Further research into the association between ethnicity and social gradients in the MCCS, in particular the role of diet and neighbourhood is therefore warranted. The small number of fatal CVD events among southern Europeans in the MCCS precluded this at time of analysis. However, linkage of the MCCS cohort with the Victorian Admitted Episodes Dataset to determine non-fatal CVD events in the cohort is underway, and this will provide a greater number of outcomes allowing for stratification of analysis by ethnic group. In addition, further development of a geographic information systems database in the MCCS will allow the effects of environmental exposures on CVD inequalities to be examined. This will also provide an opportunity for multilevel analysis of the
effects of neighbourhood over and above individual-level effects on the incidence of major risk factors for CVD. Few studies have examined the association between neighbourhood and the incidence of cardiovascular risk factors, and given the important contribution of risk factors to CVD inequalities, such work would be extremely valuable.

A further area of research is to examine the implications of adding either education or an area-based measure of disadvantage into absolute risk equations for cardiovascular disease. Only a limited number of studies have done this, [234-236] and none in an Australian context. Risk equations such as the Framingham have been shown to underestimate CVD risk in lower SES groups, yet are widely recommended for use. In Australia for example, the National Vascular Disease Prevention Alliance has recommended that the Framingham equation be used for assessment of absolute risk in the primary care setting. [65] Studies from the UK have shown that incorporating an indicator of SES into newly developed equations improves the proportion of lower SES individuals identified as being at high risk. [234, 235] Research into whether adding SES to the Framingham equation (or to new equations currently being developed in Australia) improves accuracy among lower SES groups should be conducted in the near future.

The review in chapter 7 found that few studies of the effectiveness of best-practice interventions for CVD had examined outcomes according to SES, and therefore the impact of such interventions on the SES-CVD gradient is largely unknown. A similar review of best-practice interventions to prevent or treat major cardiovascular risk factors through an equity lens is warranted. As reported in this thesis, gradients in obesity and diabetes are likely to widen in the future, and will almost certainly lead to widening gradients in CVD. Many prevention and treatment strategies for these risk factors are currently available, yet it is probable that their effectiveness among lower SES groups has not been established. Such a review will not only highlight key areas for intervening to reduce health inequalities but will also direct the evaluation of such interventions, ensuring a valuable resource for public health policy makers, practitioners and researchers.
8.6 Conclusion

From the initial study hypotheses, the following conclusions can be made: first, that there is an inverse association between SES (measured using educational attainment) and CVD mortality in an Australian setting and cardiovascular risk factors can explain a large portion of this association; second, that addressing ‘midstream’ factors such as cardiovascular risk factors and access to health systems will have an impact upon the social gradient in CVD; and third that educational attainment is a predictor of risk factor onset and change over time. Fourth, that existing cohort studies can be modified to increase their potential for use in social epidemiological research. Finally, this thesis concludes that current best-practice interventions for CVD are potentially widening the SES-CVD gradient. The findings of this thesis have major implications for public health as they contribute significantly towards increasing our understanding of health inequalities.
References


APPENDIX 1

Editorial - “The Importance of Extinguishing Secondhand Smoke”
The Importance of Extinguishing Secondhand Smoke
Andrew M. Tonkin, Alison Beauchamp and Christopher Stevenson
Circulation 2009;120;1339-1341; originally published online Sep 21, 2009;
DOI: 10.1161/CIRCULATIONAHA.109.894758
Circulation is published by the American Heart Association. 7272 Greenville Avenue, Dallas, TX 75251
Copyright © 2009 American Heart Association. All rights reserved. Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/cgi/content/full/120/14/1339

Subscriptions: Information about subscribing to Circulation is online at http://circ.ahajournals.org/subscriptions/
Permissions: Permissions & Rights Desk, Lippincott Williams & Wilkins, a division of Wolters Kluwer Health, 351 West Camden Street, Baltimore, MD 21202-2436. Phone: 410-528-4050. Fax: 410-528-8550. E-mail: journalpermissions@lww.com
Reprints: Information about reprints can be found online at http://www.lww.com/reprints

Downloaded from circ.ahajournals.org at Monash University on October 6, 2009
The Importance of Extinguishing Secondhand Smoke

Andrew M. Tonkin, MD, FRACP, FCSANZ; Alison Beauchamp, RN, BHSc, MPH; Christopher Stevenson, BHSc, MSc, PhD

Government policies have enormous influence on the health of nations. Arguably, this is illustrated most vividly with tobacco control. However, smoking continues to be a global problem and the major cause of preventable death. The countries with the highest per-capita smoking prevalence rates include (alphabetically) Bangladesh (20.9% of adults), Brazil (16.2% of adults), China (31.4% of adults), Germany (27.2% of adults), India (32.7% of men, 1.4% of women), Indonesia (34.5% of adults), Japan (43.3% of men and 12% of adults), the Russian Federation (60.4% of men, 15.5% of women), Turkey (34.6%), and the United States (23.2%).

Prevalence rates among younger people vary, but in the United States, 18.4% of youths still smoke.1

This issue of Circulation includes a meta-analysis of the impact of smoking bans on hospitalization for acute myocardial infarction.2 The findings further attest to the power of government interventions.

The meta-analysis compares rates of acute myocardial infarction after with those before legislation restricting smoking in public areas in various North American and European populations. The pooled random effects estimate of the rate of hospitalizations for acute myocardial infarction 12 months after legislation was 0.83 (95% confidence interval 0.80 to 0.87). This increased with time up to 3 years, the longest follow-up. Variability in the duration of follow-up accounted for most of the heterogeneity between findings in the individual studies. The findings were robust, with realistic simulations of individual risk and exposure scenarios.

One difficulty with assessing population-level health promotion activities, such as those aimed at secondhand smoke, is that it is difficult to measure their effect directly. Population health is complex and is influenced by many factors, and it is generally not possible to isolate and measure the effect of any specific program independently of any other factors. Hence, a particular study can measure an effect only in a specific context, which may or may not be generalizable. Meta-analysis can address this to some extent by combining the results of studies done in different environments to give an overall effect estimate. However, the results are subject to differences between the study designs and end points and so must still be interpreted with some caution.

Another approach to assessing population health interventions is to combine evidence from observational epidemiology and from clinical trials using modeling techniques. This can allow projection of the likely effect of a health intervention to a population that may differ from those in which the original studies were undertaken. However, epidemiological models often rely on assumptions that cannot be tested and on simplifications of characteristics of the intervention and the disease process. They also frequently incorporate a large number of estimated parameters. This can lead to a spuriously close fit between the model results and the data on which the model parameters were based.

