2020

An Epidemiological and Geospatial Study of Cardiometabolic Risk Factors in the Illawarra-Shoalhaven Region of NSW, Australia.

Renin Melkias Baby Selvi Toms
University of Wollongong

An Epidemiological and Geospatial Study of Cardiometabolic Risk Factors in the Illawarra-Shoalhaven Region of NSW, Australia.

Renin Melkias Baby Selvi Toms, MSc (CHN), MBA (HCS)

Supervisors:

Professor Andrew Bonney

Dr Darren John Mayne

Associate Professor Xiaoqi Feng

This thesis is presented as part of the requirement for the conferral of the degree:

DOCTOR OF PHILOSOPHY

This research has been conducted with the support of the Australian Government Research Training Program Scholarship

School of Medicine

University of Wollongong

March 2020

Abstract

Introduction: Cardiovascular disease (CVD) associated metabolic risk factors are a growing human health concern in Australia and worldwide. This thesis investigated the small-area geographic variation in the distribution of cardiometabolic risk factors (CMRFs) in the Illawarra-Shoalhaven region of Australia, their association with area-level disadvantage and access to primary care and whether area-level disadvantage and primary care access contribute to the geographic variation of CMRFs.

Materials and methods: Geographic variation in the distribution of individual CMRFs was analysed at Statistical Area Level 1 (SA1), which is the smallest unit that disaggregated census data are reported in Australia. Individual-level data used in this thesis included de-identified CMRF test data from non-pregnant adult (≥18 years) residents of the Illawarra-Shoalhaven region between 2012–2017, which was sourced from the largest pathology service provider in the study region. These data included the most recent individual-level test results for: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI); and diabetes mellitus (DM) status. The test results were dichotomised into higher and lower cardiometabolic risk values based on the existing clinical guidelines. Area-level data included: SA1-level disadvantage, sourced from the 2011 Australian Census of Population and Housing Index of Relative Socioeconomic Disadvantage; and primary care provider data retrieved from publicly available sources current in year 2016.

Choropleth maps describing the distribution of CMRFs rates were produced using an Empirical Bayes (EB) approach to smooth the rates. Spatial clustering of CMRFs was assessed using Moran's I test and Local Indicators of Spatial Autocorrelation (LISA). A two-step floating catchment area (2SFCA) method was used to calculate the primary care access index of the SA1s within the study region. Multilevel logistic regression models were used to elucidate the association of the area-level socioeconomic disadvantage and primary care access with the geographic variation of CMRFs in the study region, after adjusting for individual- and area-level covariates.

Results: Analysis of 1, 132, 016 pathology tests contributed by 256, 525 individuals revealed significant geographic variation, spatial autocorrelation and clustering of higher cardiometabolic risk findings at the SA1-level. Multilevel analyses revealed associations between area-level disadvantage and all higher risk CMRFs findings, after adjusting for individual-level covariates. Geographic access to primary care was inversely associated with higher risk levels of HDL and obesity in the study region after adjusting for the individual and area-level covariates but was not associated with the remaining CMRFs. The estimated proportions of the geographic variation in the higher cardiometabolic risk values explained by area-level disadvantage ranged from 14.3–57.8%, while geographic access to primary care explained ≤ 10.5%.

Conclusion: The findings support future investigations into whether geographically targeted public health activities or location-specific interventions in primary care can ameliorate CMRFs. The findings also call for universal approaches proportional to the need and disadvantage level of populations for the prevention and control of CMRFs. These findings can be used to inform regional health care service commissioning and related policy developments, and are highly relevant in the context of the global paradigm shift from communicable diseases to cardiovascular diseases (CVD) as the leading cause of human death and health care expenditure.

This thesis is dedicated to the people of the Illawarra-Shoalhaven region in Australia, whose de-identified data were used in this study

Acknowledgment

It would have been hardly possible for me to survive the hard work, consistency and dedication required in this thesis without the help of a few precious personalities who directly or indirectly instrumented this journey. Here is my humble tribute to all those people.

First of all, I wish to acknowledge my incredible supervisory team for their unwavering support, guidance and encouragement through a path that felt implausible at times. I wish to thank my primary supervisor ***Professor Andrew Bonney*** (School of Medicine, University of Wollongong) for introducing me into the world of multilevel analyses in population health research. It was only due to his extraordinary wisdom, cheerful enthusiasm and ever friendly nature that I was able to complete this thesis in a timely manner. I would like to extend my gratitude to co-supervisors ***Dr Darren John Mayne*** (Epidemiologist, Illawarra Shoalhaven Local Health District, NSW Health) and ***Associate Professor Xiaoqi Feng*** (School of Public Health and Community Medicine, University of New South Wales). Without Mr Mayne's able guidance I would not have been able to learn the fundamentals of the complex statistical analyses required in this thesis. I was able to communicate with him on a routine basis throughout this thesis, which helped me to learn many valuable lessons that constantly supported me throughout this journey. I am equally thankful to Associate Professor Feng, whose positive spirits, guidance and encouragement had always pushed me forward, especially in the moments I was stuck. Her mentorship continues to guide me especially in the professional paths which I am pursuing in relation with this research area.

This thesis would not have been possible without the de-identified data of the study population. I sincerely thank the data providers of this thesis, ***Southern IML Pathology and staff*** for their generosity in providing data for the SIMLR Cohort Study and ongoing support. In particular I would like to thank ***Mr Bryan Jones*** for providing the technical expertise for the data acquisition and helpful comments on the manuscript. Southern IML Pathology are the owners of the data contained within this thesis and the Illawarra Health and Medical Research Institute (IHMRI) is the custodian facilitating access to the data.

I would like to acknowledge the ***University of Wollongong*** for the infrastructure and facilities provided to me and the support of the ***Australian Government Research Training Program Scholarship*** which helped to make this thesis possible.

I am extremely thankful to my friend ***Deepti Vinod Panicker***, whose selfless support and sisterly care helped and motivated me to survive during the most difficult times in this journey. I am also privileged to have my friend and colleague ***Ramya Walsan*** who had provided great company through the up and downs of this thesis path.

Finally, I would also like to take this opportunity to thank my family for their support at different phases of this journey. Sadly my father ***Toms Melkias (****Science teacher****)*** and brother ***Sgt Bazanio Jose (Indian Air Force)*** did not survive to see this thesis, but I know they would have been immensely proud of their little girl's achievement. My mother ***Baby Selvi Melkias*** (Rtd Rural Welfare Officer, Government of India) who, while overseas, talked to me and encouraged me every day. I am also thankful to my sisters ***Sherin Toms*** and ***Merlin Toms*** for their prayerful support while writing this thesis.

To my husband ***Dr Ginu Rajan***, you have been a constant source of support. I know how proud you are of me and I love that you speak as enthusiastically about my research as you do about yours. I should also be thankful

to the cheerfulness of my toddler daughters, ***Angelina*** and ***Isabella***. Whenever things were difficult, I would think for both of you and remind myself that I cannot give up being role model for you both. This thesis proves you can do anything you put your mind to, with the right support and people around you. I hope one day this would make you proud of your mum and inspire you to do your best in whatever you pursue.

*Above all, within me I know that it is the will of **Almighty** that took me forward and lighted my paths and strengthened my spirits for which I am grateful and would remain humble and dutiful always.*

Thesis Style

This thesis is presented as a compilation of manuscripts written for publication. A *General Introduction* chapter is written in the beginning and a *General Discussion* chapter is written at the end to provide a critical overview of the thesis. *A Materials and Methods* chapter is also included describing the methodologies adopted in individual studies. All the remaining core chapters of this thesis represent individual manuscripts as prepared for a journal, with specific objectives, methods, results and conclusions. The structure of each chapter is therefore consistent with the requirement of the journal for which it is written.

Certification

I, **RENIN MELKIAS BABY SELVI TOMS**, declare that this thesis submitted in fulfilment of the requirements for the conferral of the degree, **Doctor of Philosophy** in the School of Medicine from the University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

_____

Renin Melkias Baby Selvi Toms

27h March 2020

List of Publications

Peer reviewed articles

1. **Toms, R**, Bonney A, Mayne D J, Feng X, Walsan R. *Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature*. **International Journal of Health Geographics**, 2019. 18(1), 1.
   Available from: https://link.springer.com/article/10.1186/s12942-018-0165-5.
   DOI: https://doi.org/10.1186/s12942-018-0165-5
2. **Toms R**, Mayne DJ, Feng X, Bonney A. *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256, 525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia*. **PLOS ONE**, 2019. 14(10): e0223179.
   Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223179
   DOI: https://doi.org/10.1371/journal.pone.0223179
3. **Toms, R**., Mayne, D. J., Feng, X., & Bonney, A. *Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia*. **Nature - Scientific Reports**, 2020. 10(1), 1-18.
   Available from: https://www.nature.com/articles/s41598-020-69552-4
   DOI: https://doi.org/10.1038/s41598-020-69552-4
4. **Toms, R**., Feng, X., Mayne, D. J., & Bonney, A. *Role of area-level access to primary care on the geographic variation of cardiometabolic risk factor distribution: a multilevel analysis of the adult residents in the Illawarra—Shoalhaven Region of NSW, Australia*. **International journal of environmental research and public health**, 2020. 17(12), 4297.
   Available from: https://www.mdpi.com/1660-4601/17/12/4297
   DOI: https://doi.org/10.3390/ijerph17124297

Peer reviewed abstracts

5. **Toms, R**., Bonney, A., Mayne, D. J. & Feng, X. *Geographic variance in distribution of cardiometabolic risk factors in the Illawarra-Shoalhaven region of the NSW, Australia*. **15th WONCA World Rural Health Conference,** 2018. New Delhi, India. (pp. 1-11).
   Available from: https://scholars.uow.edu.au/display/publication130220
6. **Toms, R**., Bonney, A., Mayne, D. J., & Feng, X. G*eographic distribution of cardiometabolic risk factors: a small area level approach.* (Poster presentation) Public Health Prevention Conference – 2018, **Public Health Association of Australia (PHAA)**, 4-5 May 2018. Sydney, Australia.
   Available from: https://scholars.uow.edu.au/display/publication130221
7. **Toms, R**., Bonney, A., Mayne, D. J., & Feng, X. *Area-level socioeconomic disadvantage and cardiometabolic risk distribution: an analysis of 256, 565 adult residents in the Illawarra- Shoalhaven region of the NSW Australia*. (Poster presentation). 8th Annual NHMRC Symposium on Research Translation. **National Health and Medical Research Council**. 19 - 20 November 2019, Melbourne, Australia.
   Available from: https://scholars.uow.edu.au/display/publication140392

Related publications

8. Walsan, R., Bonney, A., Mayne, D. J., Pai, N., Feng, X., & **Toms, R**. Serious Mental Illness, Neighborhood Disadvantage, and Type 2 Diabetes Risk: A Systematic Review of the Literature. **Journal of Primary Care & Community Health**, 2018 Jan-Dec; 9 : 2 1 5 0132718802025.
   Available from: https://journals.sagepub.com/doi/10.1177/2150132718802025.
   DOI: https://doi.org/10.1177/2150132718802025
9. Walsan, R., Bonney, A., Mayne, D. J., Pai, N., Feng, X., & **Toms, R**. (2018). Neighbourhood Disadvantage and Type 2 Diabetes Comorbidity in Serious Mental Illness: A Systematic Review of Literature. (Poster presentation). **The Mental Health Services Conference**, 2018, Adelaide, Australia.
   Available from: https://scholars.uow.edu.au/display/publication130397

Media/Press

10. "*What your address tells us about your health*" - Researchers investigate how cardiometabolic risk varies within the Illawarra-Shoalhaven region. IHMRI News release on 28th August 2020
    Available from: https://www.ihmri.org.au/what-your-address-tells-us-about-your-health/

Funding and Awards

- *2016 – 2020:* ***Australian Government Research Training Program Scholarship***
- *2018: Faculty of Science Medicine and Health,* ***HDR travel grant for oral presentation*** *in an international conference.*
- *2019: Faculty of Science Medicine and Health,* ***HDR travel grant for poster presentation*** *in a domestic conference.*
- *Health Impact Research Centre (HIRC),* ***HDR Publication Prize 2020*** *-* ***Toms, R****., Mayne, D. J., Feng, X., & Bonney, A.. Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia.* ***Nature - Scientific Reports****, 2020.10(1), 1-18.*

Statement of Contribution

This statement verifies that the greater part of the above-named manuscripts is attributed to the PhD candidate, *Renin Melkias Baby Selvi Toms*. Under the guidance and supervision of her supervisors, she took the primary responsibility for the design of each study, access to the data, statistical analyses, prepared the first draft of each manuscript, and prepared the papers for submission to relevant journals. Co-authors, who were also supervisors to the candidate, contributed to the thesis by providing guidance on the design and structure of each study, and provided editorial suggestions for every paper.

20/March/2020

Ms Renin Melkias Baby Selvi Toms (PhD Candidate/Main author)

20/March/2020

Ms Ramya Walsan (Co-author)

23/March/2020

Dr Darren Mayne (Co-supervisor)

23/03/2020

Associate Professor Xiaoqi Feng (Co-supervisor)

20/March/2020

Professor Andrew Bonney (Primary Supervisor)

Table of Contents

List of Names or Abbreviations

A ABS - Australian Bureau of Statistics; ACR - Albumin creatinine ratio; AIC - Akaike Information Criterion; ASED - Area socioeconomic disadvantage; AU – Australia; ASGS - Australian Statistical Geography Standard

B BMI - Body mass index; BP – Blood pressure; BS – Blood sugar

C CD -Census Collection District; CHD - Coronary heart disease;

CKD - Chronic Kidney Disease; CMRFs - Cardiometabolic Risk Factors; CVD - Cardiovascular disease; CVH – Cardiovascular health

D DM - Diabetes Mellitus

E eGFR - estimated Glomerular filtration rate; EB - Empirical Bayes; ESRI - Environmental Systems Research Institute

F FBG – Fasting blood glucose; FBSL - Fasting blood sugar level;

FI – Fasting insulin; FPG – Fasting plasma glucose

G GFR – Glomerular filtration rate; GIS - Geographic Information System;

GR- Germany; GPs - General practitioners.

H HbA1c - Glycated haemoglobin; HDL- High density lipoprotein;

HR – Heart rate; HREC - Human Research Ethics Committee; HT – Hypertension.

I ICC - Intra-cluster Correlation Coefficient; IR – Insulin resistance; IRIS - Ilôts regroupés pour l'information statistique ; IRS – Insulin resistance syndrome; IRSD – Index of Relative Socioeconomic Disadvantage ; ISLHD - Illawarra and Shoalhaven Local Health District ; ISR - Illawarra-Shoalhaven region

L LDL – Low density lipoprotein; LGA - Local Government Area; LISA- Local Indicators of Spatial Autocorrelation

M M – Model; MOR - Median Odds Ratio; MEDLINE – Medical Literature Analysis and Retrieval System Online

N NSW – New South Wales

O OR – Odds Ratio

P PCV - Proportional change in Variance; PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses; POA - Postal Area; PLOS - Public Library of Science

Q Q - Quintile

R RACGP - Royal Australian College of General Practitioners;

S SA1 – Statistical Area Level 1; SES - Socioeconomic status; SEIFA- Socio-Economic Indexes for Areas; 2SFCA - Two-step floating catchment area; SIMLR -Southern IML Research; SLA -

Statistical Local Area; SES - Socioeconomic status; strobe - Strengthening the reporting of observational studies in epidemiology

T τ^2 - Area level variance; TC- Total cholesterol; TCR – Total cardiometabolic risk; T2DM – Type 2 diabetes mellitus; TG –Triglycerides; TRIRIS - Groups of around three IRIS areas.

U UOW - University of Wollongong

W WC – Waist circumference

List of Figures and Illustrations

List of Tables

Chapter 1

General Introduction

Chapter 1: General Introduction

"If you want to learn about the health of a population, look at the air they breathe, the water they drink and the places where they live."

(Hippocrates 5th Century BC)

1.1 Introduction

Place has always been a key element in human health and disease. Written documents on the link between place and health date back to Hippocrates (~ c. 460- c.370 BC), the father of medicine who stated "airs, waters, places" all have significant impact on human health and history.[1] In 1854, John Snow (1813-1858), the father of modern epidemiology, undertook his classic work on cholera epidemics by highlighting the importance of location on disease.[2] However, subsequent quantitative research developments on the influence of 'place on health' were slow and sporadic until recent advances in Geographic Information System (GIS) and related area-level analytical approaches.[3]

Advances in area-level analytical methods over the last one quarter of a century have provided powerful tools to help understand the impact of place on health.[4] For example, area level visualization and analyses of geocoded data became possible through geospatial analytical software such as ArcGIS, GeoDa and SatScan. More robust analytical techniques for nested data such as multilevel modelling (MLM) functionalities became widely available in popular data analytical software packages, which also mean their use is more widespread. These developments have the potential to yield the environmental context of the diseases and their risk factors. Meaningful application of the combination of these developments has the potential to become a powerful and evidence based tool to plan health care service commissioning at a regional level. This is important especially in the context of the global epidemiological transition from infectious diseases to non-communicable diseases (NCDs), specifically cardiovascular disease (CVD), as the leading cause of death and health care expenditure across the world.[5]

1.2 Background

Cardiovascular diseases are the leading cause of death worldwide and also the highest consumer of health care expenditure in many developed nations.[6, 7] Once diagnosed with CVD, the ongoing costs of care and productivity loss due to consequent disabilities and premature deaths create a major economic burden not only to the individual, but to the family and the community. These burdens get exacerbated when half the people dying from CVD were in their prime productive ages.[6]

In 2017, The World Health Organization (WHO) reported that at least three-quarters of CVD deaths occur in low- and middle-income countries.[7] Age-standardised CVD events and mortality were estimated to be declining in many nations in recent years, with a more marked decline in developed nations.[8] Within developed countries, a dramatic decline was reported among regions with very high sociodemographic indices, but only a gradual decrease or no change in most other areas.[8] Despite this, CVD remains the leading cause of death and health care expenditure in developed nations, including Australia.[9]

In Australia, CVD remains the single leading cause of death; largest health problem; and the major health-related economic cost to the nation.[10] As per the Australian Bureau of Statistics (ABS) 2019 released Causes of Death in 2018, cardiovascular disease accounted for one in four of all deaths in Australia, and on average,

76 deaths each day.[11, 12] In addition, the prevalence of CVD is projected to steeply increase in the coming decades.[10]

Cardiometabolic risk factors (CMRFs) are a group of interrelated individual-level metabolic risk factors which eventually may lead to the development of CVD.[13] Nine in 10 adult Australians have at least one risk factor for CVD; and one in four have the occurrence of an aggregate of three or more risk factors present simultaneously.[14] Even though lifestyle changes have the potential to control or prevent most of these risk factors, many individuals do not achieve or maintain target risk-reduction through lifestyle changes alone.[15-17] People belonging to lower socioeconomic groups and those residing in remote areas have been found to have higher rates of hospitalisation and death due to CVD in Australia.[10, 14] This emphasises the necessity to look beyond individuals at risk to disadvantaged groups and geographic locations of people at risk, in order to effectively mitigate the risk factors of the development of CVD at an early stage.