Lightwood and Glantz2 used a combination of these approaches. Their meta-analysis and the model draw their parameter estimates from different, independent data sources. Hence, the demonstrated consistency between the results allows greater confidence in the results of each approach. This is important because policy makers are informed by such data in balancing costs and effectiveness and in prioritizing between different population health interventions.

A further feature of epidemiological models is that they can highlight important gaps in the data available to inform health policies and strategies. Typically, studies of the risks of secondhand smoke have relied on self-reporting. As discussed by the authors, individual exposures, which are presently measured by levels of cotinine, a stable metabolite of nicotine, were not generally available. However, all the scenarios modeled by Lightwood and Glantz2 that showed a close agreement with the results of the meta-analysis were based on parameters derived from the study with cotinine levels available.

There are no “control” data to enable comparison with similar contemporary communities without smoke-free ordinances. However, this was presented in some of the individual studies to support the positive impact of legislation.3-5 Longer-term data could also have permitted examination of the possible effect of other factors that might have had an impact on secular changes, such as the use of important therapies, including 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors.

The outcome examined in the meta-analysis was hospitalization for acute myocardial infarction. Coronary plaque rupture and thrombosis are critical to this. Reduction in out-of-hospital deaths has also been demonstrated,6 and those studies with a broader range of end points have shown similar effects to those on incident myocardial infarction.5-7

The term “secondhand smoke” captures the involuntary nature of exposure.8 This emanates from both “sidestream”...
smoke from a smoldering cigarette and “mainstream” smoke exhaled by a smoker. Mainstream smoke is the more important. Secondhand smoking is a leading cause of preventable death, and the 2006 US Surgeon General’s report was entirely devoted to it. The report concluded that the evidence was sufficient to support immediate adverse effects on the cardiovascular system and, importantly, to allow deduction of a causal relationship between secondhand smoke and coronary heart disease (CHD) morbidity and mortality. It was also concluded that the evidence was suggestive of, but not sufficient to infer, a causal relationship with subclinical vascular disease, particularly carotid intimal-medial thickness and with the increased risk of stroke. Although exposure to secondhand smoke has declined over recent decades, the report estimated that 60% of American nonsmokers had biological evidence of exposure to secondhand smoke. The National Health and Nutrition Examination Survey measured cotinine levels among nonsmokers aged 20 years and over in 1999 to 2000, and it was found that 46% living in jurisdictions with smoking legislation had detectable cotinine levels (≥0.05 ng/mL), compared with 13% in jurisdictions without such laws.

Secondhand smoke has been estimated to cause a 25% to 30% increase in risk of CHD events. Exposure has often been estimated by the number of cigarettes smoked daily by a spouse or partner. Even for those with low to moderate exposure (1 to 14 or 1 to 19 cigarettes daily), the relative risk was 1.16 in comparison with those without such exposure. The evidence supporting these associations has been derived from both cohort and case-control studies, which have ranged in follow-up to 20 years, with different exposure measures and variability in the other factors that were controlled for. The association was slightly stronger in the overview of case-control studies, perhaps partly reflecting the bias associated with retrospective recall. There may be other biases. Although exposures may be misclassified, failure to account for background secondhand smoke and the use of only whether or not a spouse of a nonsmoker was a smoker would lead to underestimation of the strength of the association with outcomes. Indeed, estimates of the effects were 2-fold greater when based on cotinine levels, which capture exposure in the range of environments. Whincup et al demonstrated a dose-response relationship between serum cotinine and CHD events in a 20-year prospective study and found that risk was increased by 57% in individuals in the highest quartile of events in a 20-year prospective study and found that risk was associated with active smoking and the fact that the larger than anticipated, based on the 80% increase in relative risk associated with active smoking and the fact that the measured exposure to tobacco smoke to which nonsmokers are exposed is only about 1% of that from smoking 20 cigarettes daily. However, such large effects are biologically plausible and are consistent with nonlinear effects of tobacco exposure at low doses, including important actions on platelet and endothelial function. The effects on platelet and endothelial function, arterial stiffness, oxidative stress, and inflammatory markers are approximately 80% to 100% of those associated with active smoking. Other deleterious effects of secondhand smoking include those on matrix metalloproteinases, which might contribute to instability and rupture of atherosclerotic plaques; high-density lipoprotein cholesterol, and mitochondrial energy utilization.

The effects not only are large but occur rapidly. An important early study showed that platelet activation and aggregation and endothelial cell damage occurred in nonsmokers within 20 minutes of exposure to secondhand smoke, with no further activation among active smokers. Similarly, 30 minutes of breathing secondhand smoke caused endothelial dysfunction to a degree similar to that in active smokers. The effects of endothelial function may be slow to recover after long-term higher levels of exposure to secondhand smoking ends. However, there will be a rapid decrease in platelet aggregation, a key factor in acute coronary syndromes.

Exposure to secondhand smoke can occur in several environments, particularly in the home and workplace but also in restaurants, bars, gambling venues, and automobiles. It is not shared equally. Women are less often active smokers but sustain most of the burden of secondhand smoking. Effects may be particularly hazardous for children. They have smaller airways, breathe more quickly, and take in 3 to 4 times as much air and, potentially, secondhand smoke relative to their body weights as adults. The association with subclinical atherosclerosis in children is unproven. However, a relationship with carotid intimal-medial thickness has been shown in the Atherosclerosis Risk in Communities Study and children have many years to manifest a disease such as atherosclerosis, which has a long latency period. Disturbingly, the declines in cotinine levels in the time intervals between 1988 to 1991 and 1999 to 2002 have been less among children than adults. In 1999 to 2002, 59.6% of American children aged 3 to 11 years had cotinine levels ≥0.05 ng/mL, and their median cotinine concentration was 0.09 ng/mL, compared with 0.035 ng/mL in older adults. Nonsmokers in lower socioeconomic groups are also vulnerable because of higher active smoking rates and other environmental conditions to which they are exposed. In indigenous populations with high smoking prevalence, reduction in smoking, which would extend to secondhand smoke, may be the single most important short-term action to improve their life expectancy.

The World Health Organisation Framework Convention on Tobacco Control had 166 parties as of July 2009. Protection from secondhand smoke is among the 6 most important and effective policies outlined in the Framework, together with raising taxes and prices, health warnings, QUIT programs, banning of advertising and sponsorship, and careful surveillance of the tobacco epidemic and prevention policies. This should include cotinine or alternative biomarker data if possible. The present meta-analysis strengthens the evidence base to support laws promoting smoke-free environments. The effects are significant and, although immediate, increase with time. It is also important that the very high rates of CHD magnify the public health impact of secondhand smoke to underscore the importance of legislation. The California Environmental Protection Agency estimated that in the United States alone, in 2005 secondhand smoke resulted in
46,000 deaths due to CHD, compared with 3400 due to lung cancer in adult nonsmokers and 430 deaths related to sudden infant death syndrome.18

Such regulation can also have an impact on active smoking. A systematic review of 26 studies showed that smoke-free workplaces reduced smoking prevalence by 3.8% and the amount smoked by 3.1 cigarettes daily in those continuing to smoke, together constituting a 29% decrease in total cigarette consumption.19 One of the studies included in the meta-analysis2 found that acute coronary syndrome admissions were decreased in smokers as well as nonsmokers.7 Furthermore, rather than having a negative impact on businesses, smoking bans can increase patronage of restaurants and drinking venues.

Clinicians should advise their patients to avoid public places that permit smoking, and families should be counseled not to smoke at home or in a vehicle with patients. Healthcare professionals can also be powerful advocates, and research such as that described in this issue2 strengthens the case for government action.

Source of Funding
A.B. is a recipient of National Health and Medicine Research Council Scholarship ID No. 465352.

Disclosures
None.

References

Key Words: Editorials, myocardial infarction, prevention, smoking
APPENDIX 2

Background to the MCCS - adapted from “The Melbourne Collaborative Cohort Study Databook” (Volume 1: Clinical)

The study participants

A total of 41,514 people were recruited: 24,469 women and 17,045 men. The study investigators decided to seek volunteers instead of random recruits, given that volunteers would be more likely to remain interested in the study. The participants were aged between 27 and 80 years at recruitment with almost all between 40 and 69 years.

Focusing on individuals 40 to 69 years of age allowed aspects of diet and other lifestyle habits to be examined in middle age and as the participants moved into late middle age and old age. In these databooks, results are grouped into the following age categories: <40 (n=213), 40–49 (n=12,994), 50–59 (n=13,518), 60–69 (n=14,704) and 70 (n=85).

All were originally from metropolitan Melbourne, although some participants have since moved interstate and overseas – and continue to be contacted for follow-up.

Most of the men and women recruited to the study were Australian-born, but close to one in three of all the people in the study were born outside Australia. They are migrants from Greece, Italy, Malta, England, Ireland, Scotland, Wales and New Zealand. In these databooks, they are grouped by region: Australia/New Zealand, Northern Europe and Southern Europe.

Why are one-quarter of study participants Southern European migrants?

The MCCS is one of the few multi-ethnic prospective studies of diet and lifestyle currently underway in the world and the only study of its kind in Australia.

About one-quarter of the people recruited are migrants from Southern Europe: especially Italy and Greece. This is one of the factors that make MCCS an exceptional study. We wanted to include a large proportion of people from Southern Europe so the study would include people with an extended range of dietary and possibly other factors. We also wanted to include a wider genetic variation in the sample than we would have found among only non-immigrant Melburnians.

When the MCCS began in 1989, there was some evidence that migrants from Southern Europe were less likely than their Australian-born counterparts to develop cancer or heart disease, had lower morbidity levels than people born in Australian, and lived on average four or five years longer. The MCCS was devised partly to test hypotheses about the role of the ‘Mediterranean diet’ in disease development.
The inclusion of people from Southern Europe also enables potentially valuable comparisons of factors such as obesity, alcohol use, cigarette smoking, physical activity, hypertension, diabetes, cholesterol, use of hormones in women and family history of cancer.

How participants were recruited

Recruiting such a large number of people is by no means a quick or uncomplicated task. We were pleased to have the support of representatives of the Italian and Greek groups included in the study. These included the Vice-President of the Commission for Ethnic Affairs, the Assistant Director for Government and Community Liaison and Community Relations/Education, the coordinators of multicultural radio programs and journalists from the major multicultural papers. The recruitment took four years, from 1990 to 1994.

Many people were recruited through a letter of invitation, with details taken from the electoral rolls and phone books. We also used advertisements and announcements on multicultural radio and in newspapers, clubs and churches, to attract the interest of migrants who had not taken out citizenship and therefore would not be on the electoral rolls.

Another important way of recruiting people was to invite people who took part in the study to encourage friends and family to join.

The study also recruited using a mobile unit set up at major festivals, including the Lygon Street festival and the Aegean festival. Major manufacturing companies which employed a large number of migrants were contacted and asked if they would allow the unit to be located at the work premises for a time. Approaches were also made to have the unit placed in the centre of Melbourne, at the Bourke Street Mall.

How the information was collected

Each participant was given a unique identifying code upon providing personal information at reception. Participants were read a plain language statement and a relative/friend contact form, to assist with follow up in later years. The study was reviewed and approved before commencement by the Human Research Ethics Committee of the Cancer Council.

Face-to-face interviews were conducted with all participants. The interviews included:

- health and lifestyle questionnaire (see below)
- clinical data (see below)
food frequency questionnaire (see below).

The questionnaires were rigorously reviewed in development to ensure that every question could be justified. The researchers wished to keep the questionnaires as brief as possible, given that a long questionnaire would be arduous for participants. The same strenuous review applied to the selection of clinical measures.

The questionnaires as far as possible were interviewer administered and suitable for optical scanning.

Participants often found the interview process interesting and useful in itself and were pleased to receive a free check of their cholesterol and blood glucose, a blood pressure reading and a weight assessment.

What baseline measures were taken?

**Health and lifestyle questionnaire**

All volunteers answered an interview-administered questionnaire that included questions on:

- personal medical history, including questions about asthma, angina, hypertension, diabetes, arthritis/rheumatism, cancer, kidney stones, gallstones, heart attack and stroke
- weight
- family history of heart disease, cancer, diabetes and stroke
- smoking
- alcohol
- physical activity: walking and vigorous exercise
- other details, including occupation, education, household size and social life.

Women in the study were asked about:

- menstrual history
- pregnancies
- breastfeeding
- use of the Pill and other hormones
hysterectomy.

**Clinical data**

Clinical data were collected, including:

- blood pressure measurement
- blood sample
- measurements of standing height, weight (minus shoes and jackets), waist and hips circumference (over light clothes), bra size (for women), and lean and fat mass using bioelectric impedance.

This part of the study was administered by a nurse, who also took the drug and medical history described above.