Hypercholesterolemia, diabetes mellitus (DM) and chronic kidney disease (CKD) are three major chronic cardiometabolic risk conditions. The diagnosis of one or more of these risk conditions places the person at an increased lifetime risk for the development of CVD.[13, 18-20] Hypercholesterolemia directly increases the risk for CVD through artherosclerosis.[15] Diabetes Mellitus, characterised by elevated blood glucose, over time increases the risk of atherosclerosis and microcirculation inadequacies.[18, 19] Impaired fasting blood sugar level (FBSL) and glycated haemoglobin (HbA1C) levels are two commonly used clinical laboratory parameters to diagnose and monitor DM.[21, 22] Similarly, CKD patients also have a high risk for the development of CVD, in part through inappropriate activation of the renin–angiotensin system.[20] In addition, obesity is also found to increase the risk for CVD, independently or in combination with any of the above chronic conditions.[23, 24]

The management of these chronic CMRFs demands routine clinical testing and monitoring; and over the time this creates a series of data at an individual-level. For different diagnostic groups of such chronic conditions, aggregations of similar series of individual-level data can create large datasets. Considering the nesting of such individual-level data within areas, these databases can potentially be used beyond the individual-level, especially for area-level health care service commissioning and planning activities. Figure 1.1 conceptually represents the discrete and co-existing measures of CMRFs nested within individuals, individuals nested within small geographic areas with defined characteristics, and small-areas constituting larger geographical regions. The availability of such nested data would call for appropriate analytical approaches to accurately reflect the structuring of the data.

Multilevel analyses are generally recognised as appropriate for data which are hierarchical in structure.[25-27] In these analyses, the unit of analyses would usually be individuals (at the lower level) who are nested within areas (at a higher level).[28] Multilevel models address data hierarchies by allowing for residual components at each level of the data.[29] For example, a two-level model which enables grouping of individual-level outcomes within area-level units would include residuals at the individual-level and area-level. Thus, the residual variance can be partitioned into a between-area component (the variance of the area-level residuals) and a within-area component (the variance of the individual-level residuals).[30] The area-level residuals, often called "area effects", represent unobserved area characteristics associated with individual-level outcomes. It is these unobserved variables which lead to a correlation between outcomes for individuals from the same

area.[25, 28, 29] Identifying geographic variation in CMRFs would be a useful first step in the analysis of such nested data.

There is consistent evidence in the literature on geographic inequalities of CMRFs which have been reported from multiple nations.[31-38] Even though we have a better understanding of the individual-level factors contributing to CMRFs due to decades of research, there is relatively less research and a general poorer understanding concerning area-level factors. An understanding of these factors can inform area-level health care service commissioning and related policy development. Therefore, an initial review of the literature was performed to identify existing reports on geographic variation of CMRFs and the area-level factors attributing to this variation.

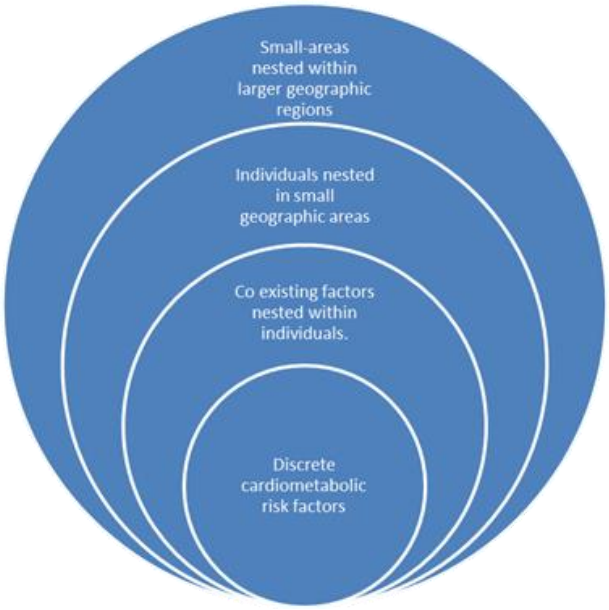

Figure 1.1: Conceptual image of the hierarchical nesting of cardiometabolic risk factors within the areas

1.3 Literature overview

An initial search for related studies was undertaken in the early phase of the study to construct a broad understanding of the literature background. A more detailed systematic review of the literature is presented in *Chapter 2* of this thesis.

Preliminary searches revealed that epidemiological analyses of the geographic location and variability in diseases had become increasingly common over the last 20 years, in parallel with advances in computing power and developments in GIS.[4] Disease atlases were initially constructed using mapping techniques and epidemiological data.[39] More recent advances in GIS and multilevel analyses of data have provided various tools for the geographical illustration and spatial analyses of area-level data.[40] Visualisation of the geographic distribution of various disease and risk factors became possible through geocoding and mapping of the data.[41] Spatial analytical methods applied to non-communicable diseases were first used in reporting diabetes incidence in Finland, based on methods reported by Whittemore et al (1987).[42] Later, geospatial analyses of health related data used geographic location information on occurrences to detect and quantify spatial patterns and clustering of area-level health and risk factors. This further extended the scope of the

statistical analysis of data at multiple levels to investigate the association between potential area-level risk factors and the prevalence of disease risk indicators.[43]

1.3.1 Geographic variation in the prevalence of cardiometabolic risk factors

Multiple studies internationally have reported geographic variation in the prevalence of CMRFs at various geographic scales. In the preliminary literature search, the reported geographic scales ranged from multinational country aggregations to smaller areas such as cities and towns [37, 44-51], although even finer geographic scales have also been used [31-38](see Chapter 2, page 48) Area level features and neighbourhood characteristics, such as area-level disadvantage and facilities associated with geographic variation in CMRFs, have been analysed more frequently in recent studies than earlier work.[37, 44-51]

Over time, studies reporting the geographic variation of various CMRFs have reduced the size of the geographic unit of analyses compared with earlier studies. Geographic variations in the prevalence of various CMRFs have been reported across cohorts of countries in Northern Europe, America and Southern Europe.[50] Also between-country level variations in CMRFs have been reported across the USA, Finland, The Netherlands, Italy, Croatia, Serbia, Greece and Japan.[49] National level studies reporting geographic variation in the prevalence of CMRFs across regions, towns and cities have been reported from Spain, Italy, Britain and USA.[34, 37, 45, 46, 48] Within-district level variations in the prevalence of cardiovascular risk factors and socioeconomic variables have been examined in Ohio, USA.[47] Within-city level variations in CMRF distribution were reported from Sweden and USA, with significant correlation between area-based scores of cardiovascular risk and disadvantaged circumstances.[44]

In addition to geographic variation in prevalence, significant geographic clustering of individual CMRFs has been reported in multiple studies.[39, 52-55] The use of GIS software such as SatScan and ArcGIS have been among the most common analytical methods for cluster detection and density plot mapping.[39, 52-55] These geospatial analytical approaches have demonstrated the potential to illustrate and visually represent the interplay of area level risk factors.[39-41, 55-57] Study reports of the geographic clustering of selected CMRFs have also pointed to the possible roles of area-level characteristics in the prevalence of risk factors.[39, 54]

In recent years there has been an explosion of research reports based on multilevel analysis of CMRF data and their association with area-level socioeconomic disadvantage.[58-60] These reports are mainly from industrialised nations across the world. The current evidence regarding the association between various CMRFs and area-level socioeconomic disadvantage is the focus of the following sections of this chapter.

1.3.2 Area-level socioeconomic disadvantage and cardiometabolic risk factors

An initial literature search regarding area-level socioeconomic disadvantage and CMRFs retrieved a total of 32 studies. The highest numbers of studies retrieved were from the USA (n=13); followed by UK, Australia, France, Germany, Sweden, Canada and China. Most studies used a cross sectional design (n= 25), with longitudinal studies less frequently reported (n=4). The majority of the studies had used samples from existing population based studies/programs, electronic medical record databases and national/regional level surveys or surveillance. Census block level aggregates were the most commonly used measure of area-level socioeconomic disadvantage. Other measures included: electoral ward level deprivation score; built environment status; mixed land use level; and area level unemployment and overcrowding.[58-67]

Both the longitudinal and cross sectional studies reported significant association between various CMRFs and area-level socioeconomic disadvantage, independent of individual-level characteristics such as socioeconomic status, education and duration of exposure to area.[58, 59, 62-68] Less disadvantage was consistently reported to have a protective effect on behavioural cardiac risk factors such as smoking, physical inactivity and obesity.[59, 63, 69] Men from highly urbanised environments have been reported to have higher incidence of heart disease with increasing area-level socioeconomic disadvantage, after adjusting for individual characteristics.[61] Type 2 diabetes and high body mass index (BMI) had been reported more prevalent in disadvantaged areas.[33, 68, 70-76] However, LDL management had been reported as not being associated with area-level disadvantage.[77] Type of neighbourhood food outlets,[70, 78, 79] poor physical activity resources,[78] individual perception of area level features,[80] residential density and service availability[58] had all been reported as explanatory variables associated with cardiometabolic risk prevalence among people living in disadvantaged areas.

As the CMRFs are generally chronic in nature, area-level accesses to primary health care services were assumed to have an effect on CMRFs resulting from their identification and ongoing management. Primary health care relates to people who are not hospitalized for treatment and generally the first contact a person has with the health system.[81] Thus, access to primary care may have a direct effect on the geographic distribution of CMRFs. Therefore CMRFs studies reporting on the access to primary care were also included in the initial review.

1.3.3 Geographic access to health care services and cardiometabolic risk

Studies examining the relationship between health care service outcomes and travel time using multilevel logistic regression models found that general practitioner (GP) consultations were less likely to happen when the travel time was longer in rural areas.[82] When access to adequate treatment and the geographical pattern of end stage renal disease (ESRD) across 46 counties in southern California in USA were investigated, ESRD incidence was found to be consistently higher in rural compared to urban counties.[51] In addition, the rates were found to be inversely associated with physician density, suggesting that access to adequate treatment facilities had a role in preventing ESRD.[51]

A multinational study by the World Health Organisation (WHO), the DiaMond complications study (DiaComp), examined the role of availability and access to health care facilities in relation to the complications of type 1 DM across 12 countries.[83] Using clinical laboratory data, markers of diabetes complications in 14 clinical centers across the study nations were assessed and then linked to health care access, cost, and local social and economic landscapes. Results of the DiaComp study suggested that health system performance and the social distribution of wealth played roles in explaining the geographic variations in resulting complications from DM.[83]

1.4 The extent of current evidence

This section describes the extent of evidence based on the preliminary literature review. The section outlines what we already know; and sets direction on what we need to know with further research which will be addressed in this thesis.

1.4.1 What we know:

Area-level disadvantage and health

In 2008, the World Health Organization's pioneering Commission on Social Determinants of Health reported that a young boy living in a disadvantaged area of Glasgow, Scotland had an average life expectancy of 54 years when compared with a boy living in an affluent area, only 12 km away, who could expect to live to 82 years: a loss of 28 years.[84] Michael Marmot, the then chair of the WHO's Commission, later reflected this in his book 'The Health Gap' that this "….was a tale of two cities…both in Glasgow".[85] This is a stark example of the contextual effect of area-level disadvantage on people.

The role of socioeconomic contexts on health has been recognised for centuries.[86] Whether it is the mill towns of Victorian England [87] or the slums of contemporary India [88], the poorer in every society succumbed to morbidity and early death [86]. Moreover, this was observed across history, not as a threshold effect of poverty, but a gradient effect where the middle class enjoyed better health than the poor and the affluent enjoyed better health than the middle class.[86] A social gradient in health is an exceedingly widespread phenomenon.[89] In general, individuals from lower socioeconomic circumstances are at greater risk of poorer health outcomes, disability and death.[89] Life and death inequalities underpinned by the differences in income, wealth, and power[90], directly points to the levels of justice and fairness in a society.[91]

In the research literature, a direct relationship between the socioeconomic position of populations and their health has been a focus for centuries.[92] The earliest available literature on this can be found in the 1567 writings of Paracelsus—the father of toxicology and pioneer of the "medical revolution" of the Renaissance—who noted unusually higher rates of diseases among miners of the medieval period Europe.[93] This relationship has always been present in society in sometimes prominent and sometimes subtle ways.[86] For instance, it is notable that in the Titanic event (1912), women who travelled in the 3rd class had a 20 times higher likelihood of dying from drowning mainly due to the socioeconomic stratification in the distribution of the resource, here the lifeboat.[94] In recent decades, there has been a notable increase in studies reporting on this relationship from various nations across the world and is one of the most consistent findings in epidemiology.[95] Area socioeconomic measures reported in the literature include both the component or composite measures such as a median or per capita income, median education, unemployment rate, the percentage in poverty, white-collar job rates and area socioeconomic scores/indices, or deprivation scores/indices.[86] Even though the period and the population are different across historical and recent research literature, the socioeconomic pattern of morbidity and early mortality remains the same: that the socioeconomically disadvantaged population get sick and die earlier.[86] However, the literature also shows that this relationship is not the same in all contexts with political, cultural and institutional factors having influential and modifying effects.[86]

Theoretical and empirical constituents

Following is an overview of the theoretical and empirical constituents of the socioeconomic position of individuals and populations.

There are three major theories of social stratification: Marxian theory, Weberian theory, and Functionalism. The Marxian theory views social stratification as imposed by exploitative resource capitalism.[96] In contrast, Weberian theory views the stratification of society as a multidimensional reflection of the interplay between

wealth, prestige and power and its role in creating an individual's "life chances" in terms of factors such as education, occupation and income which are vital in the unequal distribution of resource within a population.[97] Finally, Functionalist concepts originated in the United States of America (USA), which views social stratification as a natural and essential part of complex modern societies[98]. To a greater or lesser extent, various combinations of these three schools of thought can be found in the measurements of socioeconomic position in the health literature.[97, 98] However, it is the individualist Weberian school of thought that seems to lead epidemiological researchers to measure indices of "life chances" such as education, occupation, income and housing of individuals and their averages to the areas.[86]

Social determinants of health are the factors present in society that influence individual health.[99] Even though individual-level factors such as education provide a logical structural link between the occupation and income of an individual, there also exists extra-individual dynamics beyond these individual-level factors.[100] For example, beyond the education or income or employment status of an individual, their residential area access to primary care services can have a direct influence on the identification and ongoing management of CMRFs in an individual.[51, 82, 83] Thus, social and economic resources have the potential to configure the health of individuals beyond individual-level factors.[101]Individual-level health is often a result of the social determinants of health.[102] It is not an accident that an individual consumes a diet high in saturated fat and salt. It also reflects on the food choices available, affordability, cultural practices, and other social and marketing influences.[102] Thus the biochemical causes of diseases in an individual are largely influenced by the social circumstances of an individual, which been termed “a cause of the cause”.[102] Social determinants of health are considered to be mostly responsible for health inequities among populations.[89, 99]

Measurement of area-level disadvantage

The measurement of the socioeconomic position of an area can be either absolute or relative.[91] Both methods have their own merits and disadvantages.[103] For example, based on an absolute measure, the average financial income level of an area may have increased over time, but in relative terms, this increased income may not be sufficient to afford the cost of material resources to benefit health outcomes to the same extent as neighbouring areas.[91] However, absolute measures may help in studying changes over time, whereas the relative measures are mostly based on concurrent data and statistics, and are often criticised for overestimating effects and their interpretive complexity.[104, 105]. It is for this reason, methodological experts and organisations including the World Health Organization's Commission on Social Determinants of Health, recommend a balanced reporting of both absolute and relative measures, when feasible, for an overall picture of socioeconomic inequalities.[101] However, in the absence of access to absolute data, research has often relied on relative measures.[104] Review reports indicate that population health inequality studies are most commonly reported using relative measures.[104]

Socio-Economic Indexes for Areas (SEIFA) are relative measures of area-level socioeconomic indices developed by the Australian Bureau of Statistics (ABS).[106] SEIFA rank areas in Australia based on how advantaged or disadvantaged they are relative to other areas.[106] SEIFA consists of four area indices which are constructed using the principal components scores of the data from the five-yearly Census of Population and Housing.[106] The four SEIFA indices consist of: i) The Index of Relative Socio-economic Disadvantage

(IRSD); ii) The Index of Relative Socio-economic Advantage and Disadvantage (IRSAD); iii) The Index of Education and Occupation (IEO); iv) The Index of Economic Resources (IER).[106]

The thesis uses only one measure of area-level socioeconomic disadvantage (ASED), which is the Index of Relative Socioeconomic Disadvantage (IRSD). The IRSD is chosen for the analyses primarily as it is the only unitary index of the area-level disadvantage from SEIFA (in distinction to IRSAD), and the other indices are more oriented to education, occupation and economic resources. Australian Government agencies use IRSD indices for planning of services and funding allocation.[108] As socioeconomically disadvantaged people are a priority population for health monitoring, researchers routinely use IRSD to analyse and report the health outcomes.[99] The IRSD was chosen as the independent variable for analyses in this thesis because it is primarily designed to compare the relative socio-economic characteristics of areas at a given point in time[107] and various Australian Government agencies use IRSD by itself or with other more targeted information to assist in determining the allocation of resources and services[108]. As one of the main intentions of this research programme is informing regional health care service commissioning, the thesis uses IRSD as the only measure of ASED.