**Diet**

A third element of the data collection was the questions on eating habits and the food frequency questionnaire.

There were 13 questions relating to dietary habits including the types and frequency of fats and oils used, and the use of milk, sugar, garlic and diet supplements.

The food frequency questionnaire was developed for the MCCS and was based on eight days of weighed food records from a large group of Australian, Italian and Greek-born people recruited for the pilot studies. People were then asked to report the frequency with which they ate foods from the categories:

- cereal foods, cakes and biscuits
- dairy foods and eggs
- meat, poultry, seafood and mixed dishes
- soups, salads and cooked vegetables
- fruit: dried, fresh, stewed and canned
- beverages and other food items.

**Why and how was blood collected?**
Nurses collected a sample of 15 ml of blood from each participant. Total plasma cholesterol and glucose were measured immediately. It was decided before the study began that participants would be informed if these tests showed elevated cholesterol or glucose levels and advised to see their doctor.

Samples of buffy coats and plasma have been stored in liquid nitrogen for future analyses of DNA and other substances of interest, such as hormones and markers of diet.
APPENDIX 3

The Australian diabetes, obesity and lifestyle study (AusDiab) - methods and response rates
The Australian Diabetes, Obesity and Lifestyle Study (AusDiab)—methods and response rates

David W. Dunstan a,*, Paul Z. Zimmet a, Timothy A. Welborn b, Adrian J. Cameron a, Jonathan Shaw a, Maximilian de Courten a, Damien Jolley c, Daniel J. McCarty a, AusDiab Steering Committee

a The International Diabetes Institute, 250 Kooyong Road, Caulfield, Vic. 3162, Australia
b Department of Medicine and Department of Public Health, University of Western Australia, Perth, Australia
c School of Health Sciences, Deakin University, Melbourne, Australia

Received 2 November 2001; received in revised form 14 January 2002; accepted 15 February 2002

Abstract

The Australian Diabetes, Obesity and Lifestyle Study (AusDiab) addresses the urgent need for data on diabetes prevalence, risk factors and associated conditions in Australia. Here we describe the methods used and the response rates obtained. AusDiab was a population-based cross-sectional survey of national diabetes mellitus prevalence and associated risk factors in people aged ≥ 25 years, conducted between May 1999 and December 2000 in the six states and the Northern Territory of Australia. The study involved an initial household interview, followed by a biomedical examination that included an oral glucose tolerance test (OGTT), standard anthropometric tests, blood pressure measurements and the administration of questionnaires. Of the 20,347 eligible people (aged ≥ 25 years and resident at the address for ≥ 6 months) who completed a household interview, 11,247 (55.3%) attended for the biomedical examination. Of those who completed the biomedical examination 55.1% were female. Comparisons with the 1998 Australian population estimates showed that younger age responders were under-represented at the biomedical examination, while the middle-aged and older age groups were over-represented. Weighting of the AusDiab data for age and gender have corrected for this bias. AusDiab, which is the largest national diabetes prevalence study undertaken in a developed nation to have used an OGTT, provides a valuable national resource for the study of the prevalence and possible causes of diabetes, as well as identifying possible risk factors that may lead to diabetes. Furthermore, it generates the baseline data for a prospective 5-year cohort study. The data will be important for national and regional public health and lifestyle education and health promotion programs. © 2002 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: AusDiab; Response rates; Diabetes survey; Australia; Diabetes prevalence

1. Introduction

Globally, the prevalence of diabetes, particularly Type 2 diabetes is rapidly increasing [1].
Indeed, it has been predicted that the global figure of people with diabetes will rise from current levels of about 150 million in 2000 to 300 million by 2025 [2]. However, with the exception of the USA [3], nationally representative, population based diabetes prevalence data among developed nations is scarce. In particular, few studies have involved an oral glucose tolerance test (OGTT).

In Australia, estimates of diabetes prevalence and other categories of glucose intolerance are confined to studies conducted 10–20 years ago on a small sample of residents from a rural town in Western Australia [4]. Most recent estimates of diabetes prevalence in Australia have relied on self-reported data, but since Type 2 diabetes can be asymptomatic for many years before it is diagnosed in a clinical situation, reliance on self-reported information invariably contributes to an underestimation of the true prevalence. Furthermore, such studies fail to provide information on the extent of other states of glucose intolerance, which are known to substantially increase risk of future diabetes.

To address the urgent need for more definitive data on the true current prevalence of diabetes and its associated risk factors in Australia, the Australian Diabetes, Obesity and Lifestyle Study (AusDiab) was a cross-sectional study involving a standard OGTT conducted during 1999–2000 in all Australian States and the Northern Territory. The present paper provides a detailed description of the survey methods including the design, sampling techniques and survey protocols. Data on weighting of the sample, response rates and statistical techniques are also presented. The survey methods conform to those recommended by the World Health Organisation (WHO) [5], and the study was approved by the International Diabetes Institute ethics committee.

The AusDiab study aimed to determine the national prevalence of diabetes and other selected non-communicable diseases and their risk factors in a representative sample of adults aged 25 years and over from each of the states and the Northern Territory of Australia.

More specifically, the objectives of the study were to:

1. estimate the national and regional prevalence of diabetes and other forms of abnormal glucose tolerance;
2. estimate the prevalence of the cardiovascular risk factors within the Metabolic Syndrome, including obesity, hypertension, and lipid profile abnormalities;
3. assess the distribution and relationships of the cardiovascular risk factors indicated above;
4. assess temporal trends in risk factor prevalences with reference to previous Australian surveys;
5. describe health knowledge and attitudes and utilization of health services, and
6. provide baseline data for longitudinal cohort studies.

2. Methods

2.1. Target population/eligibility requirements

Non-institutionalised adults aged 25 years and over residing in private dwellings in each of the six states and the Northern Territory of Australia were included in the survey if they had resided permanently at the address for a minimum of 6 months prior to the survey. Persons with physical or intellectual disabilities that precluded participation in the study were not included.

2.2. Sampling frame

A stratified cluster sampling method was used, involving seven strata (six states and the Northern Territory) and clusters based on census collector districts (CDs—the smallest geographic unit defined by the Australian Bureau of Statistics at each census, with an average of 225 dwellings each). Within each state, 6 CDs were randomly selected with a selection probability proportional to the population size (population aged over 25 years). Due to the logistic and economic constraints of the survey, and to avoid the bias of including an unrepresentative number of high prevalence groups, the following exclusion criteria were adopted:
1. CDs containing fewer than 100 persons aged 25 years and over
2. CDs that formed part of a statistical local area that was classified as 100% rural according to 1996 census data [6]
3. CDs that contained more than 10% indigenous population.

Of the total pool of CDs available (34 410), 4141 CDs (12%) were excluded from selection on these grounds. From the excluded CD’s, 762 (18.4%) had >10% indigenous population, 1464 (35.4%) were rural, 1100 (26.6%) had <100 persons aged ≥25 years, while 815 (19.7%) had more than one factor of the exclusion criteria. The three exclusion categories meant that the total eligible population (adults aged ≥25 years) was reduced by 6.44% from 11 341 070 to 10 610 855. This comprised 241 931 (33.1%) adults from CDs that had >10% indigenous population, 349 716 (47.9%) adults from CDs that were rural, 74 723 (10.2%) adults from CDs that had <100 people aged ≥25 years and 63 845 (8.7%) adults from CDs that had more than one factor of the exclusion criteria.

2.3. Sample size determination

The sample size was selected based on precision of estimates to identify a national diabetes prevalence of 7.0% (an estimation based on results of previous surveys, and the expectation that the diabetes rate had increased over time). As a secondary objective of the study was to deliver useful state-specific prevalence estimates, the sampling frame was stratified at the state level. With very little loss of efficiency, an accurate national estimate can be obtained from weighted samples of equal size from the six states and the Northern Territory. Accounting for the clustering of the survey design, a sample size of 10 500 (1500 per state) was predicted to provide 95% confidence intervals of 6.2–7.8, around a diabetes estimate of 7.0%. This level of precision was regarded as acceptable, and the sample size was considered achievable and within the funding constraints of the survey. It should be noted however, that the sample size was calculated for total diabetes prevalence only and would be expected to have limited power to describe the prevalence of type 1 diabetes in this sample.

2.4. Sample selection

It was calculated that 6 CDs were required to provide the required sample size (1500 per state) within each state. Following an initial field visit, if the CD was considered inappropriate for sampling in that location, the selected CD was replaced with another randomly selected CD from the same state. Replacements occurred in seven instances during the course of the survey, for the following reasons:
1. The low population density of the CD made it economically and logistically impossible to conduct the survey activities within the allocated timeframe (3 CDs)
2. The area selected was predominantly an industrial/business zone (2 CDs)
3. No eligible ‘neighbouring’ CD was available (see below) (1 CD)
4. The area had been recently involved in a large-scale health survey, including diabetes testing (1 CD).

After the first three sites had been surveyed, it became clear that a single CD would not provide the required sample size at each location surveyed. Clusters were subsequently formed by combining the randomly selected index CD and its largest adjoining neighbours to achieve a minimum cluster size of 250 participants. The final sample comprised 3 single CDs, 22 pairs of CDs, 16 triplets and 1 quad.

2.5. Survey protocol and procedures

The AusDiab survey activities occurred over a 21-month period between May 1999 and December 2001. Approximately 2 months were allocated to the collection of data in each state and the Northern Territory. The AusDiab survey activities were divided into two phases—the household interview and the biomedical examination.

2.5.1. Household census and interview

Following a local media advertising campaign involving news items in local community newspa-
pers and local radio and/or television, all private dwellings within the sampled cluster received a hand-delivered (non-addressed) letter informing residents about the survey and advising that an AusDiab interviewer would visit to conduct the household interview. A brochure describing the study objectives, the interview and examination process, and study confidentiality was supplied in the initial contact letter. This brochure was provided only in English.

The first visit by the interviewer occurred approximately 3 days after the letter had been delivered. If the interviewer could not make contact with household members, a letter was left requesting the household to telephone a toll-free number to arrange a suitable interview time. The interviewers made a minimum of 2 visits and up to 5 visits before a household was classified as a non-contact.

Where possible, at each participating household a personal interview was conducted with every adult member aged 25 years and over who met the eligibility requirements. The interview ascertained marital status, level of education, date and country of birth, language spoken at home and history of diabetes or high blood sugar levels. In some instances, adult household members were unable to answer for themselves because of old age, illness, intellectual disability or difficulty with the English language. In these cases, a responsible ‘proxy’ was interviewed on their behalf. There were no provisions for interviews to be conducted in languages other than English. In order to obtain a personal interview with all eligible household members, interviewers made appointments to visit as often as was necessary to the household. In a small number of cases interviews were conducted over the telephone with the Household Survey Coordinator.

At the completion of the interview, all household members aged 25 years or older were invited to attend a local test site for the biomedical examination. Participants were provided with a brochure explaining the biomedical examination procedures, together with the self-administered SF-36 General Health and Well-Being questionnaire, which they were asked to complete and bring to their biomedical examination appointment.

2.5.2. Biomedical examination

The biomedical examination was conducted at a local test site on weekdays (except Friday) and weekend days over an 8-day period in each sampled area. Local survey sites included community centres, scout headquarters, sporting venues, church halls and schools. Survey activities at the testing site commenced at 7:00 a.m. and typically finished at 2:00 p.m. On average, approximately 40 participants attended daily.

All responders gave written informed consent to participate in the survey upon arrival at the testing site. The AusDiab biomedical examination protocol followed closely the WHO recommended model for diabetes and other non-communicable disease field surveys [5,7]. The components of the biomedical examination are shown in Table 1. Following the initial collection of the fasting blood sample, an OGTT was performed on all participants, except those on insulin or oral hypoglycaemic drugs or those who were pregnant. The OGTT was performed according to WHO specifications. Participants moved through the biomedical examination procedures in a circuit-like manner that took approximately 2.5–3 h to complete. The SF-36 and dietary questionnaires were self-administered, while all other questionnaires were interviewer administered. All data from the participant record forms were entered both electronically and manually.

3. Results

3.1. Survey response

Response rates to the household interview and the biomedical examination are shown in Fig. 1. In total, the AusDiab interviewers approached 25,984 households in the 42 selected clusters. Of these, 6,769 (26%) were classified as non-contacts. Reasons for non-contact (and hence non-participation) in the household interview included language difficulties (318 households), no access gained to the residence (e.g. because of dangerous dogs, security fences) (941), the householders not being contactable despite several attempts (5358), and other reasons such as drunkenness or disability of the householders (152).
Of the 19,215 residential properties where contact was achieved, 1,095 were excluded because none of the occupants met the residency criteria of the survey, and a further 991 were excluded because all of the residents in the household were less than 25 years of age. Of the remaining 17,129 eligible households, 5,178 refused to be interviewed and 472 were away for the duration of the study period, giving rise to a total of 11,479 households (70.2%) where an interview was achieved. Reasons for refusal included health concerns (486, 9.4%), being unable to attend because of work commitments (1,159, 22.4%), feeling they were too old to participate (368, 7.1%), medical problems (1,317, 25.4%), and ‘other’ reasons (1,848, 35.7%).