The IRSD summarises the socioeconomic disadvantage of an area into an index score based on the aggregate characteristics of its usual resident population using variables such as low income, low educational attainment, high unemployment, and jobs in relatively unskilled occupations.[106] The areas are then ranked based on the scores and grouped based on the ranks, often in quintiles.[99] It should be noted that IRSD reflects the overall or average level of disadvantage of an area but does not identify differences between individuals living in the same area.[99] Being an average, the score is likely to reduce obvious differences between individuals within an area, and between areas.[99] For this reason, IRSD is recommended to be used with the smallest available area unit and thus the smallest population size.[99] As socioeconomically disadvantaged people are a priority population for health monitoring, researchers routinely use IRSD to analyse and report the health outcomes.[99] Thus, IRSD was chosen as the independent variable for analyses in this thesis because it is primarily designed to compare the relative socio-economic characteristics of areas at a given point in time[107] and various Australian Government agencies use IRSD by itself, or in conjunction with other (more targeted) information to assist in determining the allocation of resources and services[108].

Area-level disadvantage had been reported as being inversely associated with different CMRFs in multiple studies. Most of these studies were reported from industrialised nations mainly across the USA and European regions of the world, but only a few from the Oceania region. The reported associations were mostly identified in cross-sectional studies and independent of individual-level factors such as, age, sex, education and income. Area-level access to primary care is a correlate of geographic variation of certain CMRFs. Primary care consultations have been reported as less likely to occur when the access to care was poor. Therefore access is assumed to have a direct effect on the detection and ongoing management of CMRFS, as they are generally chronic in nature and require continuous but non-hospitalised care.

1.4.2 What we need to know:

Small-area level analyses of CMRFs could be a useful step forward in understanding the geographic variation of CMRFs in regional Australia. In selecting a suitable geographic unit for the analyses, the smallest possible area-level unit is preferred as it lowers the risk of ecological fallacy.[26] Small-areas are likely to have the smallest possible population size and thus a less heterogeneous population than larger geographic areas.[26]

This potentially minimises the within-area variations and maximises the between area variations.[26] Previous research has demonstrated that it is feasible to produce small-area geospatial maps from health care service record data, but that this required significant technical expertise.[109] Previous research also demonstrates that when smaller area level units were used, the measured health outcome inequalities were larger in comparison with the use of larger area level units.[110] Thus, the geocoding and mapping of CMRF data in this thesis may facilitate not only area-level visualisation of the data, but also facilitate generation of hypothesis for further area-level analyses. In Australia, regional planning for the prevention and management of CMRFs lacks information about its epidemiology within small-areas.[68] Centralised approaches of disease prevention and management may not suit regional requirements.[111] The geographic distribution patterns at larger geographic scales may not adequately represent significant local geographic variations. Analysis at small-area levels is important in order to understand local patterns and requirements.[68]

Area-level disadvantage had been reported to be associated with individual CMRFs in previous studies.[58-67] An analysis of this association across multiple risk factors may actually be more informative for the local management and prevention of CMRFs. Multiple risk factors occurring simultaneously in disadvantaged areas may focus attention and improve area-appropriate preventive approaches and region specific health care service commissioning.

Estimating and analysing primary care access at a small-area level may further extend our understanding of the distribution of CMRFs within a geographic region. Previous studies from Australia have reported that using the remoteness of the areas alone for health care service planning has significant limitations, but including measures of socioeconomic disadvantage and workforce supply may better target health inequities and improve resource allocation.[112] Therefore extending the analyses to include access to primary care may help provide a better picture of the need for health care services in the region, especially in disadvantaged areas.

In summary, the geospatial analysis of a wide range of CMRFs. in conjunction with appropriate multilevel analyses has the potential to provide valuable evidence for area-level health care service commissioning in Australia. The findings may reveal geographic variation in CMRF distribution and the area-level factors associated with these inequalities. Geospatial mapping of CMRFs may facilitate the visualisation of the geographic variation and generate hypotheses for multilevel analyses.[113] The evidence created from analysis of multiple risk factors (if consistent across the factors) could inform future planning of targeted health care service commissioning in regional areas.

1.4.3 Key challenges

The data and expertise required for area-level analyses is a key challenge in regional studies. The use of hospital linkage data would not cover community level distribution of CMRFs, as identification and ongoing management of CMRFs would not generally require hospitalisation until severe. However, de-identified data from community service providers may be a possible source of population derived data. The feasibility of this method has been reported earlier.[68] With area-level analyses, the choice of fitting algorithm and its implementation within accessible analytical software would be a key decision. The major analytical platforms available for area-level analyses include, but are not limited to R, SAS, MLwiN, Stata, SPSS, S-Plus, GLLAMM, HLM, MIXREG, SYSTAT, and WinBUGS.[114]

1.5 Research plan

As de-identified data on multiple CMRFs were available from the Southern IML Research (SIMLR) database for the Illawarra-Shoalhaven region of NSW Australia, the region was chosen to analyse their geographic variation and association with selected contextual factors including area-level socioeconomic disadvantage. The Illawarra-Shoalhaven has a diverse socio-economic profile, making it a useful region for area-level population health studies. [68] Figure 1.2 shows the study area with SA1 units and the major landmarks of the region.

The Illawarra-Shoalhaven region covers a land area of 5615 square kilometres and had an estimated residential population of 369,469 persons at the 2011 Australian Census of Population and Housing conducted by the Australian Bureau of Statistics (ABS), of which 285, 385 (77.24%) were aged 18 years and over.[115] Statistical Area Level 1 (SA1) was the smallest geographical unit of the 2011 Census at which data were released.[115] SA1s typically have a population size of 200 to 800 persons (average 400)[115], and the Illawarra-Shoalhaven region comprises a total of 980 conterminous SA1s.

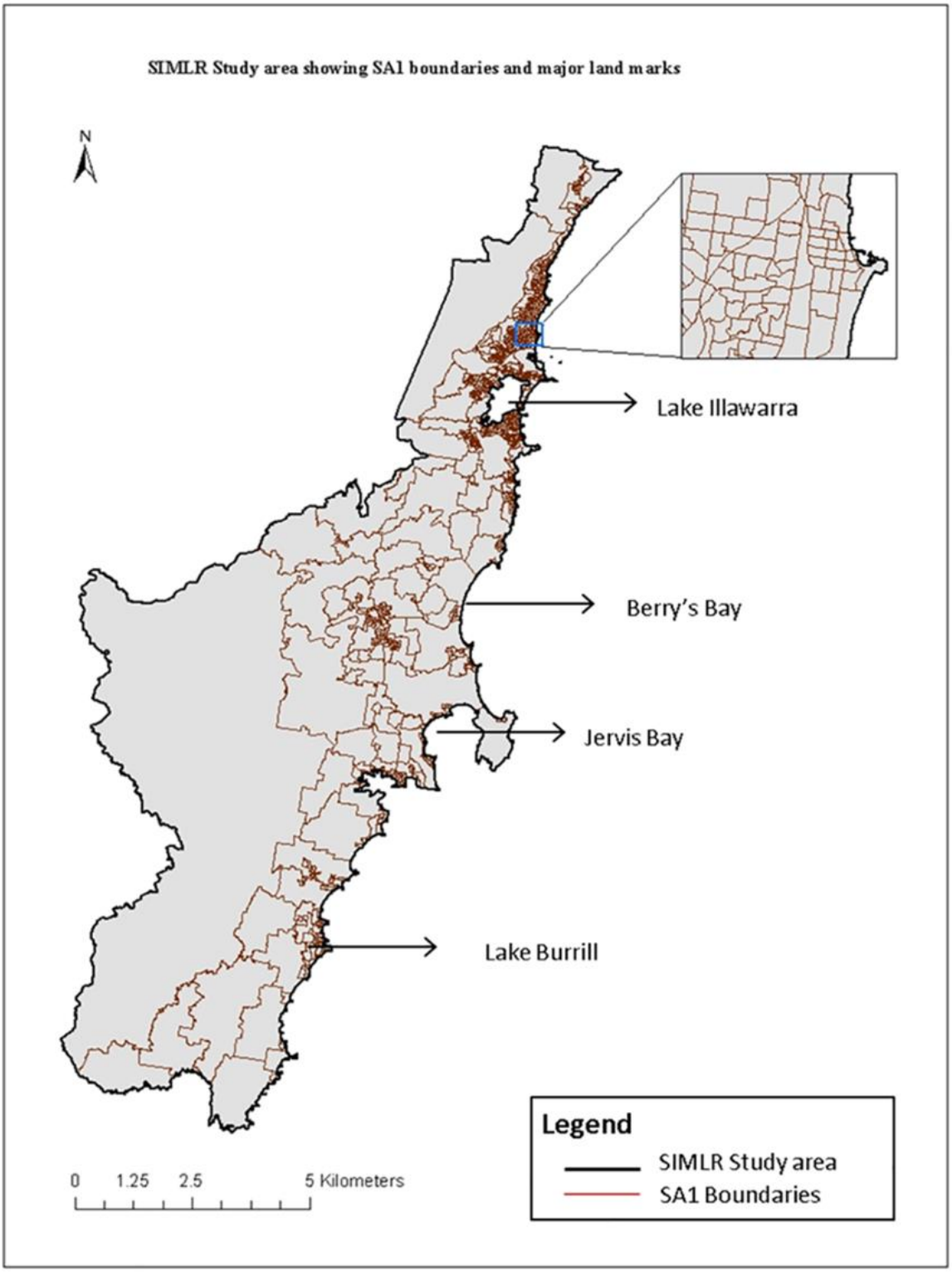

Figure 1.2: Map of the study region with the SA1 unit boundaries and major land marks.

The population profile of the Illawarra-Shoalhaven region is culturally and linguistically diverse, with a significant proportion of non-English speaking people (10.5%) residing in this region who have migrated from overseas.[116] In addition, at the 2011 Census the region is identified to have more than the NSW state and Australian national averages of: 1) Aboriginal and/or Torres Strait Islander peoples (3% versus 2.5% NSW and 2.5% Australia); 2) aged (>=65 years) population (17.6% versus 14.5% NSW and 13.8% Australia) ; 3) single-parent households (5.8% versus 5.3% NSW and 5.2% Australia); and 4) unemployment (7.1% versus 5.1% NSW and 5.1% Australia) and lower labour force participation rate (57.9% versus 64.6% NSW and 66.2% Australia).[116] ABS 2011 census data indicate that more than 31 % people in the Illawarra-Shoalhaven reside in Inner Regional areas, and 9.1% households within this region did not have a motor vehicle.[116] The Illawarra-Shoalhaven geography and a limited public transport system, especially in isolated communities, make it difficult for many people to access health services quickly.[116] These characteristics of the study region directly indicate the vulnerability of its population to poorer health outcomes. Thus, the Illawarra-Shoalhaven geographic area was chosen to study due to data availability and a population likely to benefit from the outcomes of the research. Based on the research needs identified in the preliminary literature review, a research proposal was prepared, which included the following questions, objectives, hypotheses and expected outcomes.

1.5.1 Research questions

1. What is the existing level of evidence on the geographic and socioeconomic variation in the distribution of CMRFs internationally?
2. What is the small-area level geographic distribution pattern of cardiometabolic risk factors, within the Illawarra-Shoalhaven region of NSW Australia?
3. What proportion of any geographic variability in the distribution of cardiometabolic risk factors is due to small-area level socioeconomic status, within the Illawarra-Shoalhaven region of NSW Australia?
4. What proportion of any geographic variability in cardiometabolic risk factor distribution is due to differences in small-area level primary care access, within the Illawarra-Shoalhaven region of NSW Australia?

Based on the above research questions, following objectives and hypotheses were derived.

1.5.2 Objectives

1. Systematically review the existing literature on the geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution.
2. Quantify small-area geographic variation in the distribution of cardiometabolic risk factors, within the Illawarra-Shoalhaven region of NSW Australia.
3. Quantify the proportion of small-area geographic variation in cardiometabolic risk factors explained by the area-level socioeconomic disadvantage, within the Illawarra-Shoalhaven region of NSW Australia.
4. Quantify the proportion of small-area geographic variation in cardiometabolic risk factors explained by the differences in geographic access to primary care, within the Illawarra-Shoalhaven region of NSW Australia.

1.5.3 Hypotheses

Within the Illawarra-Shoalhaven region of NSW Australia:

H_0 – Cardiometabolic risk factors are distributed in random.

H_1 – Cardiometabolic risk factor distributions are not related to small-area socioeconomic disadvantage.

H_2 – Cardiometabolic risk factor distributions are not related to small-area access to primary care.

1.6 Thesis overview

The thesis design is based on the research questions, objectives and hypotheses outlined in the research plan. Figure 1.3 provides an illustration of the thesis design and how the individual studies link together.

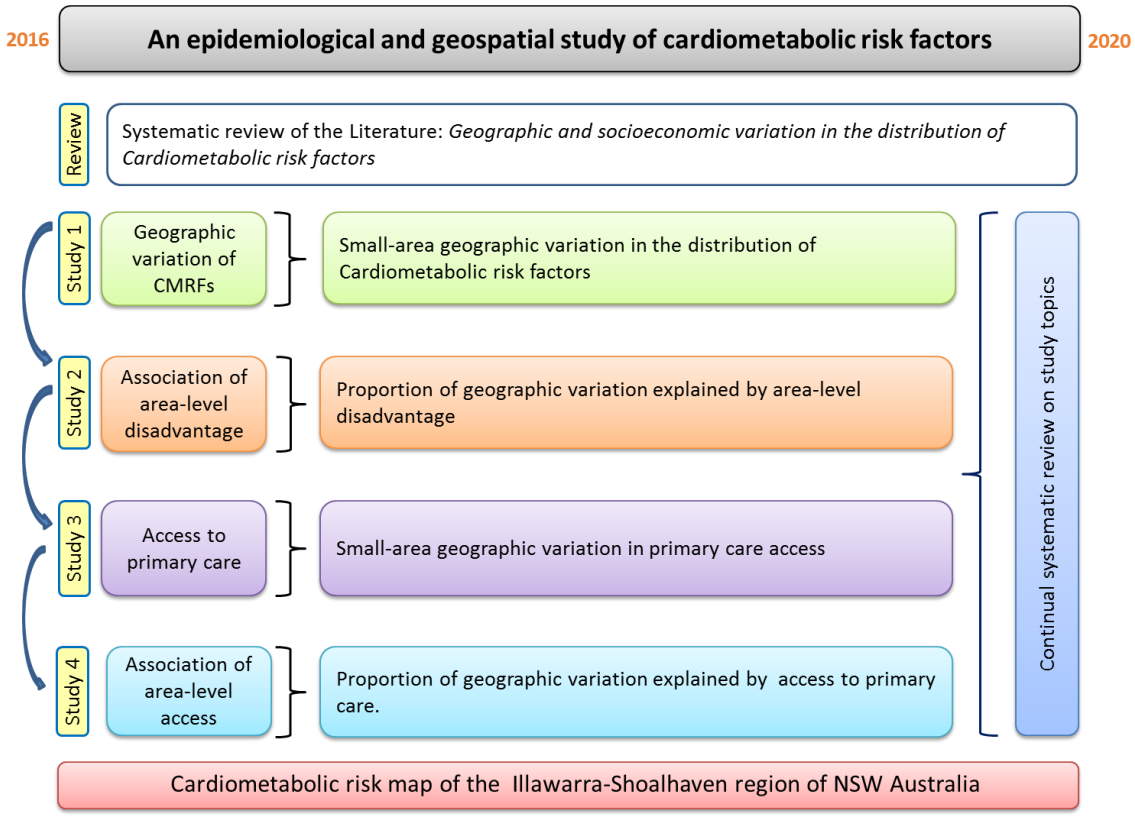

Figure 1.3: Thematic representation of the thesis design and the supporting studies

1.7 Thesis structure

The thesis is structured as a compilation of publications. A *General Introduction* is given in chapter 1 and a *General Discussion and Conclusion* is presented in chapter 8. Also, chapter 3 *Materials and methods* is written to supplement the thesis, describing the methodologies adopted in individual studies. All other chapters in this thesis are based on the manuscripts written for publication of individual articles. A summary and bibliography are provided at the end of each chapter.

The chapters outlined in this thesis are:

Chapter 1: General introduction

Chapter 2: Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature.

Chapter 3: Materials and methods

Chapter 4: Geographic variance in cardiometabolic risk distribution: a cross sectional study of 256, 525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia.

Chapter.5: Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia.

Chapter 6: Does access to primary care reduce the geographic variation of cardiometabolic risk factor distribution? A multilevel analysis of the adult residents in the Illawarra-Shoalhaven region of Australia.

Chapter 7: General discussion and conclusion

1.8 Thesis significance

The thesis has the potential to inform regional health service commissioning. Information on the geographic distribution of cardiometabolic risk factors is essential to that aim. It is important to note that the health care management protocols of the individual CMRFs are different from each other.[117-120] Resources required for the prevention control and management of diabetes is different from that of hypercholesteremia, or that of low renal function.[118, 119] When it comes to areas, some may require more resources for diabetes management, such as diabetes clinics, whereas others might require more dialysis units or renal care nurses. Some areas might require more exercise or walking facilities, whereas others might require more primary care locations. Therefore it is important to analyse data on the distribution of each of these CMRFs. This thesis presents a potential approach using a geocoded regional data source.

The thesis also aims to explore the association of higher risk CMRFs with area-level disadvantage. As disadvantaged areas are vulnerable to poor health outcomes[99], it is important to prioritise and plan area-appropriate approaches to improve the health outcomes of these areas. This study plans to focus on area-level disadvantage to describe its association with various CMRFs. This would help to demonstrate the pattern of association of a range of CMRFs across the levels of disadvantage.

Further, the thesis also attempts to explain the geographic variation in CMRFs in relation to the level of primary care access of the small areas within the study region. It should be noted that a proportion of the population in the study region do not have private vehicles and the public transport system does not cover all the regional and remote areas within this study region.[116] In such a context, geographic access to primary care services is of vital importance, especially for the early identification and ongoing management of various CMRFs.

Overall, the measurement of inequalities in health is essential to define, describe, and understand the nature of the public health problem.[121] It is an important step in the development of strategies and policies to tackle health inequalities, and in the monitoring and evaluation of the effectiveness of existing approaches. The study has the potential to identify the geographic associations of areas and CMRFs, and thus the appropriateness of targeted or universal approaches for the prevention and control of CMRFs in the study region. Often it would require collaborative policymaking approaches between different sectors such as health, education, urban/rural planning, and employment sectors to act together to reach the health and welfare goals of a population.[122] Overall the study can generate evidence for resource allocation, planning and informing local solutions to locally raised problems for the long term prevention and management of cardiometabolic risk factors in the region and potentially nationally.