Assuming that the proportion of ineligible households was similar between the contacted (2,086/19,215 = 10.9% ineligible) and the non-contacted households, 49.6% (11,479/23,163) of eligible households participated in the household interview. The denominator here (23,163) is calculated as all private dwellings (25,984) minus all ineligible households (2,821), which is comprised of

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Measurement instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic characteristics</td>
<td>Age, sex, ethnicity</td>
<td>Household interview and interviewer-administered questionnaires at survey site</td>
</tr>
<tr>
<td></td>
<td>Socio-economic status (education, occupation, income)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diabetes status</td>
<td>Interviewer-administered questionnaires at survey site</td>
</tr>
<tr>
<td>Medical and family history</td>
<td>Family history (diabetes)</td>
<td>SF-36 Questionnaire</td>
</tr>
<tr>
<td></td>
<td>Chronic health conditions (cardiovascular disease, gout)</td>
<td>Interviewer-administered questionnaires at survey site</td>
</tr>
<tr>
<td></td>
<td>Women’s health</td>
<td>Anti-Cancer Council of Victoria Dietary Questionnaire (self-administered)</td>
</tr>
<tr>
<td>Life-style related factors</td>
<td>General health and well-being</td>
<td>Interviewer-administered questionnaires at survey site</td>
</tr>
<tr>
<td></td>
<td>Alcohol/tobacco</td>
<td>Anti-Cancer Council of Victoria Dietary Questionnaire (self-administered)</td>
</tr>
<tr>
<td></td>
<td>Physical activity</td>
<td>Anti-Cancer Council of Victoria Dietary Questionnaire (self-administered)</td>
</tr>
<tr>
<td></td>
<td>Diet</td>
<td>Anti-Cancer Council of Victoria Dietary Questionnaire (self-administered)</td>
</tr>
<tr>
<td>Health-behaviour related factors</td>
<td>Health knowledge, attitudes and practice data</td>
<td>Interviewer-administered questionnaires at survey site</td>
</tr>
<tr>
<td>Physical measurements</td>
<td>Height</td>
<td>Stadiometer</td>
</tr>
<tr>
<td></td>
<td>Weight</td>
<td>Beam balance scales</td>
</tr>
<tr>
<td></td>
<td>Waist and hip circumference</td>
<td>Tape measure</td>
</tr>
<tr>
<td></td>
<td>Body fat determination</td>
<td>Bioimpedance</td>
</tr>
<tr>
<td></td>
<td>Blood pressure</td>
<td>Dinamap/mercury sphygmomanometer</td>
</tr>
<tr>
<td></td>
<td>12-lead ECG</td>
<td></td>
</tr>
<tr>
<td>Blood measurements (fasting)</td>
<td>Blood glucose</td>
<td>Glucose oxidase</td>
</tr>
<tr>
<td></td>
<td>Blood lipids</td>
<td>Enzymatically—Olympus AU600 analyser</td>
</tr>
<tr>
<td></td>
<td>HbA1c</td>
<td>Boronate affinity high performance liquid chromatography</td>
</tr>
<tr>
<td>Urine measurements (spot morning sample)</td>
<td>Albumin</td>
<td>Immunoturbidimetric method—Olympus AU600 analyser</td>
</tr>
<tr>
<td></td>
<td>Creatinine</td>
<td>Olympus AU600 analyser</td>
</tr>
</tbody>
</table>
2086 contacted ineligible households, plus 735, which is an equivalent percentage of non-contacted ineligible households. This response rate is a conservative estimate, as more of the non-contacted households are likely to be unoccupied or have fewer occupants than contacted households.

In the 17129 households that were confirmed as containing at least one eligible participant, 20347 eligible adults were interviewed. Of those who participated in the household interview, 11247 (55.3%) took part in the biomedical examination. This response rate for the biomedical examination ranged from 49.5% in Queensland and 49.6% in South Australia to 61.8% in Western Australia (Table 2). Assuming that the numbers of eligible adults residing in the 5178 households that refused the household interview was the same as in those which participated, and combining the household response rate (11479/17129 = 67%) with the biomedical examination response rate (11247/20347 = 55.3%) the overall response rate can be estimated to be 37%.

3.2. Profile of responders and weighting of the survey sample

To account for the clustering and stratification of the survey design, and to adjust for non-response, the data have been weighted to match the age and gender distribution of the 1998 estimated residential population of Australia aged over 25 years [8]. The weighting factor is based on the probability of selection in each cluster. The number of males and females in each cluster aged 25 years and over identified in the 1996 census was used to calculate the probability of selection in each cluster. The weight was then calculated based on the probability of selection, adjusted to reflect the age and sex structure of the 1998 estimated residential population over the age of 25. Groups based on age deciles and gender defined the weighting variable. As there are two
distinct populations in our sample—one who participated in the household interview and a subset of this population who attended the biomedical examination, two weighting factors have been applied, one to all responders to the biomedical examination and another to all responders to the household interview.

Among the responders to the biomedical examination \((n = 11247)\), 44.9% were male, with the mean age being 51.5 years (Table 3). This compares to 49.0% male in the 1998 Australian population, and a mean age in the 1996 census [6] over 25 years of 48.1 years. Among the non-responders \((n = 9049)\), 51.2% were male and the mean age was 47.7 years. Weighting of the sample to the estimated 1998 residential Australian population corrected the gender and age bias, with 49.0% (95% CI, 47.9–50.1) of the weighted responders to the biomedical examination being male, and the mean age being 48.2 years (95% CI, 46.6–49.9).

Table 4 provides a comparison between responders and non-responders to the biomedical examination for both unweighted and gender and age-adjusted estimates, with respect to various demographic characteristics. For the crude, unadjusted estimates, significant differences were observed for percent married, the percentage of English speaking participants, the percent born in the UK, and the percentage who suspected they had diabetes, but no differences were noted for the percentage born in Australia, the percentage born outside the UK or Australia, the percentage who had completed the highest year of school, and the percentage who had ever been told they had diabetes. Adjustment for age and gender rectified the difference between responders and non-responders for the percentage married, but differences were still observed in the percentage who suspected that they had diabetes, the percent born in the UK and the percentage of English speakers. Additionally, after adjustment for age and gender, the percentage who had completed the final year of high school, technical education or University was higher for responders.

4. Discussion

AusDiab is the largest cross-sectional study of the prevalence of diabetes and its precursors ever performed in a developed nation. Through its capacity to provide the first definitive data on the true magnitude of the diabetes epidemic in Australia, AusDiab will not only be a valuable resource for health care planners in Australia, but will also serve as an important research tool for the study of diabetes and associated diseases on a longitudinal basis.

The AusDiab experience provides a valuable insight into the execution of population-based,
Table 3
Response rates of eligible residents to the biomedical examination by age and gender

<table>
<thead>
<tr>
<th>Gender/age group</th>
<th>Eligible residents</th>
<th>Respondents to household interview</th>
<th>Respondents to biomedical examination</th>
<th>Biomedical examination response rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male 25–34</td>
<td>1757</td>
<td>1747</td>
<td>590</td>
<td>33.6</td>
</tr>
<tr>
<td>Male 35–44</td>
<td>2342</td>
<td>2331</td>
<td>1093</td>
<td>46.7</td>
</tr>
<tr>
<td>Male 45–54</td>
<td>2290</td>
<td>2281</td>
<td>1345</td>
<td>58.7</td>
</tr>
<tr>
<td>Male 55–64</td>
<td>1516</td>
<td>1515</td>
<td>928</td>
<td>61.2</td>
</tr>
<tr>
<td>Male 65–74</td>
<td>1125</td>
<td>1122</td>
<td>731</td>
<td>65.0</td>
</tr>
<tr>
<td>Male 75+</td>
<td>677</td>
<td>677</td>
<td>362</td>
<td>53.5</td>
</tr>
<tr>
<td>Female 25–34</td>
<td>1894</td>
<td>1890</td>
<td>803</td>
<td>42.4</td>
</tr>
<tr>
<td>Female 35–44</td>
<td>2510</td>
<td>2503</td>
<td>1465</td>
<td>58.4</td>
</tr>
<tr>
<td>Female 45–54</td>
<td>2355</td>
<td>2352</td>
<td>1546</td>
<td>65.6</td>
</tr>
<tr>
<td>Female 55–64</td>
<td>1649</td>
<td>1647</td>
<td>1096</td>
<td>66.5</td>
</tr>
<tr>
<td>Female 65–74</td>
<td>1286</td>
<td>1285</td>
<td>837</td>
<td>65.1</td>
</tr>
<tr>
<td>Female 75+</td>
<td>927</td>
<td>927</td>
<td>451</td>
<td>48.7</td>
</tr>
</tbody>
</table>

* Nineteen eligible people refused to give their age and are thus missing from this table, of whom 16 were respondents to the household interview, and none were respondents to the biomedical examination.

b Calculated as biomedical examination responders as a percentage of eligible residents.

cross-sectional surveys involving the use of an OGTT. Since AusDiab required careful consideration of the logistics required to achieve a national sample within the funding and timeframe constraints imposed, particular emphasis was given to the establishment of a study design that reflected the ‘best available’ model. This extensive 12-month planning process was crucial to the successful implementation of the study.

Several aspects of the methods used in sample selection and the study design of AusDiab warrant further discussion. First, the inclusion criteria contained only those CDs that contained less than 10% indigenous population. Existing data provide clear evidence of a very high prevalence of diabetes among the indigenous population in Australia [9]. To overcome the chance selection of one or more CDs with a large proportion of indigenous people, and thus minimize the potential bias introduced to the national and state diabetes estimates, we considered it more practical to restrict the inclusion criteria to those CDs likely to contain smaller proportions of indigenous people rather than account for any potential bias at the analysis stage. Furthermore, this approach was considered important for the operations of the study, since aspects such as questionnaire design would have required extensive modifications to reflect the cultural differences. It is unlikely that this restriction would have impacted greatly on the generation of national estimates since the indigenous population is numerically a very small minority group in Australia (~2% of the total Australian population), and indeed, represented only 0.8% of the total AusDiab sample. Preparations are presently underway to address these issues through a survey that will employ similar survey methods used within AusDiab in urban indigenous Australians living in Darwin, Northern Territory.

The decision to sample equal numbers from each stratum reflects a compromise between the primary and secondary objectives of the survey. It is probable that a study design that sampled from the states proportional to their size would have been more efficient in terms of providing a more accurate national diabetes prevalence estimate, however accurate estimates for all states (in particular the smaller states) would have been compromised. Since weighting of the data prior to the analysis stage enables us to allow for over-representation of the smaller states and under-representation of the larger states, it is unlikely that our primary objective was compromised unduly by this decision.
Table 4
Comparison between biomedical examination responders and non-responders, both unweighted, and age and gender-adjusted to the 1998 estimated residential population aged over 25 years

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Responders to biomedical exam. (unweighted)</th>
<th>Non-responders to biomedical exam. (unweighted)</th>
<th>Responders (age and gender adjusted to 1998 population*)</th>
<th>Non-responders (age and gender adjusted to 1998 population*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>71.5 (68.8–74.2)</td>
<td>67.0 (64.4–69.5)</td>
<td>68.5 (65.5–71.5)</td>
<td>67.6 (65.0–70.2)</td>
</tr>
<tr>
<td>Country of birth: Australia</td>
<td>76.0 (72.9–79.1)</td>
<td>77.1 (73.8–80.5)</td>
<td>77.6 (74.6–80.6)</td>
<td>76.7 (73.3–80.1)</td>
</tr>
<tr>
<td>Country of birth: UK</td>
<td>11.3 (9.7–12.8)</td>
<td>8.7 (7.2–10.2)</td>
<td>10.3 (8.9–11.7)</td>
<td>8.8 (7.3–10.4)</td>
</tr>
<tr>
<td>Country of birth: Other</td>
<td>12.7 (10.2–15.2)</td>
<td>14.1 (10.9–17.3)</td>
<td>12.1 (9.7–14.5)</td>
<td>14.4 (11.1–17.7)</td>
</tr>
<tr>
<td>Language spoken: English</td>
<td>96.0 (94.6–97.4)</td>
<td>93.7 (91.2–96.3)</td>
<td>96.1 (94.9–97.4)</td>
<td>93.6 (91.0–96.2)</td>
</tr>
<tr>
<td>Education: completed high school/ university/ technical education</td>
<td>55.8 (51.4–60.1)</td>
<td>51.7 (46.4–57.1)</td>
<td>58.2 (54.0–62.4)</td>
<td>51.3 (46.0–56.6)</td>
</tr>
<tr>
<td>Ever told have DM?</td>
<td>6.4 (5.7–7.1)</td>
<td>6.2 (5.2–7.1)</td>
<td>5.9 (5.3–6.5)</td>
<td>6.4 (5.4–7.3)</td>
</tr>
<tr>
<td>Suspect have DM?</td>
<td>1.5 (1.3–1.7)</td>
<td>0.5 (0.4–0.7)</td>
<td>1.5 (1.3–1.7)</td>
<td>0.5 (0.4–0.7)</td>
</tr>
</tbody>
</table>