Therefore, this program of study has the potential to make significant impact on regional planning and implementation of preventive health care services. Description of small-area geographic variation in CMRFs and their environmental contexts in relation to area-level disadvantage and access to primary care are the chief outcomes of this research. Quantification of small-area geographic variation in CMRFs associated with area-level disadvantage at small-area level could assist our understanding of the socioeconomic context of CMRFs in the study region. Analysis of the geographic access to primary care may further help to understanding its link with the CMRFs, especially in disadvantaged areas. Cardiometabolic risk mapping of the region visually

translates the geospatial analyses of the data along with epidemiological findings. This may make the findings of this research more accessible to the end users than traditional graph and table methods to report research findings.[123]

1.9 Summary of the chapter

Chapter 1 provides a general introduction to the thesis. This chapter summarises the importance of the place of living on human health and provides a short literature background on the importance of location on the development of CMRFs. Based on the existing research, a research plan was articulated in this chapter.

Geographic variation in cardiometabolic risk factor prevalence at a small-area level is the major concept dealt with in this thesis. The thesis attempts to describe 1) the geographic variation in CMRF distribution in the Illawarra-Shoalhaven region of the NSW, in Australia. Also, the contextual associations of 2) area-level disadvantage and 3) area-level access to primary health care services – with the geographic variation observed.

The thesis is presented by compilation of publications. The importance of the thesis is explained in terms of its possible impacts on regional planning services, policy initiatives and preventive health care services at a small-area level.

References

1. Yapijakis, C., M. Bartsakoulia, and G. Patrinos, *Hippocrates, the father of clinical medicine and Asclepiades, the father of molecular medicine.* Archives of Hellenic Medicine/Arheia Ellenikes Iatrikes, 2013. **30**(1).
2. Snow, J., *On the mode of communication of cholera.* The Challenge of Epidemiology: Issues and Selected Readings, 1988. **505**: p. 42.
3. Lawson, A.B., *Statistical methods in spatial epidemiology*. 2013: John Wiley & Sons.
4. Auchincloss, A.H., et al., *A review of spatial methods in epidemiology, 2000–2010.* Annual review of public health, 2012. **33**: p. 107-122.
5. Mathers, C., *The global burden of disease: 2004 update*. 2008: World Health Organization.
6. Bloom, D.E., et al. *"The global economic burden of noncommunicable diseases". PGDA Working Papers 8712, Program on the Global Demography of Aging(PGDA).* 2012; Available from: https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1288/2013/10/PGDA_WP_87.pdf.
7. World Health Organization. *Cardiovascular diseases (CVDs): Key facts* [cited 2016 Nov 5]; Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
8. Roth, G.A., et al., *Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015.* Journal of the American College of Cardiology, 2017. **70**(1): p. 1-25.
9. World Health Organization. *The top 10 causes of death*. 24 May 2018 [cited 2020 Sept 5]; Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
10. Australian Institute of Health and Welfare, *Australia's health 2014*. Canberra: AIHW; 2014.
11. Australian Bureau of Statistics, *Causes of Death 2018*, in *cat. no. 3303.0*. 2019.
12. National Heart Foundation of Australia. *Key Statistics: Cardiovascular Disease*. 2020 [cited 2020 Sept]; Available from: https://www.heartfoundation.org.au/Activities-finding-or-opinion/key-stats-cardiovascular-disease.
13. Snyder, S. and N. Gangeri, *Obesity, Cardiometabolic risk, and chronic kidney disease*, in *Obesity*. 2016, Springer. p. 181-198.
14. Australian Institute of Health and Welfare, *Cardiovascular disease, diabetes and chronic kidney disease—Australian facts: Morbidity–Hospital care in Cardiovascular, diabetes and chronic kidney disease* in *series no. 3.* : Canberra: AIHW; 2014.
15. Brunzell, J.D., et al., *Lipoprotein management in patients with cardiometabolic risk: consensus statement from the American Diabetes Association and the American College of Cardiology Foundation.* Diabetes care, 2008. **31**(4): p. 811-822.
16. Cannon, C.P. *Cardiovascular disease and modifiable cardiometabolic risk factors*. Clinical cornerstone 2007 [cited 8 3]; 11-28].
17. Mahmood, S.S., et al., *The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective.* The lancet, 2014. **383**(9921): p. 999-1008.
18. Castelli, W.P., et al., *Lipids and risk of coronary heart disease The Framingham Study.* Annals of epidemiology, 1992. **2**(1-2): p. 23-28.
19. Garvey, W.T., et al., *Lifestyle therapy in the management of cardiometabolic risk: diabetes prevention, hypertension, and dyslipidemia*, in *Lifestyle Medicine*. 2016, Springer. p. 245-267.

20. Weiner, D.E., et al., *Kidney disease as a risk factor for recurrent cardiovascular disease and mortality.* American Journal of Kidney Diseases, 2004. **44**(2): p. 198-206.
21. Santaguida, P.L., et al., *Diagnosis, Prognosis, and Treatment of Impaired Glucose Tolerance and Impaired Fasting Glucose: Summary*, in *AHRQ Evidence Report Summaries*. 2005, Agency for Healthcare Research and Quality (US).
22. Selvin, E., et al., *Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults.* New England Journal of Medicine, 2010. **362**(9): p. 800-811.
23. Akil, L. and H.A. Ahmad, *Relationships between obesity and cardiovascular diseases in four southern states and Colorado.* Journal of health care for the poor and underserved, 2011. **22**(4 Suppl): p. 61.
24. Ritchie, S. and J. Connell, *The link between abdominal obesity, metabolic syndrome and cardiovascular disease.* Nutrition, Metabolism and Cardiovascular Diseases, 2007. **17**(4): p. 319-326.
25. Steenbergen, M.R. and B.S. Jones, *Modeling multilevel data structures.* american Journal of political Science, 2002: p. 218-237.
26. Allik, M., et al., *Creating small-area deprivation indices: a guide for stages and options.* J Epidemiol Community Health, 2020. **74**(1): p. 20-25.
27. Duncan, C., K. Jones, and G. Moon, *Context, composition and heterogeneity: using multilevel models in health research.* Social science & medicine, 1998. **46**(1): p. 97-117.
28. Gelman, A. and J. Hill, *Data analysis using regression and multilevel/hierarchical models*. 2006: Cambridge university press.
29. Hox, J.J., M. Moerbeek, and R. Van de Schoot, *Multilevel analysis: Techniques and applications*. 2010: Routledge.
30. Merlo, J., et al., *An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health.* PloS one, 2016. **11**(4): p. e0153778.
31. Alkerwi, A., et al., *Geographic variations in cardiometabolic risk factors in Luxembourg.* International journal of environmental research and public health, 2017. **14**(6): p. 648.
32. Astell-Burt, T., et al., *Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia.* Diabetes research and clinical practice, 2014. **106**(3): p. e68-e73.
33. Barker, L.E., et al., *Geographic distribution of diagnosed diabetes in the US: a diabetes belt.* American journal of preventive medicine, 2011. **40**(4): p. 434-439.
34. Lawlor, D., et al., *Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study.* Journal of Epidemiology & Community Health, 2003. **57**(2): p. 134-140.
35. Oh, W.S., et al., *Geographical variations and influential factors in prevalence of cardiometabolic diseases in South Korea.* PloS one, 2018. **13**(10): p. e0205005.
36. Paquet, C., et al., *Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia.* International journal of environmental research and public health, 2016. **13**(5): p. 519.
37. Valdés, S., et al., *Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@ bet. es study.* Revista Española de Cardiología (English Edition), 2014. **67**(6): p. 442-448.

38. Zhou, M., et al., *Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults.* Diabetes care, 2015. **38**(1): p. 72-81.
39. Hahn, R.A., G.W. Heath, and M.-H. Chang, *Cardiovascular disease risk factors and preventive practices among adults—United States, 1994: A behavioral risk factor atlas.* MORBIDITY AND MORTALITY WEEKLY REPORT: CDC Surveillance Summaries, 1998: p. 35-69.
40. Samuelsson, U. and O. Löfman, *Geographical mapping of type 1 diabetes in children and adolescents in south east Sweden.* Journal of Epidemiology & Community Health, 2004. **58**(5): p. 388-392.
41. Stevens, C.D., et al., *Geographic clustering of diabetic lower-extremity amputations in low-income regions of California.* Health Affairs, 2014. **33**(8): p. 1383-1390.
42. Ranta, J., et al., *Detection of overall space-time clustering in a non-uniformly distributed population.* Statistics in medicine, 1996. **15**(23): p. 2561-2572.
43. Waller, L.A. and C.A. Gotway, *Applied spatial statistics for public health data*. Vol. 368. 2004: John Wiley & Sons.
44. Engström, G., et al., *Geographic distribution of stroke incidence within an urban population: relations to socioeconomic circumstances and prevalence of cardiovascular risk factors.* Stroke, 2001. **32**(5): p. 1098-1103.
45. Gabriel, R., et al., *Prevalence, geographic distribution, and geographic variability of major cardiovascular risk factors in Spain. Pooled analysis of data from population-based epidemiological studies: the ERICE Study.* Revista Española de Cardiología (English Edition), 2008. **61**(10): p. 1030-1040.
46. Jarvie, J.L., et al., *Geographic variance of cardiovascular risk factors among community women: the national Sister to Sister campaign.* Journal of Women's Health, 2011. **20**(1): p. 11-19.
47. Jenum, A.K., I. Stensvold, and D.S. Thelle, *Differences in cardiovascular disease mortality and major risk factors between districts in Oslo. An ecological analysis.* International journal of epidemiology, 2001. **30**(suppl_1): p. S59.
48. Laccetti, R., et al., *Evidence on the prevalence and geographic distribution of major cardiovascular risk factors in Italy.* Public health nutrition, 2013. **16**(2): p. 305-315.
49. Menotti, A., et al., *Cardiovascular risk factors as determinants of 25-year all-cause mortality in the seven countries study.* European journal of epidemiology, 2001. **17**(4): p. 337-346.
50. Menotti, A., et al., *Relationship of some risk factors with typical and atypical manifestations of coronary heart disease.* Cardiology, 1998. **89**(1): p. 59-67.
51. Occelli, F., et al., *Mapping end-stage renal disease (ESRD): spatial variations on small area level in northern France, and association with deprivation.* PloS one, 2014. **9**(11): p. e110132.
52. Oggioni, C., et al., *Shifts in population dietary patterns and physical inactivity as determinants of global trends in the prevalence of diabetes: An ecological analysis.* Nutrition, Metabolism and Cardiovascular Diseases, 2014. **24**(10): p. 1105-1111.
53. Ranta, J., et al., *Detection of overall space-time clustering in a non-uniformly distributed population.* Statistics in medicine, 1996. **15**(23): p. 2561-2572.
54. Schlundt, D.G., M.K. Hargreaves, and L. McClellan, *Geographic clustering of obesity, diabetes, and hypertension in Nashville, Tennessee.* The Journal of ambulatory care management, 2006. **29**(2): p. 125-132.
55. Faruque, L.I., et al., *Spatial analysis to locate new clinics for diabetic kidney patients in the underserved communities in Alberta.* Nephrology Dialysis Transplantation, 2012. **27**(11): p. 4102-4109.

56. Rodriguez, R.A., J.R. Hotchkiss, and A.M. O'Hare, *Geographic information systems and chronic kidney disease: racial disparities, rural residence and forecasting.* Journal of nephrology, 2013. **26**(1): p. 3.
57. Vanasse, A., J. Courteau, and M. Courteau, *The interactive atlas on health inequalities.* Spatial and spatio-temporal epidemiology, 2012. **3**(2): p. 129-140.
58. Chaix, B., *Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda.* Annual review of public health, 2009. **30**: p. 81-105.
59. Dragano, N., et al., *Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany.* BMC Public Health, 2007. **7**(1): p. 255.
60. Schmitz, N., et al., *Association between neighborhood-level deprivation and disability in a community sample of people with diabetes.* Diabetes care, 2009. **32**(11): p. 1998-2004.
61. Silhol, R., et al., *Investigating the spatial variability in incidence of coronary heart disease in the Gazel cohort: the impact of area socioeconomic position and mediating role of risk factors.* Journal of Epidemiology & Community Health, 2011. **65**(2): p. 137-143.
62. Andersen, A., et al., *Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study.* Diabetic medicine, 2008. **25**(12): p. 1462-1468.
63. Cubbin, C., et al., *Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects.* Scandinavian journal of public health, 2006: p. 228-237.
64. Laraia, B.A., et al., *Place matters: neighborhood deprivation and cardiometabolic risk factors in the Diabetes Study of Northern California (DISTANCE).* Social science & medicine, 2012. **74**(7): p. 1082-1090.
65. Lawlor, D.A., et al., *Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study.* American journal of public health, 2005. **95**(1): p. 91-97.
66. Roux, A.V.D., et al., *Neighbourhood environments and mortality in an elderly cohort: results from the cardiovascular health study.* Journal of Epidemiology & Community Health, 2004. **58**(11): p. 917-923.
67. Strachan, D.P., et al., *Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood.* International journal of epidemiology, 2007. **36**(3): p. 522-531.
68. Bonney, A., et al., *Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users–a cross-sectional study.* PLoS One, 2015. **10**(8): p. e0137261.
69. Mobley, L.R., et al., *Environment, obesity, and cardiovascular disease risk in low-income women.* American journal of preventive medicine, 2006. **30**(4): p. 327-332. e1.
70. Astell-Burt, T. and X. Feng, *Geographic inequity in healthy food environment and type 2 diabetes: can we please turn off the tap?* Medical Journal of Australia, 2015. **203**(6): p. 246-248.
71. Chaikiat, Å., et al., *Neighborhood deprivation and inequities in coronary heart disease among patients with diabetes mellitus: a multilevel study of 334,000 patients.* Health & place, 2012. **18**(4): p. 877-882.
72. Feng, X. and A. Wilson, *Do neighbourhood socioeconomic circumstances not matter for weight status among Australian men? Multilevel evidence from a household survey of 14 691 adults.* BMJ open, 2015. **5**(9): p. e007052.

73. Grintsova, O., W. Maier, and A. Mielck, *Inequalities in health care among patients with type 2 diabetes by individual socio-economic status (SES) and regional deprivation: a systematic literature review.* International journal for equity in health, 2014. **13**(1): p. 43.
74. Maier, W., et al., *Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update'GEDA 2009 and 2010.* PloS one, 2014. **9**(2): p. e89661.
75. Mezuk, B., et al., *Depression, neighborhood deprivation and risk of type 2 diabetes.* Health & place, 2013. **23**: p. 63-69.
76. Stoddard, P.J., et al., *Neighborhood deprivation and change in BMI among adults with type 2 diabetes: the Diabetes Study of Northern California (DISTANCE).* Diabetes Care, 2013. **36**(5): p. 1200-1208.
77. Geraghty, E.M., et al., *Using Geographic Information Systems (GIS) to assess outcome disparities in patients with type 2 diabetes and hyperlipidemia.* The Journal of the American Board of Family Medicine, 2010. **23**(1): p. 88-96.
78. Christine, P.J., et al., *Longitudinal associations between neighborhood physical and social environments and incident type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA).* JAMA internal medicine, 2015. **175**(8): p. 1311-1320.
79. Millstein, R.A., et al., *Food availability, neighborhood socioeconomic status, and dietary patterns among blacks with type 2 diabetes mellitus.* The Medscape Journal of Medicine, 2009. **11**(1): p. 15.
80. Baldock, K., et al., *Associations between resident perceptions of the local residential environment and metabolic syndrome.* Journal of environmental and public health, 2012. **2012**.
81. Australian Government Department of Health. *Fact Sheet: Primary Health Care*. Available from: https://www1.health.gov.au/internet/main/publishing.nsf/Content/Fact-Sheet-Primary-Health-Care.
82. Hiscock, R., et al., *Is neighborhood access to health care provision associated with individual-level utilization and satisfaction?* Health services research, 2008. **43**(6): p. 2183-2200.
83. Walsh, M.G., et al., *The socioeconomic correlates of global complication prevalence in type 1 diabetes (T1D): a multinational comparison.* Diabetes research and clinical practice, 2005. **70**(2): p. 143-150.
84. Commission on Social Determinants of Health, *Closing the gap in a generation: health equity through action on the social determinants of health: final report of the commission on social determinants of health*. 2008, Geneva, Switzerland World Health Organization.
85. Marmot, M., *The health gap: the challenge of an unequal world.* The Lancet, 2015. **386**(10011): p. 2442-2444.
86. El-Sayed, A.M., *Social Epidemiology. Edited by Lisa F. Berkman, Ichiro Kawachi, and M. Maria Glymour*. 2015, Oxford University Press.
87. Ittmann, K., *Work, gender and family in Victorian England*. 2016: Springer.
88. Ezeh, A., et al., *The history, geography, and sociology of slums and the health problems of people who live in slums.* The lancet, 2017. **389**(10068): p. 547-558.
89. Marmot, M., *Status syndrome.* Significance, 2004. **1**(4): p. 150-154.
90. *WHO (2010). A conceptual framework for action on the social determinants of health. Geneva, Switzerland: World Health Organization. Available from: http://www.who.int/socialdeterminants/publications/9789241500852/en/.*

91. Conway, D.I., et al., *Measuring socioeconomic status and inequalities.* 150 cours Albert Thomas, 69372 Lyon Cedex 08, France© International Agency for Research on Cancer, 2019 Distributed by WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland, 2019: p. 71.
92. Antonovsky, A., *Social class, life expectancy and overall mortality.* The Milbank Memorial Fund Quarterly, 1967. **45**(2): p. 31-73.
93. Jacobi, W., *The history of the radon problem in mines and homes.* Annals of the ICRP, 1993. **23**(2): p. 39-45.
94. Hall, W., *Social class and survival on the SS Titanic.* Social science & medicine, 1986. **22**(6): p. 687-690.
95. Kaplan, G.A. and J.W. Lynch, *Whither studies on the socioeconomic foundations of population health?* American journal of public health, 1997. **87**(9): p. 1409-1411.
96. Acker, J., et al., *Social class and stratification: Classic statements and theoretical debates*. 2006: Rowman & Littlefield Publishers.
97. Brennan, C., *Max Weber on power and social stratification: an interpretation and critique*. 2020: Routledge.
98. Buckley, W., *Social stratification and the functional theory of social differentiation.* American Sociological Review, 1958. **23**(4): p. 369-375.
99. Australian Institute of Health and Welfare. *Social determinants*. 2014; Available from: https://www.aihw.gov.au/reports-data/behaviours-risk-factors/social-determinants/about.
100. Lynch, J. and G. Kaplan, *Socioeconomic position.* Social epidemiology, 2000. **1**: p. 13-35.
101. World Health Organization. *Poverty: Assessing the distribution of health risks by socioeconomic position at national and local levels* [cited 2020 Oct 5]; Available from: https://www.who.int/quantifying_ehimpacts/publications/en/ebd10.pdf.
102. Marmot, M. and R. Wilkinson, *Social determinants of health*. 2005: OUP Oxford.
103. Habitat for Humanity. *What is relative vs. absolute poverty*. [cited 2020 Sept 5]; Available from: https://www.habitatforhumanity.org.uk/blog/2018/09/relative-absolute-poverty/.
104. Houweling, T.A., et al., *Using relative and absolute measures for monitoring health inequalities: experiences from cross-national analyses on maternal and child health.* International journal for equity in health, 2007. **6**(1): p. 15.
105. Clarke, P.M., et al., *On the measurement of relative and absolute income-related health inequality.* Social science & medicine, 2002. **55**(11): p. 1923-1928.
106. Australian Bureau of Statistics. *Socio-Economic Indexes for Areas*. 27 March 2018 [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/seifa.
107. Australian Bureau of Statistics. *How to Use SEIFA*. 27 March 2018 [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2011~Main%20Features~How%20to%20Use%20SEIFA~10008.
108. Australian Bureau of Statistics. *Using SEIFA to Target Areas for Services*. 27 March 2018 [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100102011.
109. Noble, D., et al., *Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning.* BMJ open, 2012. **2**(1).
110. Krieger, N., et al., *Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter? the Public Health Disparities Geocoding Project.* American journal of epidemiology, 2002. **156**(5): p. 471-482.