Estimates are percentages (95% CI).
* June 30, 1998 Australian population [8].
It is also noted that, due to the exclusion criteria of the study, the results may not be generalisable to either the indigenous population or the rural population of Australia. The primary aim of the study, however, was to provide estimates which were accurate for the Australian population over 25 years as a whole and these exclusion criteria should not significantly affect that aim.

The response rates to AusDiab can be interpreted in several ways. In many studies where a defined population is used as the sample pool, an absolute response rate can be accurately calculated. For example, when using an electoral role in the sample selection, the number of residents in each household is accurately known, allowing the demographic profile of both the non-responders and the responders to be calculated. In AusDiab, the sample pool was comprised of households in CDs based on the 1996 Australian census, conducted 2 years prior to commencement of the AusDiab survey. An accurate estimate of the number of residents in households where contact was not achieved, as well as the age and gender profile of these households cannot be accurately obtained. This is due to the possibility that in those households where contact could not be achieved, many may have been unoccupied, or the resident population in each household may be lower (assuming that the more people residing in a household, the more likely it is that someone will be home when an interviewer calls).

Our estimates suggest, however, that reasonably good response rates were obtained from those households where contact could be achieved. Furthermore, considering the duration and nature of the testing procedures involved in the biomedical examination for each individual, the response to the biomedical examination is acceptable. Nevertheless, additional in-depth analyses will be necessary to explore whether specific non-response biases exist at both the national and state level.

Regarding the analysis of non-response bias presented, there are several points worth noting. Firstly, the difference in the percentage of English speakers between respondents and non-responders shown in Table 4, while being significant, was fairly small (96.1 vs. 93.6%) and is unlikely to have had a significant impact on diabetes or other prevalence estimates. Similarly, the percentage of responders born in the UK was only slightly greater than the percentage of non-responders born in the UK (10.3 vs. 8.8%), although again, this difference was significant. Most of the difference in country of birth between responders and non-responders was removed by age and gender-standardisation. It is unlikely that the percentage of people born in the UK would have an important effect on diabetes prevalence estimates, since many cultural similarities exist for those born in the UK and those born in Australia.

The greatest differences between responders and non-responders were observed in suspicion of diabetes and level of education. Firstly, the percentage of those who suspected they had diabetes (but have never been told they do) was significantly higher in the responders (1.5%) compared to the non-responders (0.5%). Only one in 12 of those who suspected they had diabetes were actually found to have the disease, compared to one in 25 of those who did not suspect they had diabetes. Taking into account the very low prevalence of those who suspected they had diabetes, and the low prevalence of those found to actually have diabetes when they suspected they had diabetes, the difference between responders and non-responders with respect to suspicion of diabetes would have increased the total number of newly diagnosed cases of diabetes by 6 or 7 persons at most. This would be expected to have only a negligible effect on the total prevalence estimates for diabetes.

Participants who attended the biomedical examination were more likely to have completed the final year of high school, University or other higher education (58.2 vs. 51.3%) than non-responders. This would indicate that the higher socio-economic groups were over-represented in AusDiab. This difference could potentially bias estimates of diabetes, as well as other studied variables. However, for glucose intolerance, as well as other cardiovascular disease risk factors such as dyslipidaemia, physical activity, alcohol consumption and smoking, there is a negative association with socio-economic status [10]. Therefore, our estimates of these disease states, if a socio-economic bias does indeed exist, are likely
to underestimate the true prevalence. Of course, education level is only one indicator of socio-economic status, and other variables such as income level, occupation and type of residence will need to be considered in further analyses of response bias. Detailed comparisons between responders and the Australian population aged over 25 (using both census data and other previous surveys), particularly in the areas of socio-economic status, language spoken and suspicion of diabetes, will be valuable in assessing more precisely the impact of any response bias in the AusDiab survey.

AusDiab is a major achievement in the study of diabetes in Australia. The study not only provides much needed data on the current magnitude of the diabetes epidemic that exists in Australia but also fills a 10-year void in knowledge on current levels of many of the associated cardiovascular disease risk factors that can only be determined through blood collection. Furthermore, an important extension to this initiative will be the follow-up of the AusDiab cohort, that will provide the first opportunity ever in Australia to examine the natural history of diabetes and its complications, as well as the incidence of cardiovascular disease among this representative sample of Australians with diabetes or impaired glucose metabolism.

Acknowledgements

We are most grateful to the following for their support of the study: The Commonwealth Department of Health and Aged Care, Eli Lilly (Aust) Pty Ltd, Janssen–Cilag (Aust) Pty Ltd, Knoll Australia Pty Ltd, Merck Lipha s.a., Alphapharm Pty Ltd, Merck Sharp & Dohme (Aust), Roche Diagnostics, Servier Laboratories (Aust) Pty Ltd, SmithKline Beecham International, Pharmacia and Upjohn Pty Ltd, BioRad Laboratories Pty Ltd, HITECH Pathology Pty Ltd, the Australian Kidney Foundation, Diabetes Australia (Northern Territory), Queensland Health, South Australian Department of Human Services, Tasmanian Department of Health and Human Services, Territory Health Services, Victorian Department of Human Services and Health Department of Western Australia.

For their invaluable contribution to the field activities of AusDiab, we are enormously grateful to Annie Allman, Marita Dalton, Adam Meehan, Claire Reid, Alison Stewart, Robyn Tapp and Fay Wilson.

And our special thanks goes to the local collaborating centers, including Sir Charles Gairdner Hospital (Western Australia), the Prince of Wales Hospital (New South Wales), the Menzies Centre for Population Health Research (Tasmania), the Queen Elizabeth Hospital (South Australia), the Menzies School of Health Research (Northern Territory), Queensland Health, the Monash Medical Centre, Department of Nephrology (Victoria) and the Centre for Eye Research Australia (Victoria).

References