111. Gray, I., I.W. Gray, and G. Lawrence, *A future for regional Australia: Escaping global misfortune*. 2001: Cambridge University Press.
112. Butler, D.C., et al., *Use of measures of socioeconomic deprivation in planning primary health care workforce and defining health care need in Australia.* Australian Journal of Rural Health, 2010. **18**(5): p. 199-204.
113. Bazemore, A., R.L. Phillips, and T. Miyoshi, *Harnessing geographic information systems (GIS) to enable community-oriented primary care.* The Journal of the American Board of Family Medicine, 2010. **23**(1): p. 22-31.
114. University of Bristol: Centre for multilevel modelling. *Multilevel modelling software reviews*. [cited 2020 Sept 5]; Available from: http://www.bristol.ac.uk/cmm/learning/mmsoftware/.
115. Australian Bureau of Statistics. *2011 Census data* [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280.
116. Ghosh A, M.K., Marshall K. . *Illawarra-Shoalhaven Medicare Local - Population Health Profile: 2013.* 2013; Available from: https://www.gph.org.au/assets/Main-Site/Uploads/Resources/Improving-population-health/ISML-Population-Health-Profile-2013-FINAL.pdf.
117. The Royal Australian College of General Practitioners & Diabetes Australia. General Practice Management of Type 2 Diabetes 2016-2018. The Royal Australian College of General Practitioners (2016). doi:10.1007/s00125-010-2011-6.
118. *National heart foundation of Australia. Lipid management profile for health professionals. Available at: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management.*
119. *National Kidney foundation(USA). Albumin creatinine Ratio (ACR). (2018). Available at: https://www.kidney.org/kidneydisease/siemens_hcp_acr.*
120. *WHO. Obesity : Preventing and managing the global epidemic. World Health Organization: Technical Report Series. WHO Technical Report Series, no. 894. (2000). doi:ISBN 92 4 120894 5.*
121. Arcaya, M.C., A.L. Arcaya, and S. Subramanian, *Inequalities in health: definitions, concepts, and theories.* Global health action, 2015. **8**(1): p. 27106.
122. *World Health Organisation(WHO) . Social determinants of Health: Key concepts. Commission on Social determinants of Health final report 2005-2008. Geneva.2019 Available from: https://www.who.int/social_determinants/thecommission/finalreport/key_concepts/en/.*
123. Brewer, C.A. and L. Pickle, *Evaluation of methods for classifying epidemiological data on choropleth maps in series.* Annals of the Association of American Geographers, 2002. **92**(4): p. 662-681.

Chapter 2

Geographic and Area-level Socioeconomic Variation in Cardiometabolic Risk Factor Distribution: A Systematic Review of the Literature

Chapter 2: Geographic and Area-level Socioeconomic Variation in Cardiometabolic Risk Factor Distribution: A Systematic Review of the Literature

2.1 Publication profile

This chapter presents the substantive content of research published in: ***International Journal of Health Geographics***, on 08th January 2019.

Journal article

Toms R, Bonney A, Mayne D J, Feng X, Walsan R. *Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature*. **International Journal of Health Geographics**, 2019. 18(1), 1.

Available from: https://link.springer.com/article/10.1186/s12942-018-0165-5

DOI: https://doi.org/10.1186/s12942-018-0165-5

Published version of the article

The published version of the article is appended within the **'Supplementary Materials'** section of the thesis as **Supplementary Material 1**. Additional files of this chapter are appended in **Appendix I, II and III.**

2.2 Abstract

Introduction: A growing number of publications report variation in the distribution of cardiometabolic risk factors (CMRFs) at different geographic scales. A review of these variations may help inform policy and health service organisation.

Aim: To review studies reporting variation in the geographic distribution of CMRFs and its association with various measures of area-level socioeconomic disadvantage (ASED) among adult (>=18 years) population across the world.

Method: A systematic search for published articles was conducted in four databases (MEDLINE (Ovid), PubMed, Scopus and Web of science) considering the interdisciplinary nature of the review question. Population-based cross-sectional and cohort studies on: Geographic variations of one or more biological proxies of CMRFs, with/without an analysed contextual association with ASED were included. Two independent reviewers screened the studies and PRISMA guidelines were followed in the study selection and reporting.

Result: A total of 265 studies were retrieved and screened to 24 eligible studies. The review revealed reports of variation in the distribution of CMRFs, at varying geographic scales, in multiple countries. In addition, consistent inverse associations between ASED and of CMRFs were demonstrated. The reports were mainly from industrialised nations and small-area geographic units were frequently used.

Conclusion: Geographic variation in cardiometabolic risk exists across multiple spatial scales and is positively associated with ASED. This association is independent of individual-level factors and provides an imperative for area-based approaches to informing policy and health service organisation. The study protocol is registered in International prospective register of systematic reviews (Register No: CRD42018115294) PROSPERO 2018.

Keywords: *Cardiometabolic risk factors, Area-level socioeconomic disadvantage, Geographic variation*

2.3 Introduction

Cardiovascular disease (CVD) associated metabolic risk factors represent major global public health concerns. CVD is the leading cause of human death, accounting for 17.7 million (31%) of the 56.4 million total deaths reported worldwide in 2015.[1] Coronary heart disease (7.4 million) and stroke (6.7 million) were responsible for the greatest mortality within CVD and have remained the leading cause for mortality for the last 15 years.[2] CVD and its associated metabolic risk factors are listed in the top 15 causes of Disability Adjusted Life Years (DALY) globally.[3] In keeping with historical trends, deaths due to CVD are projected to increase steeply and reach more than 23.6 million annually by 2030.[4]

An important way to control CVD is by focussing on reducing associated metabolic risk factors. In low resource settings, vulnerable and disadvantaged groups are more likely to be exposed to unhealthy products and practices and develop metabolic risk factors for the development of CVD.[5] Cardiometabolic risk factors (CMRFs) such as diabetes mellitus (DM), hyperlipidaemia, high body mass index (BMI) and chronic kidney disease (CKD) can predispose and worsen CVD. Individual level approaches to prevent and control these risk factors have demonstrated limited success as evidenced by their increasing rates.[6-8] Thus it is important, in addition, to discern the contextual associations of development of these risk factors to assist in mitigating this global epidemic.

Geographic inequalities in the distribution of CMRFs at varying scales are reported in multiple studies from different countries in association with area-level socioeconomic disadvantage (ASED). Critically examining the area-level distribution patterns and associated area-level disadvantages reported in these studies may deepen our understanding of the higher prevalence of CMRFs in some geographic areas. Most recent relevant reviews in this area have broadly covered the influence of physical, social and service environment characteristics on CVD risk.[9-12] Systematic synthesis of evidence regarding this globally reported variation and association may inform policy development and healthcare service planning to detail area-level approaches, in addition to the individual level measures, to prevent and control CMRFs effectively.

Therefore, the questions attempted to answer in this review are: Is there any geographic variation in the distribution of CMRFs among adult population (aged 18 years and above) across the world and is this variation associated with ASED. The studies expected to include were epidemiological or population based cross sectional and/or cohort studies.

2.4 Methods

A review protocol was developed and registered in International prospective register of systematic reviews, PROSPERO 2018 (Register No: CRD42018115294) Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018115294.

Four databases: MEDLINE (Ovid), PubMed, Scopus and Web of science databases were chosen for the search, considering the breadth of fields they cover and the interdisciplinary nature of the review question. Also, hand-search of related articles served as an ‘other source’ of studies. The database search strategy commenced with two general search domains: 1) studies on CMRFs in single and multiple reporting forms; and 2) geographic and spatial health studies. An intersectional retrieval of studies from both these domains yielded a narrower list of studies on geographic variation in CMRFs. A third domain 3) studies addressing area-level measures of socioeconomic disadvantage were further intersected with the retrieved studies to

create a focal list of studies addressing geographic association of CMRFs with ASED. This approach maximised the number of potentially eligible studies identified compared to using single domain searches. Figure 2.1 conceptualizes the major search domains and their intersections used in the review.

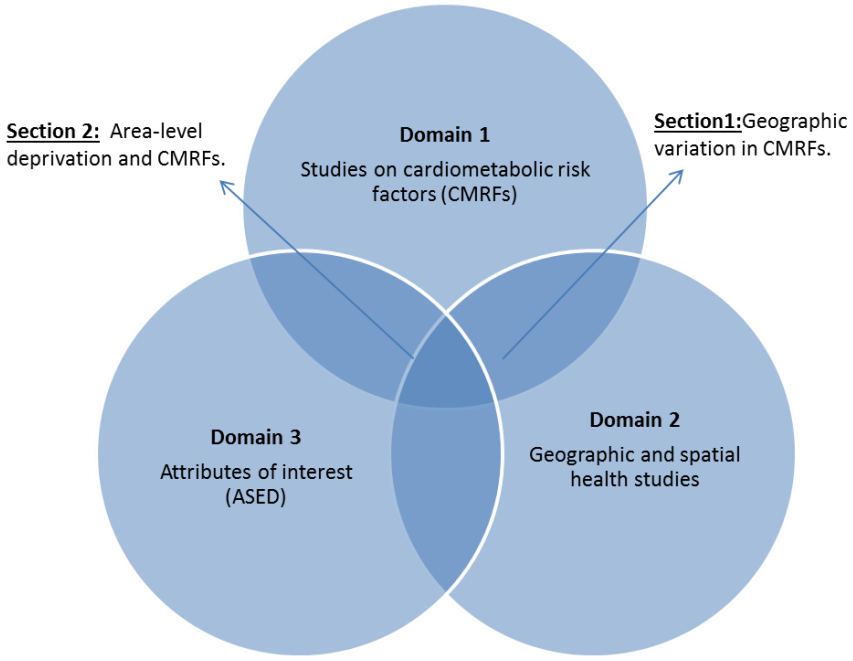

Figure 2.1: Conceptual representation of the literature search strategy

The review included epidemiological or population-based cross-sectional and cohort studies on: geographic variation of one or more biological proxies of CMRFs, with/without an analysed contextual association with ASED. Obesity, diabetes mellitus (DM), hyperlipidaemia and indices of low kidney function were the included biological proxies of CMRFs. Hypertension is included only when reported with other biological proxies of CMRFs, but not independently considering its limited summation into an overall cardiometabolic risk in an individual. Studies involving: type 1 DM and gestational DM were excluded as they were out of scope for the current review pertaining the geographic or area based contexts of the CMRFs. Studies measuring area-level characteristics other than ASED were also excluded.

All search outcomes were limited to: human studies; adult population (>=18 years) and availability in English language. The initial search included studies from year 1995; and latter it was modified to 01/01/2001 due to minimal publications on the review topic between the years 1995 to 2000. The search was last updated on 30/11/2018. Adopted search strategy in Ovid MEDLINE and search result URLs of remaining databases are available in Appendix I.

All retrieved studies were screened by two independent reviewers in three stages to reduce the risk of bias. In stage 1, articles from all databases were combined and screened to remove duplicates. Titles and abstracts of remaining articles were screened for eligibility, in stage 2. The final stage of study selection was done after full text reading of the remaining studies. Qualities of the individual studies were assessed using the STROBE checklist for cohort, case-control and cross-sectional studies (www.strobe-statement.org). The second coder repeated all three stages in parallel and selected studies were matched at the conclusion of each stage and any differences were resolved by consensus and arbitration. Other review team members served as additional reviewers when required.

Data extraction and coding of the chosen studies were carried out using two pilot-tested templates for consistency. Template 1 focused on the geographic variation in CMRFs and was used to extract information on: author, year, nation, study design, sample size and characteristics, geographic unit of reporting, studied CMRFs and the study outcome. Data on behavioural risk factors were not extracted as these were not included in the current review. Template 2 addressed the association of ASED and cardiometabolic risk factor distribution and extracted additional data on the reported proxies of ASED and its association status. An additional template was used for thematic mapping of the data in included studies for further qualitative syntheses. Study origin, representation, nature of problem, ecological context and evidence strength were the mapped themes.

The two independent review authors extracted and coded the data and any discrepancies were resolved through discussions between the authors. Summary measures used in this review are descriptive and based on the frequency of relevant studies to its denominator. Endnote software was used to keep track of the bibliographic details of the studies throughout the selection and data extraction process.

2.5 Results

A total of 265 individual studies were retrieved from four electronic databases (n=251) and hand searches of reference lists (n=14). Studies from electronic data bases included 91 Ovid Medline, 80 PubMed, 58 Scopus and 22 Web of science. Figure 2.2 shows the screening process as per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses PRISMA guidelines (www.prisma-statement.org).

Stage 1 screening combined studies from all sources and removed the duplicates (n=99). Duplicates in removed order: Ovid Medline (n=0), PubMed (n=80), Scopus (n=10), Web of science (n=3) and hand-searches (n=6). After removing duplicates, 166 studies were forwarded for stage 2 screening.

Stage 2 screening excluded 130 studies based on title and abstract screens, forwarding 36 studies for the full text screen. Studies excluded in stage 2 mainly addressed genetic, cellular, instrumental or pharmacological research regarding CMRFs. Studies on type 1 DM, paediatric or juvenile DM and gestational DM were also excluded at this stage as per the exclusions stated. Stage 3 screening carefully considered the whole full text of articles and 12 records were excluded with reason (list available in Appendix II) leaving 24 studies for the systematic synthesis.

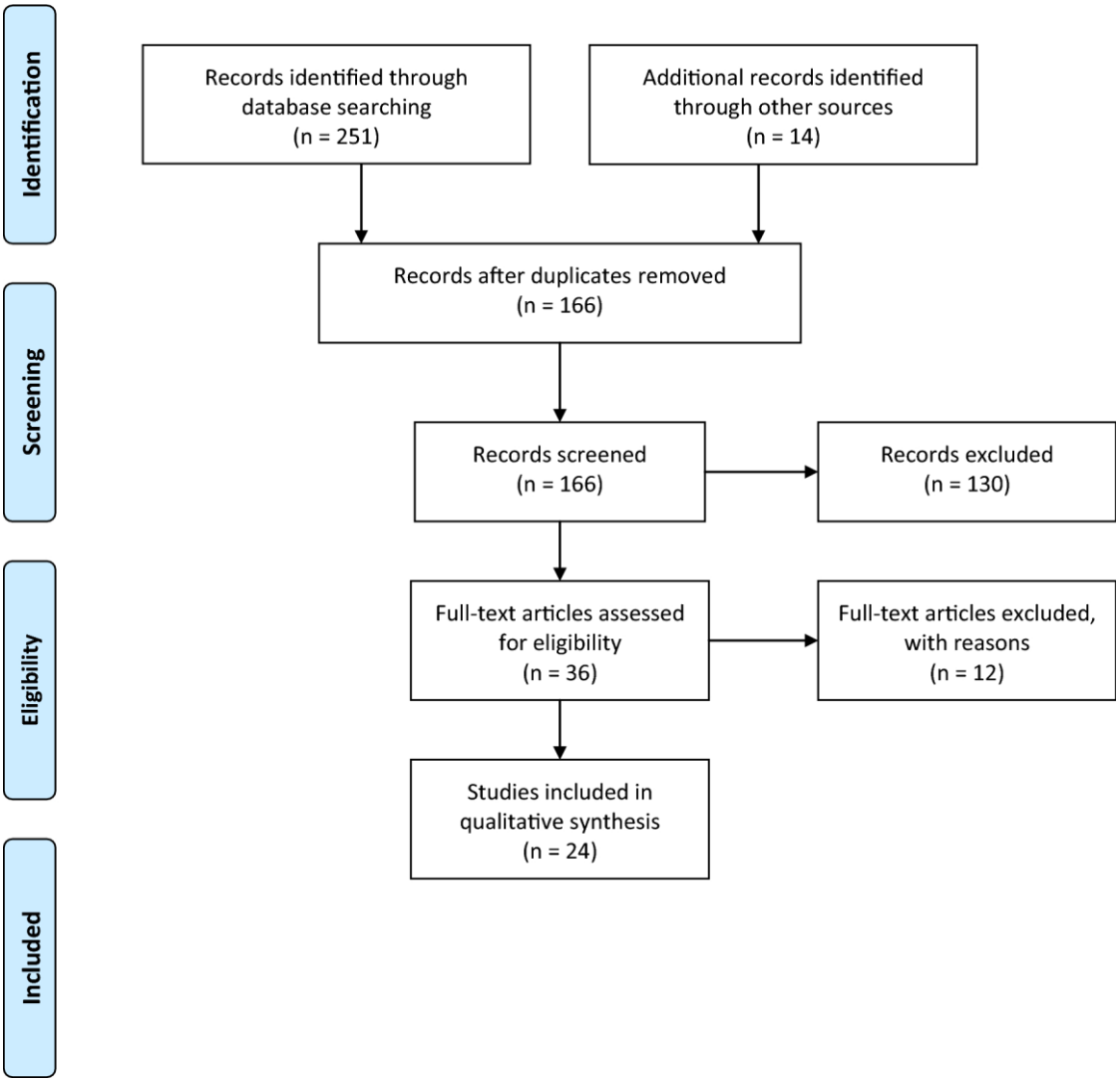

Figure 2.2: Combined PRISMA Flow Chart of the study selection.

The review is structured into three sections. Screened research articles retrieved through 'AND' intersections of search domain 1 and 2 (n=8) are reviewed in section 1: Geographic and spatial variation in cardiometabolic risk factors. Screened articles retrieved by intersecting domains 1 and 3 (n=16) are reviewed in section 2: Area level deprivation and cardiometabolic risk. Overall synthesis based on the total reviewed studies (n=24) are presented in section 3: Overall synthesis of the studies.

2.5.1 Geographic and spatial variation in cardiometabolic risk factors.

Table 2.1 summarizes the eight studies reviewed under this section.[13-20] Geographic variation in the population distribution of one or more CMRFs is reported in each of these studies. Most of the studies (7/8) reported hyperglycaemia as an important biomarker displayed geographic variation in cardiometabolic risk[13-15, 17-20], followed by dyslipidaemia (4/8), body mass index (4/8), blood pressure (BP) (3/8) and reduced glomerular filtration rate (GFR) (1/8).

All studies reported geographic variation in the distribution of CMRFs, regardless of the geographic unit of analysis used.[13-20] Most of these studies were from Europe (4/8), predominantly from Western Europe (3/8).[13, 16, 18, 19] These reports were from UK[16], Spain[19], France[18], and Luxembourg[13]. In the UK, geographic variation in the prevalence of risk factors such as obesity,

smoking, diabetes, hypertension and high cholesterol were reported across four main regions: South England, Midlands and Wales, Scotland and North England.[16] A higher prevalence of CMRFs was reported in southern Spain (Andalusia), which was found in close association with sedentary lifestyle and markers of socioeconomic disadvantage.[19] Variation in the distribution of diabetes, high BMI (≥25 kg/m^2), abdominal obesity, hypertension, high cholesterol and low glomerular filtration rate were reported at both canton and municipality levels in Luxemburg, Western Europe.[13] BMI and resting heart rate were reported to have greater geographic variation among matched cohorts in France and Australia.[18]

Other reports in this section were from Oceania (2/8), East Asia (2/8) and North America 1/8) - sourced from Australia, China, South Korea and US.[14, 15, 17, 18, 20] A geographic variation of 42% was reported in the odds of being diagnosed with DM among adults in Sydney, Australia.[14] In another Australian metropolitan based cohort, Glycated haemoglobin (HbA1c) was reported to have greatest geographic variation among matched cohorts in Australia and France.[18] In China, significant variation in the regional prevalence of diabetes was reported after adjusting for age, sex and urban/rural socioeconomic circumstances.[20] Geographic clustering of CMRFs were reported at administrative district level in South Korea.[17] The presence of a 'diabetic belt' with higher prevalence of diagnosed diabetes (>11.0%) was reported in the United States, consisting of 644 counties in its 15 mostly southern states.[15] Though the risk profiles and parameters varied, all these studies consistently reported geographic variation in its CMRFs.

The geographic scales of area-based units reported in all these studies ranged from large regions[15, 16, 18], within countries to smaller jurisdictional administration units [13-15, 17, 18], and trended towards smaller geographic areas over time. Easily accessible pre-existing geographic units and boundaries were used in these studies but most weren't explicit on the spatial extension and average population within their geographic units. Three studies had relied only on self-reports on anthropometric, behavioural, biochemical, physiological and diagnostic categories of data, risking for recall bias and misclassifications.[14, 15, 17]

Table 2.1: List of studies reviewed on geographic variation in CMRFs.

First Author Country (year)	Sample Age group	Design Geographic units	CMRFs [a] (Data source)	Outcome
Lawlor D A UK (2003)	4, 286(women) 60-79 years	Cross-sectional 4 Regions within country	HT, BMI , LDL, TC (Data collected)	Geographic variation
Barker L E USA (2011)	813, 498 DM ≥18 years	Cross-sectional 644 counties in proximity	DM prevalence >=11.0%, (Behavioral Risk Factor Surveillance)	Geographic diabetic belt
Valdes S Spain (2014)	5, 103 adults ≥18 years	Cross-sectional 2 region within country	BP, BMI , FPG, TC, WC (Di@bet.es study)	Geographic coherence
Astell-Burt T Australia (2014)	114, 755 adults ≥45 years	Cross-sectional ~40 Local Government areas	DM (The 45 and Up Study)	Geographic variation
Zhou M China (2015)	98, 058 adults >18 years	Cross-sectional multilevel 31 Provinces in country	DM (National health survey)	Geographic variation
Paquet C AU-France (2016)	Au: 3, 893(≥18 years) Fr: 6, 430 (30- 79 years)	Cross-sectional multilevel Au: 767 CDs (SS, POA, SLA LGA). Fr: 1866 IRIS (TRIRIS, Municipalities)	BP, BMI, WC, FG, HbA1c, HR, TC, HDL, TG, (Au: NWAHS study , Fr: RECORD Cohort Study)	Geographic clustering(ICCs)
Alkerwi A Luxemberg(2017)	1, 432 subjects 18-69 years.	Cross-sectional multilevel 106 Municipalities (12 cantons)	BMI, FPG, TC, GFR (ORISCAV-LUX national survey)	Geographic variation
Oh W S South Korea(2018)	228, 921 people ≥ 19 years	Cross-sectional 230 administrative districts	HT, DM (Korean Community Health Surveys)	Geographic clustering

Abbreviations: AU – Australia; CD -Census Collection District; POA -Postal Area; SLA -Statistical Local Area; LGA - Local Government Area; IRIS - Ilôts regroupés pour l'information statistique; TRIRIS - Groups of around three IRIS areas; BP – Blood pressure; BMI – Body mass index; DM – Diabetes mellitus; FBG – Fasting plasma glucose; FPG – Fasting glucose; HbA1c – Glycated haemoglobin; HR – Heart rate, HT – Hypertension, TC- Total cholesterol; TG – Triglycerides; LDL – Low density lipoprotein; GFR – Glomerular filtration rate; WC – Waist circumference

2.5.2 Area level deprivation and cardiometabolic risk.

Table 2.2 summarises the 16 studies reviewed under this section.[21-36] Reported studies were mainly from Europe (7/16) and North America (7/16), followed by Oceania (1/16) and South America (1/16). Studies from Europe were predominantly reported from the western region and sourced from UK, Germany, Czech and France. Reports from North America were mainly from USA (6/7) and Canada (1/7). There was only one study from Oceania, sourced from Australia.[25] Most of these studies were sourced from industrialised nations, except one study from Brazil[21], a developing nation in South America.

All studies reported inverse associations of various CMRFs with ASED.[21-36] Various measures of the biological proxies of CMRFs reported include biochemical, anthropometric, physiologic, behavioural and diagnostic categories of data. Census sourced data on ASED were used in most of these studies (12 /16), whereas other survey sourced data were used in the remaining studies (4/16) to construct summary scores or indices on ASED. The categories of measures used to calculate ASED in these studies were area-level proportions of: median income, education, occupation, housing, transport, dependent population, social class, social capital, environment, security, family structure, disability, internet access and insurance coverage. A minimum of one category of these measures are used in all the studies.[21-36]

The samples characteristics and variables considered were notably heterogeneous across studies. Sampling frame of most (7/16) of these studies were population based lists, however service provider given (4/16) and employees (3/16) lists were also used. Two studies had used a combination of both population lists and service provider given lists.[28, 32] Though subjects in all studies qualified adult age limits (>=18 years), divergent age groups were sampled across all of the studies. Also gender[33, 34], and race[22, 24], specific sampling were done in two studies each. Heterogeneity of these sample characteristics makes a comparison and further quantitative synthesis more difficult.

The samples are mostly accessed from existing study cohorts, laboratory databases, national surveys and audit lists. The sample size of studies ranged from 342 adults to a maximum of 91, 776 adults, mostly larger in size. Census administration units were the most commonly used neighbourhood proxy, followed by other administrative units and electoral wards. Pre-existing geographic boundaries were mostly adopted to define the spatial unit, but their spatial extents of the unit of analyses were not stated in most of the studies.

Table 2.2: List of studies reviewed on the association of area-level deprivation and cardiometabolic risk factors.

First Author Country (Year)	Sample Age group	Design Spatial unit	CMRFs* (Data source)	Proxies of ASED (Data source)	Association
Bonney A Australia (2015)	91, 776 adults 55.2 ±15.66	Cross-sectional higherarchical 631 Census collection districts	BMI (The SIMLR Study)	Index of Relative Socioeconomic Disadvantage ((Australian Census 2006)	+ve (women)
Unger E USA (2014)	5, 805 adults 45 -84 years	Prospective cohort higherarchical Census tract level	BMI, BP, BS, TC - CVH score (The MESA study)	Neighbourhood SES (constructed summary score)	+ve
Maier W Germany (2014)	33, 690 adults <30 years	cross-sectional design 412 Districts	T2DM, obesity (GEDA national health interview survey)‘	German Index of Multiple Deprivation score (assessed by GIMD)	+ve (women)
Silhol R France (2011)	19, 808 adults 35-50 years	cross-sectional cohort Municipality level	Incidence of CHD (French GAZEL cohort Data)	Area socio - economic position (French Census 1990)	-ve
Naimi A I Canada (2009)	342 adults 18–55 years	cross-sectional, 250m respondent-centred moving window buffer within 7 census tracts.	BMI, HbA1c, TG, TC, HDL – TCR (Montreal Neighbourhood Survey of Lifestyle and Health)	Area-level unemployment (Canada Census 2001)	+ve
Cox M Scotland (2007)	3, 917 adults < 35 years	cross-sectional 3382 Census- Output Areas(OA)	T2DM (DARTS Diabetes Audit and Research Tayside Scotland dataset)	Area deprivation (The Carstairs score based on 2001 Scotland census data)	+ve
Andersen A UK (2008)	4, 286 women 60 -79 years	Cross-sectional 457 British Electoral wards	T2DM, FBG, IR (British Women’s Heart and Health Study)	Area deprivation (The Carstairs score based on 2001 census data)	+ve
Gabert R USA (2016)	63, 053 DM 18-74 years	Retrospective observational 120 zip code areas	BP, HbA1c, LDL (Minnesota Community Measurement electronic health records)	Area-level indicators of SES (based on American Community Survey 2013)	+ve
Dragano N GR-Czech (2007)	GR: 4, 814 adults CZ: 8, 856 adults 57.7±6.6 years	2 longitudinal cohort studies 326 pre-existing administrative units	Obesity, HT (GR: 'Heinz Nixdorf Recall (HNR) Study', Czech: 'Health,	Area-level socioeconomic status (based on census data)	+ve

			Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) Study')		
Cubbin C Sweden (2006)	18, 081 adults 25 - 64 years	Pooled cross-sectional data 8624 SAMS neighbourhoods	Obesity, DM, HT (Swedish Annual Level of Living Survey (SALLS), 1988–89)	Neighbourhood deprivation (assessed by Care Need Index (CNI) 1997 data)	+ve
Mujahid M S USA (2005)	13, 167 adults 45 -64 years	Crosssectional and longitudinal (3-9 years) Census block	BMI (The Atherosclerosis Risk in Communities ARIC Study)	Neighbourhood SES score (1990 U.S. Census1990)	-ve
Lawlor D A UK (2005)	4, 286 women 60 -79 years	Cross-sectional 457 Electoral wards	Coronary heart disease (British Women's Heart and Health Study)	Residential area deprivation(The Carstairs score based on 1991 UK census data)	+ve
Roux A V D USA (2002)	3, 093 adults 28–40 years	Cross-sectional 10 years follow up 2, 260 Census block (in 45 states).	BMI, HDL, TG, BP, FI & FG -IRS (Coronary Artery Risk Development in Young Adults CARDIA Study)	Neighbourhood SES score (1990 U.S. Census)	-ve
Keita A D USA (2014)	19, 079 black/white age > 45 years	Cross-sectional cohort Census block group	Obesity, WC, BP, FBG, TG, low-HDL (REGARDS study).	Neighborhood socioeconomic deprivation(US Census 2000)	+ve (black/ white)
Clark, C.R USA (2013)	3, 909 Afro-Americans 35 - 84 years	Cross-sectional cohort 102 Census tracts	TG, FBG, BP, WC, low-HDL (Jackson Heart Study).	Neighborhood socioeconomic disadvantage (US Census 2000)	+ve (women)
Barber et al Brazil (2018)	10617 adults 35 -75 years	Cross sectional cohort Study defined clusters of contiguous census tracts	DM and HT (Brazilian Longitudinal Study of Adult Health)	Area level economic residential segregation (IBGE census 2010)	+ve

Abbreviations: BMI – Body mass index; BP – Blood pressure; BS – Blood sugar; CHD – Coronary heart disease; CVD – Cardiovascular disease; CVH – Cardiovascular health; DM – Diabetes mellitus; eGFR – estimated Glomerular filtration rate; FBG – Fasting blood glucose; FG – Fasting glucose; FI – Fasting insulin; GR- Germany; HbA1c – Glycated haemoglobin; HDL – High density lipoprotein; HT – Hypertension; IR – Insulin resistance; IRS – Insulin resistance syndrome ; LDL – Low density lipoprotein; SES – Socioeconomic status; TC- Total cholesterol; TCR – Total cardiometabolic risk; T2DM – Type 2 diabetes mellitus; TG – Triglycerides; SAMS– Small area market statistics.

2.6 Overall synthesis of the studies

Significant features of the included studies were identified to aid synthesis of the findings. These features were the origin of the study, its representativeness, nature of the CMRFs studied, the ecological context and the strength of evidence presented. These features were then formulated into five themes, mapping the related data for further analyses (Table 2.3).

We had plotted all the studies to identify their global region of origin and the economic nature of the source country. Most of the studies published were from Europe (11/24), closely followed by America (9/12), (two studies were cross national, hence counted under both the nations and corresponding regions). Fewer publications were found from Oceania (3/24) and Asia (2/24).However no identified studies were from Africa. Studies from developing nations were fewer (3/24) compared with studies from industrialised nations (21/24). This emphasises a gap in related publications from Asia-pacific and African regions, especially from nations of developing and underdeveloped economies. The global representativeness of this review is hence limited and the review findings may be more generalizable to industrialised nations.

The target populations for included studies are shown in Table 2.3. The sample frame of most of the studies were population based lists (13/24 studies), however service providers' lists (5/24) and employees lists (3/24) were also used. Both population and service providers' lists were used in three studies (3/24). All the population based studies used a random sampling technique to ensure the population representativeness. However, the response rates varied (15 - 90.5%) in these studies. Two studies had a response rate < 50%, suggesting a risk of responder bias despite a probability sampling method being employed.[31, 35]

Ecological contexts of the included studies were analysed by extracting area level characteristics (Table 2.3). Area level units used in these studies extended from small areas (10/24), to medium areas (9/24) and large areas (5/24). Small-area units were mostly based on census, administrative or zip code area with an average ~1000 residing population. Medium area units had an average ~ 5000 population and the large area units were mostly regions, provinces and districts. ASED gradients were based on area level measures of ranged from 1 to 7 measures, however single measures of income or overcrowding as an indirect proxy of ASED raised concerns regarding their comprehensiveness in comparison to aggregate measures of ASED.

The nature of CMRFs and the strength of evidence in relation to associations with outcomes were mapped by extracting data on the categories of CMRFs measured, the source of data and the mode of analyses (Table 2.3). Biological proxy categories of CMRFs were mostly biochemical (18/24), followed by anthropometric (18/24), physiologic (15/24) and diagnostic (4/24) in nature. Self-reported data on these categories of CMRFs had the highest risk for misclassification due to reporting bias or errors. Studies which adopted a combined mode of both statistical and spatial analyses provided a better ecological context of CMRFs than with statistical analyses alone.

Table 2.3: Thematic mapping of data categories from all included studies

Theme Data Map	Study origin : Reference Nation (status) Region	Representation: Sample frame Sampling Response or retention %	Ecological context: Geographic unit and/or ASED	Nature of problem: Cardiometabolic risk nature	Evidence Strength: Data source Analyses
	Roux *et al.* (2002) USA(Industrialised) North America	Population & Service provider's lists 79% retention	Small area ASED: income, education, Occupation.	Biochemical, Anthropometric Physiological	Self-report, PE, Specimen tests Statistical
	Lawlor *et al.* (2003) UK (Industrialised) Western Europe	Service provider's list Random 60% response	Large area	Biochemical, Anthropometric, Physiological	Self-report, PE, Specimen tests, MR Statistical
	Mujahid *et al.* (2005) USA (Industrialised) North America	Population list1 Random 81% retention	Small-area ASED: income, education, Occupation.	Anthropometric	Self-report, PE Statistical
	Lawlor *et al.* (2005) UK (Industrialised) Western Europe	Service provider's list Random 60% response	Median area ASED: Employment, housing, transport, social class	Biochemical, Anthropometric, Physiological	Self-report, PE, Specimen tests, MR Statistical
	Cubbin *et al.* (2006) Sweden(Industrialised) Northern Europe	Population list2 Random4 ~80% response	Small-area ASED: population structure, education, unemployment etc.	Anthropometric, Physiological, Diagnostic: DM	Self-report Statistical
	Cox *et al.*(2007) Scotland (Industrialised) Western Europe	Service provider's list ~Purposive	Small-area ASED: Employment, housing, transport, social class	Diagnostic – T2DM	Medical record Spatial & Statistical

	Dragano *et al.* (2007) GR-Czech(Industrialised) Western - Central Europe	Population list Random 56% & 55% responses	Small-area ASED :unemployment overcrowding	Anthropometric, Physiological,	Self-report, PE Statistical
	Andersen *et al.* (2008) UK (Industrialised) Western Europe	Service provider's lists Random 60% response	Small-area ASED: Employment, housing, transport, social class	Biochemical	Self-report, PE, MR Statistical
	Naimi *et al.* (2009) Canada (Industrialised) North America	Population list Stratified cluster sampling 15% response	Medium area ASED: Education employment	Anthropometric, Biochemical	Self-report, PE, Specimen tests Statistical
	Barker *et al.* (2011) USA (Industrialised) North America	Population list Random 50.6 % response	Medium area	Anthropometric, Biochemical	Self-report Statistical
	Silhol *et al.* (2011) France (Industrialised) Western Europe	Employees lists ~purposive	Medium area ASED:Higher-job, education	Anthropometric, Biochemical Physiological,	Self-report, Employers data, Insurance data Spatial & Statistical
	Keita *et al.* (2014) USA (Industrialised) North America	Population list	Small-area ASED: income, housing education & occupation	Biochemical, Anthropometric Physiological	Self-report, PE, Specimen tests Statistical
	Clark *et al.* (2013) USA (Industrialised)	Population list Random	Medium area ASED :10components	Biochemical, Anthropometric Physiological	PE, Specimen tests Statistical

	North America 6				
	Valdes *et al.* (2014) Spain (Industrialised) Southern Europe	Population list Cluster –random 54.6% response	Large area	Anthropometric Physiological Biochemical	Self-report, PE, Specimen tests Statistical
	Astell-Burt *et al.* (2014) Australia (Industrialised) Oceania	Population (insurance) lists Random	Medium area	Biochemical, Physiological	Self-report Spatial & Statistical
	Unger *et al.* (2014) USA (Industrialised) North America	Population and service provider's list (~purposive)	Medium area ASED: Income, housing, education, occupation.	Anthropometric Biochemical Physiological	Self-report, PE, Specimen tests Statistical
	Maier *et al.* (2014) Germany(Industrialised) West - Central Europe	Population list Random 29.1% response	Large area. ASED :income, employment, education, revenue, social capital, environment, security	Anthropometric Diagnostic – T2DM	Self-report Statistical
	Zhou *et al.* (2015) China (Developing) East Asia	Population (survey) list Random 90.5% response	Large area	Anthropometric Biochemical	Self-report, Specimens Statistical
	Bonney *et al.* (2015) Australia (Industrialised) Oceania	Service provider's list (~purposive)	Small-area ASED : income, education, employment, family structure, disability, housing,	Anthropometric	Medical record Statistical

			transport and internet connection		
	Gabert *et al.*(2016) USA (Industrialised) North America 8	Employees list 83.6% response	Small-area .ASED: income, education, insurance	Biochemical	Medical record Spatial & Statistical
	Paquet *et al.*C (2017) AU-France (Industrialised) Oceania - West Europe	Australia: Population list (Random 13/ 49.4% response 3) France: Employees list (Purposive /83.6% response)	Small-area	Anthropometric Biochemical Physiological	PE, Specimen tests Spatial & Statistical
	Alkerwi *et al.* (2017) Luxemberg (Industrialised) Western Europe	Population (survey) list stratified random 32.2% response	Medium area	Physiological Biochemical	Self-report, PE, Specimen tests Spatial & Statistical
	Oh *et al.* (2018) South Korea (Developing) East Asia	Population (ministry) lists ~purposive	Medium area	Biochemical Physiological Diagnostic	Self-report Spatial & Statistical
	Barber *et al.* (2018) Brazil (Developing) South America	Employees lists ~purposive	Large area ASED: income	Biochemical Anthropometric Physiological	Self-report, PE, Specimen tests Spatial & Statistical

2.7 Discussion

The systematic review presented in this chapter provides a critical review of the geographic and socioeconomic variations of CMRFs reported from various nations. Studies which met the inclusion criteria were mostly from the industrialised nations of Europe and America (20/24). This emphasises a gap in reports from Asia-Pacific and African regions, especially from nations of developing and underdeveloped economies. Thus, the findings from these reports may be more generalisable to industrialised nations rather than providing overall global representativeness. In addition, heterogeneity of sample characteristics and divergence of reported CMRFs in these studies makes comparisons and further quantitative synthesis difficult. Also, the presence of self-reported data and reportedly low sampling response rates (< 50%) suggest risks of responder and recall bias in at least some of these studies. From a methodological perspective, place of living as the primary focus in these studies; an over-emphasis on fixed effects estimates; and underreporting of the components of geographic variance are important reporting gaps. Based on the findings from this review, it is recommended greater attention is given to a more balanced and consistent reporting of studies and the components of variation in the reporting of geographic inequalities of CMRFs' distribution.

ASED was repeatedly demonstrated to be associated with higher cardiometabolic risk. Higher ASED was consistently reported to have an association with cardiovascular risk; whereas lower ASED was associated with reduced cardiovascular risk. Such associations were often demonstrated independently of individual level characteristics such as socioeconomic status, education and duration of exposure to area. Type 2 diabetes and high body mass index (BMI) were reported to be more prevalent in disadvantaged areas. Related studies report that the type of neighbourhood food outlets [37-39], poor physical activity resources[38], individual perception of area level features[40], residential density and service availability[9], were all explanatory variables associated with cardiometabolic risk factors among people living in disadvantaged neighbourhoods.

Related systematic reviews published in this area of research investigate associations for different geographically distributed factors with CVD. Chaix (2009) reviewed the associations between neighbourhood social environments and CHD and proposed a theoretical model of a mediating mechanism focussing on the social interactional environment.[9] Consistent associations of obesity or hypertension with lower levels of area socioeconomic status, urbanization, street intersection, accessibility to supermarkets, social cohesion, service availability and residential density; and higher levels of noise pollution and density of convenience stores, were reviewed and reported by Leal (2011).[10] Frequent inverse associations of the common indices of ASED with childhood obesity were reported in the UK.[12] Consistent associations between socioeconomic disadvantage and central adiposity were reported by Slopen (2013).[11] All these reviews report important methodological inadequacies and the need for further research in this area, which support the findings of the current review.

Recent advances in Geographic Information Systems (GIS) and analytical approaches were utilised in the studies reporting geographic variation in CMRFs. These studies have demonstrated advances in various analytical tools and the potential for plotting area level risk parameters. Geocoding and mapping of existing large population based datasets has become feasible with newer computational tools through linking location data; such as map co-ordinates, addresses or postcodes.[41] These tools have the capacity to visually display area based factors, in contrast with traditional table and graph methods and this has the potential to enhance impact on subsequent area level health care policy development and resource allocation.[42-44] In addition,

systematic quantitative analyses are possible with these spatial tools which create opportunities to investigate the role of environmental factors in explaining any geographic aggregations beyond random effects.[45]

National estimates of CVD have limited utility in informing prevention and management of CVD within discrete communities. The disease patterns at smaller areas may significantly differ from national and regional prevalence reports, thus small-area analysis is important in order to understand local patterns and requirements.[46] Small-area level analyses also have the potential to reveal area level contexts and dependencies of CMRFs and such analyses can highlight areas for targeted preventive interventions.

CVD and its associated CMRFs continue to evolve as a major global health threat. It is the highest cause of mortality and the highest absorber of health care expenditure in many developed nations.[7, 47, 48] Once diagnosed, the ongoing costs of care and productivity loss due to consequent disability and premature death creates a large economic burden not only to the individual and family, but to the nation—especially when half the people dying are found to be in their prime productive years.[49] Thus, CVD and its associated metabolic risk factors emerge as a threat not only to human health and life, but to the sustainable development and economies of nations. Hence, improving public health program effectiveness in reducing CVD must be a research priority.

2.8 Limitations

Firstly, the cross sectional nature of the reviewed studies precluded causative interpretations. Second, the global representativeness of the review is limited mainly due to publication gaps from Asia-pacific and African regions of the World. Third, the scope of our review excluded examination of behavioural, dietary and activity related risk factors and also other area level characteristics to focus only on the biological proxies of CMRFs. The risk factors reviewed in this chapter are largely considered as modifiable. However, an exact distinction of familial, genetic or non-modifiable risk factors was not possible as the review focus was on the biological proxies of CMRFs such as obesity, diabetes mellitus, hyperlipidaemia, and indices of low kidney function. Also, the scope of review excluded examination of behavioural, dietary and activity related risk factors, as per the registered review protocol. Fourth, methodological heterogeneity within the retrieved studies prohibited a meta-analytical synthesis of the findings. The sample characteristics, geographical scales and the CMRFs' risk profiles varied substantially across the studies impeding any further quantitative synthesis. Finally, the systematic review presented in this chapter does not exclude the possibility of any simultaneity and/or endogeneity bias within the reported studies. Simultaneity refers to the co-occurrence and symbiotic existence of CMRFs with ASED[50-52], whereas endogeneity would indicate on the influence of unmodeled covariates which might be having an effect on both CMRFs and ASED[53]. Endogeneity can also occur when individuals choose to move to a higher ASED area because of an existing CMRF (reverse causation).[54] Examining the temporal sequence of occurrences would be required to learn more regarding the possibility of such effects within theses studies. However, this has not frequently been undertaken.

2.9 Recommendations and future directions

Finding geographic variation in CMRFs (if any) and its association with ASED may assist in understanding the contexts of risk. Such studies have the potential to inform contextual planning of interventions for prevention and management of cardiometabolic risk. However, most of the studies in this review do not report the spatial extents of their units of analysis. This is important as associations are likely to be different at different levels of aggregation and limits the ability to assess the likelihood of spatial scale effects in these studies[22, 24],

known as the Modifiable Areal Unit Problem (Openshaw & Taylor, 1979; Openshaw, 1984).[55, 56] When data are aggregated to larger geographic units, small-area anomalies may be diluted or smoothed over.[24] Using smaller rather than larger area scales can help to reduce the likelihood of missing important small-area anomalies.[57] Similarly, supplementing individual level data along with area level data could minimise group effects due to area level aggregation of data.[57] Leveraging both individual- and area-level data provides a more complete picture to inform planning, policy and practice.[45, 57] Future research directions should include hierarchical multilevel analyses to yield comprehensive picture of the contextual aspects of risk factors, to help aid both individual and area-level better preventive initiatives.

2.10 Conclusion

Cardiometabolic risk distribution varied significantly across different geographic scales reported in multiple studies. In addition, there is strong evidence that area-level disadvantage is significantly associated with CMRFs, irrespective of individual-level characteristics. This review highlights the need for area-based preventive approaches in addition to individual-level approaches to prevent and control CMRFs and their consequent CVD outcomes.

Reference

1. World Health Organization. *Cardiovascular diseases (CVDs): Key facts* [cited 2016 Nov 5]; Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
2. Schutzer, S.E., et al. *Whole-genome sequences of thirteen isolates of Borrelia burgdorferi*. Journal of bacteriology 2011 [cited 193 4]; 1018-1020]. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=med5&AN=20935092.
3. Murray, C.J. *Phil., Alan D. Lopez. Measuring the Global Burden of Disease*. New England Journal of Medicine 2013 [cited 369; 448-57]. Available from: http://www.nejm.org/doi/10.1056/NEJMra1201534.
4. Benjamin, E.J., et al., *Heart disease and stroke statistics—2017 update: a report from the American Heart Association.* circulation, 2017. **135**(10): p. e146-e603.
5. World Health Organization. *The top 10 causes of death*. 24 May 2018 [cited 2020 Sept 5]; Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
6. Brunzell, J.D., et al., *Lipoprotein management in patients with cardiometabolic risk: consensus statement from the American Diabetes Association and the American College of Cardiology Foundation.* Diabetes care, 2008. **31**(4): p. 811-822.
7. Cannon, C.P. *Cardiovascular disease and modifiable cardiometabolic risk factors*. Clinical cornerstone 2007 [cited 8 3]; 11-28].
8. Mahmood, S.S., et al., *The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective.* The lancet, 2014. **383**(9921): p. 999-1008.
9. Chaix, B., *Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda.* Annual review of public health, 2009. **30**: p. 81-105.
10. Leal, C. and B. Chaix, *The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda.* Obesity reviews, 2011. **12**(3): p. 217-230.
11. Slopen, N., et al., *Socioeconomic and other social stressors and biomarkers of cardiometabolic risk in youth: a systematic review of less studied risk factors.* PLoS One, 2013. **8**(5): p. e64418.
12. El-Sayed, A.M., P. Scarborough, and S. Galea, *Socioeconomic inequalities in childhood obesity in the United Kingdom: a systematic review of the literature.* Obesity facts, 2012. **5**(5): p. 671-692.
13. Alkerwi, A., et al., *Geographic variations in cardiometabolic risk factors in Luxembourg.* International journal of environmental research and public health, 2017. **14**(6): p. 648.
14. Astell-Burt, T., et al., *Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia.* Diabetes research and clinical practice, 2014. **106**(3): p. e68-e73.
15. Barker, L.E., et al., *Geographic distribution of diagnosed diabetes in the US: a diabetes belt.* American journal of preventive medicine, 2011. **40**(4): p. 434-439.
16. Lawlor, D., et al., *Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study.* Journal of Epidemiology & Community Health, 2003. **57**(2): p. 134-140.
17. Oh, W.S., et al., *Geographical variations and influential factors in prevalence of cardiometabolic diseases in South Korea.* PloS one, 2018. **13**(10): p. e0205005.

18. Paquet, C., et al., *Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia.* International journal of environmental research and public health, 2016. **13**(5): p. 519.
19. Valdés, S., et al., *Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@ bet. es study.* Revista Española de Cardiología (English Edition), 2014. **67**(6): p. 442-448.
20. Zhou, M., et al., *Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults.* Diabetes care, 2015. **38**(1): p. 72-81.
21. Barber, S., et al., *At the intersection of place, race, and health in Brazil: Residential segregation and cardio-metabolic risk factors in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil).* Social Science & Medicine, 2018. **199**: p. 67-76.
22. Clark, C.R., et al., *Neighborhood disadvantage, neighborhood safety and cardiometabolic risk factors in African Americans: biosocial associations in the Jackson Heart study.* PloS one, 2013. **8**(5): p. e63254.
23. Cox, M., et al., *Locality deprivation and Type 2 diabetes incidence: a local test of relative inequalities.* Social science & medicine, 2007. **65**(9): p. 1953-1964.
24. Keita, A.D., et al., *Associations of neighborhood area level deprivation with the metabolic syndrome and inflammation among middle-and older-age adults.* BMC Public Health, 2014. **14**(1): p. 1319.
25. Bonney, A., et al., *Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users–a cross-sectional study.* PLoS One, 2015. **10**(8): p. e0137261.
26. Gabert, R., et al., *Identifying high-risk neighborhoods using electronic medical records: a population-based approach for targeting diabetes prevention and treatment interventions.* PLoS One, 2016. **11**(7): p. e0159227.
27. Mujahid, M.S., et al., *Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics.* Obesity research, 2005. **13**(8): p. 1412-1421.
28. Unger, E., et al., *Association of neighborhood characteristics with cardiovascular health in the multi-ethnic study of atherosclerosis.* Circulation: Cardiovascular Quality and Outcomes, 2014. **7**(4): p. 524-531.
29. Cubbin, C., et al., *Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects.* Scandinavian journal of public health, 2006: p. 228-237.
30. Dragano, N., et al., *Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany.* BMC Public Health, 2007. **7**(1): p. 255.
31. Maier, W., et al., *Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update'GEDA 2009 and 2010.* PloS one, 2014. **9**(2): p. e89661.
32. Roux, A.V.D., D.R. Jacobs, and C.I. Kiefe, *Neighborhood characteristics and components of the insulin resistance syndrome in young adults: the coronary artery risk development in young adults (CARDIA) study.* Diabetes care, 2002. **25**(11): p. 1976-1982.
33. Andersen, A., et al., *Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study.* Diabetic medicine, 2008. **25**(12): p. 1462-1468.
34. Lawlor, D.A., et al., *Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study.* American journal of public health, 2005. **95**(1): p. 91-97.

35. Naimi, A.I., et al., *Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area.* International journal of environmental research and public health, 2009. **6**(12): p. 3082-3096.
36. Silhol, R., et al., *Investigating the spatial variability in incidence of coronary heart disease in the Gazel cohort: the impact of area socioeconomic position and mediating role of risk factors.* Journal of Epidemiology & Community Health, 2011. **65**(2): p. 137-143.
37. Astell-Burt, T. and X. Feng, *Geographic inequity in healthy food environment and type 2 diabetes: can we please turn off the tap?* Medical Journal of Australia, 2015. **203**(6): p. 246-248.
38. Christine, P.J., et al., *Longitudinal associations between neighborhood physical and social environments and incident type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA).* JAMA internal medicine, 2015. **175**(8): p. 1311-1320.
39. Millstein, R.A., et al., *Food availability, neighborhood socioeconomic status, and dietary patterns among blacks with type 2 diabetes mellitus.* The Medscape Journal of Medicine, 2009. **11**(1): p. 15.
40. Baldock, K., et al., *Associations between resident perceptions of the local residential environment and metabolic syndrome.* Journal of environmental and public health, 2012. **2012**.
41. Stevens, C.D., et al., *Geographic clustering of diabetic lower-extremity amputations in low-income regions of California.* Health Affairs, 2014. **33**(8): p. 1383-1390.
42. Angier, H., et al., *Using geographic information systems (GIS) to identify communities in need of health insurance outreach: an OCHIN practice-based research network (PBRN) report.* The Journal of the American Board of Family Medicine, 2014. **27**(6): p. 804-810.
43. Auchincloss, A.H., et al., *A review of spatial methods in epidemiology, 2000–2010.* Annual review of public health, 2012. **33**: p. 107-122.
44. Bazemore, A., R.L. Phillips, and T. Miyoshi, *Harnessing geographic information systems (GIS) to enable community-oriented primary care.* The Journal of the American Board of Family Medicine, 2010. **23**(1): p. 22-31.
45. Elliott, P. and D. Wartenberg, *Spatial epidemiology: current approaches and future challenges.* Environmental health perspectives, 2004. **112**(9): p. 998-1006.
46. Occelli, F., et al., *Mapping end-stage renal disease (ESRD): spatial variations on small area level in northern France, and association with deprivation.* PloS one, 2014. **9**(11): p. e110132.
47. World Health Organization. *WHO | Noncommunicable diseases* 2017 [cited 2018 10 March]; Available from: : http://www.who.int/mediacentre/factsheets/fs355/en/.
48. World Health Organization; World Heart Federation and World Stroke Organization. *Global atlas on cardiovascular disease prevention and control*. 2011; Available from: https://www.cabdirect.org/cabdirect/abstract/20123402600%0Afile:///C:/Users/USER/Downloads/9789241564373_eng (2).pdf.
49. Bloom, D.E., et al., *The global economic burden of noncommunicable diseases*. 2012, Program on the Global Demography of Aging.
50. Gerstman, B.B., *Epidemiology kept simple: an introduction to traditional and modern epidemiology*. 2013: John Wiley & Sons.

51. Katz, M., *Study design and statistical analysis: a practical guide for clinicians*. 2006: Cambridge University Press.
52. Stephanie Glen. *Reverse Causality: Definition, Examples From StatisticsHowTo.com: Elementary Statistics for the rest of us!* . [cited 2020 Nov]; Available from: https://www.statisticshowto.com/reverse-causality/.
53. Antonakis, J., et al., *Causality and Endogeneity*, in *The Oxford Handbook of Leadership and Organizations*. 2014.
54. Kawachi, I. and S.V. Subramanian, *Neighbourhood influences on health.* Journal of epidemiology and community health, 2007. **61**(1): p. 3-4.
55. Openshaw, S., P.J. Taylor, and N. Wrigley, *Statistical applications in the spatial sciences.* London: Edited by N. Wrigley. Pion, 1979: p. 127-144.
56. Openshow, S., *A million or so correlation coefficients, three experiments on the modifiable areal unit problem.* Statistical applications in the spatial science, 1979: p. 127-144.
57. Wakefield, J.H., *Spatial aggregation and the ecological fallacy. Handbook of spatial statistics. 2010*.

Chapter 3

Materials and Methods

Chapter 3: Materials and Methods

In order to appropriately address the hypotheses stated in Chapter 1, several methodological approaches were required. Methodologies relevant to this thesis are described in this chapter.

3.1 Introduction

The central aims of this thesis are the geospatial and multilevel analysis of the distribution of cardiometabolic risk factors (CMRFs) in the Illawarra-Shoalhaven region of NSW, Australia. This chapter describes the materials and methods adopted to reach these aims, through the objectives listed in section 1.5 of the Chapter 1. That is to:

1. Systematically review the existing literature on the geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution.

2. Quantify small-area geographic variation in the distribution of cardiometabolic risk factors, within the Illawarra-Shoalhaven region of NSW Australia.

3. Quantify the proportion of small-area geographic variation in cardiometabolic risk factors explained by area-level socioeconomic disadvantage, within the Illawarra-Shoalhaven region of NSW Australia.

4. Quantify the proportion of small-area geographic variation in cardiometabolic risk factors explained by the differences in access to primary care, within the Illawarra-Shoalhaven region of NSW Australia.

A Directed Acyclic Graph (DAG) of the relationship between exposure, outcome and mediating or moderating variables of cardiometabolic risk factors are presented in Figure 3.1. The DAG presents a conceptual illustration of the single exposure and multiple outcomes being estimated, the covariates being adjusted for in the analyses and other relevant covariates or confounders not been adjusted for given the study scope and data access limitations. The DAG was informed based on the literature review, presented in chapter 1 and 2. The figure illustrates the variables been included and not included in the analyses, and provide a clear basis for the interpretation of results which are dealt in the inferential analyses and presented in the subsequent chapters.

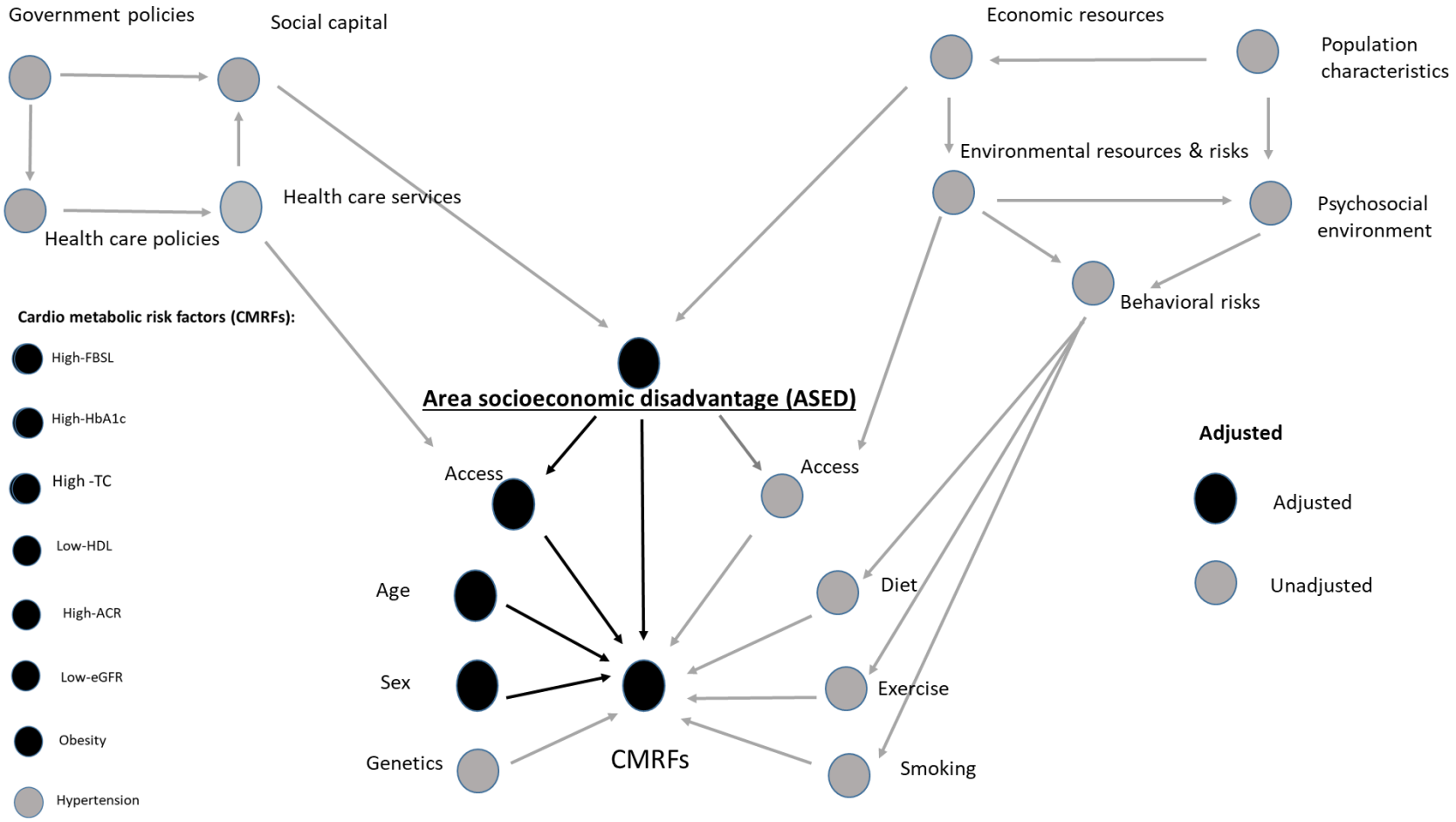

Figure 3.1: Directed acyclic graph illustrating the relationships among the exposures and mediating variables of cardiometabolic risk factors which are analysed and not analysed in this thesis

3.2 Study region

The research was conducted in the Illawarra-Shoalhaven region of NSW, Australia. This region extends south of the metropolitan boundaries of Sydney and stretches along the eastern seaboard of the NSW—bordered by the Pacific Ocean to the east and the Southern Tablelands to the west. The region consists of multiple cities, towns and rural areas; and encompasses the four local government areas of Wollongong, Shellharbour, Kiama and Shoalhaven. The study region covers a land area of 5615 km^2; and had an estimated residential population of 369, 469 people at the time of 2011 Australian Census of Population and Housing.[1]

The Illawarra-Shoalhaven region is the third-largest regional economy in NSW, Australia.[2] The main industries in the area have traditionally been farming, coal mining and steel making. [2] Manufacturing is the largest contributor to this regional economy.[2] Australia's largest steel-works, BlueScope, operates at Port Kembla in the Illawarra-Shoalhaven region. The region's natural beauty, diverse economy and relaxed coastal lifestyle make it an attractive tourism destination.[2]

The University of Wollongong is the main tertiary education centers, having multiple campuses across the region and attracts thousands of local and international students and staff every year. Health care is also a major employment provider in the region. Close proximity of the northern part of the region to Sydney also allows for industry, business and residents to connect with metropolitan economic and employment opportunities.[2] Table 3.1 shows the age and sociodemographic distribution of the local government areas of the study region.

Table 3.1a: Age distribution of the Local Government Areas of the Illawarra-Shoalhaven region, based on 2011 ABS census estimates.

Local Govt. Areas	Population		Age distribution of the population													
	Total	Density*	0-4 years		5-14 years		15-24 years		25 – 44 years		45 – 64 years		65 – 74 years		≥ 75 years	
			n	%	n	%	n	%	n	%	n	%	n	%	n	%
Wollongong	201215	769.6	12494	6.2	24040	11.9	29414	14.6	52990	26.3	49926	24.8	16516	8.2	15835	7.9
Kiama	20832	80.8	1076	5.2	2540	12.2	2565	12.3	4215	20.2	6253	30.0	2181	10.5	2002	9.6
Shellharbour	66054	448.2	4453	6.7	9415	14.3	9120	13.8	16912	25.6	16840	25.5	5276	8.0	4038	6.1
Shoalhaven	96043	89.1	5409	5.6	11460	11.9	10741	11.2	19278	20.1	27241	28.4	11864	12.4	10050	10.5
Total	**384144**	**67.6**	**23432**	**6.1**	**47455**	**12.4**	**51840**	**13.5**	**93395**	**24.3**	**100260**	**26.1**	**35837**	**9.3**	**31925**	**8.3**

* - persons/km^2

Table 3.1b: Sociodemographic distribution of the Local Government Areas of the Illawarra-Shoalhaven region, based on 2011 ABS census estimates.

Local Govt. Areas	Population		Sociodemographic distribution								
	Total	Density*	Indigenous		NESB		Labour force		Unemployment		SEIFA Score
			n	%	n	%	n	%	n	%	IRSD**
Wollongong	201215	769.6	4229	2.2	27478	14.3	9168	57.9	1093	6.9	980
Kiama	20832	80.8	285	1.4	906	4.5	1181	59.0	76	3.8	1055
Shellharbour	66054	448.2	1930	3.0	6029	9.5	2390	59.2	299	7.4	962
Shoalhaven	96043	89.1	4318	5.2	4484	4.8	5854	58.3	839	8.4	951
Total	**384144**	**67.6**	**10762**	**2.8**	**38897**	**10.1**	**18594**	**58.3**	**2307**	**7.2**	

* - persons/km^2; ** -weighted average scores of the census collection districts; IRSD - Index of relative socio economic disadvantage; NESB - Non-English speaking background; SEIFA - Socioeconomic indices for Areas.

Statistical Area level 1 (SA1) was used as the geographic unit of analysis in this research, which was the smallest geographic unit for the release of Australia census data in 2011.[3] The study region consisted of a total of 980 conterminous SA1s, with populations approximately between 200 and 800 people (400 averages). Figure 3.2 shows the map of the study area with SA1 units and the major landmarks of the region. As the SA1s are also based on their population size, the very small and crowded SA1s, similar to the areas shown the inset map, indicate densely populated SA1s.

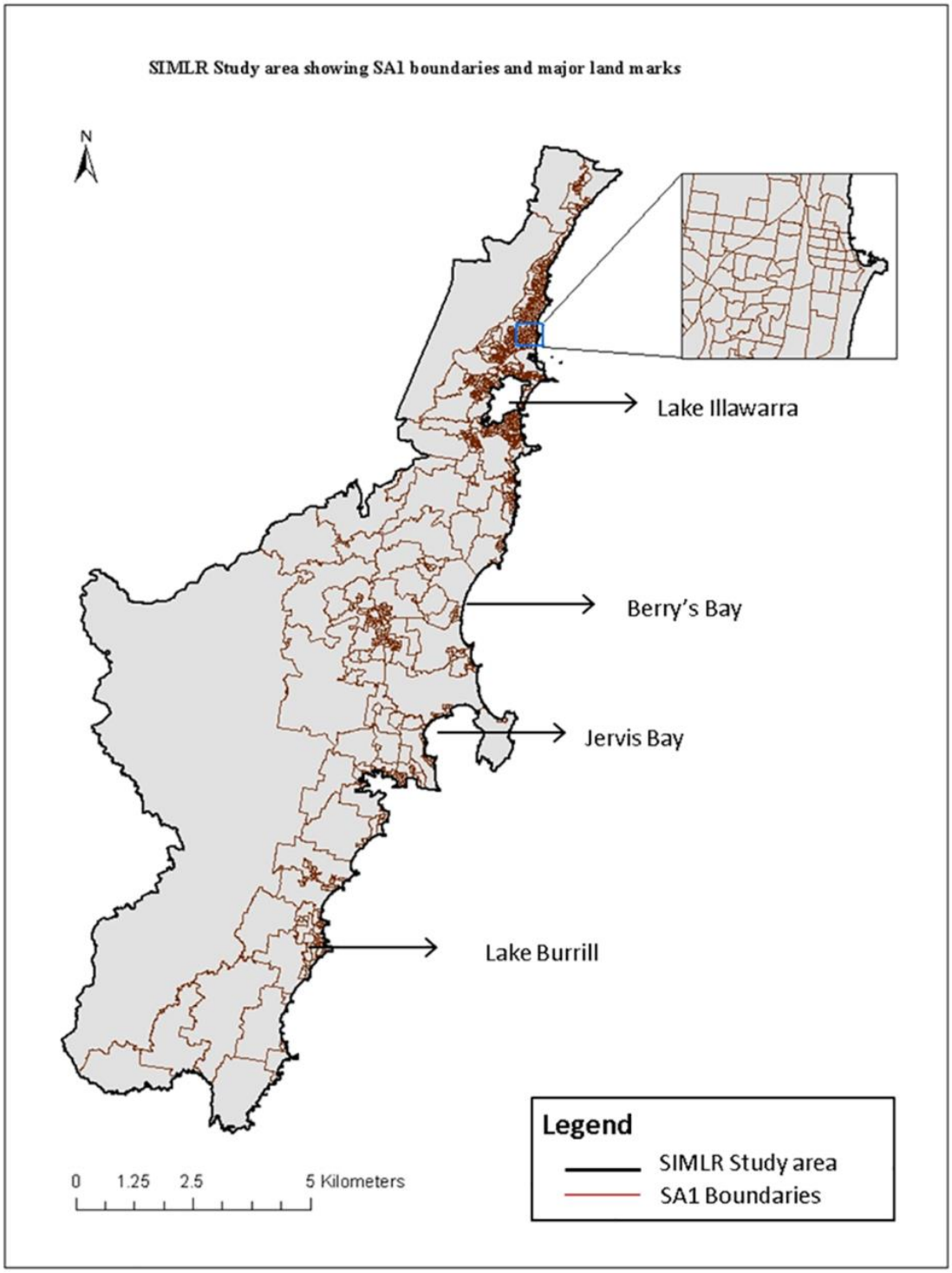

Figure 3.2: Map of the Illawarra-Shoalhaven region of NSW Australia

3.3 Data source and variables

Five main data sources were used in this thesis:

1. Literature search databases: MEDLINE (Ovid), PubMed, Scopus, and Web of Science;
2. Southern IML Research (SIMLR) data;
3. Australian census of population and housing 2011 data [4];
4. Australian Statistical Geography Standard (ASGS) reference data 2011[3]; and
5. Primary care provider data 2016.

The details of the study data sources, their location, access and the variables extracted for the analyses are detailed in the following.

3.3.1 Literature search databases

The systematic review of the existing literature on the geographic and area-level socioeconomic variation in CMRFs was undertaken using four databases. They are: MEDLINE (Ovid); PubMed; Scopus; and Web of science databases. These databases were chosen for the review, considering the breadth of fields they cover and the interdisciplinary nature of the review question. In addition, hand-searching of related articles was also undertaken from the retrieved articles. More details on these databases, the search strategy and the number of studies retrieved were presented in Chapter 2 of this thesis.

3.3.2 Southern IML Research (SIMLR) data

The de-identified, internally linked and geocoded routine pathology data of the Southern IML Pathology Pty Ltd, known as the Southern IML Research (SIMLR) data, was the primary data source for the study. Southern IML Pathology Pty Ltd. is the largest provider of pathology services in the Illawarra-Shoalhaven region; consisting of over 60 licensed collection centres across multiple geographic locations in the study region. The individual-level data in SIMLR database are geocoded to their corresponding SA1 areas, but not to their residential address, for privacy and confidentiality concerns. More detail on the procurement of SIMLR data is beyond the scope of the current thesis, but available in published literature elsewhere.[5, 6] Appendix IV details the data specification of the extracted variables and documentation of the data extraction program used.

The SIMLR data are stored and maintained in the Spatial Analysis Laboratories, within the School of Earth and Environmental Sciences of the University of Wollongong. Southern IML Pathology has granted the University of Wollongong a non-exclusive royalty free license to access and use these data for research purposes, provided the investigators pass through a two-tier approval system and adhere with the data access agreements.

Data access

Access to the SIMLR data is strictly controlled through a two-tier structure consisting of the Southern IML Research Cohort Management Committee (SIMLR-CMC) and approval through an appropriately constituted Human Research Ethics Committee (HREC). Data access were granted for a defined period as specified in the ethics approval, and renewed after the annual progress reporting to the HREC. The studies within this thesis were all approved by the University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical Human Research Ethics Committee (HREC protocol No: 2017/124). The initial approval was obtained in April 2017, and annually renewed thereafter until the completion of thesis related analyses in 2020.

Data extraction

Data were obtained from the SIMLR database for all non-pregnant persons aged ≥18 years with ≥1 cardiometabolic analyte result between 01 January 2012 and 31 December 2017. For each person, only the most recent cardiometabolic test result was extracted to maximise their temporal alignment with the residential location at the time of pathology testing.

Extracted Variables

1. *Unique person identifier*

The unique person identifier is a 7 digit numeric variable that is project specific and cannot be linked back to the individual or service provider.

2. *Sex*

The sex of individual clients as recorded at their most recent test result. Southern IML Pathology Pty Ltd have indicated that this value is considered the most accurate record of the gender status of their individual clients.

3. *Cardiometabolic analytes*

The extract includes the following cardiometabolic analyte variables:

a. Fasting blood sugar levels (FBSL);
b. Glycated haemoglobin (HbA1c);
c. High density lipoprotein (HDL);
d. Total (TC) cholesterol;
e. Albumin creatinine ratio (ACR);
f. estimated Glomerular Filtration Rate (eGFR); and
g. Body mass index (BMI).

For each of these analytes the following related variables were provided:

a. collection date;
b. age group at time of collection;
c. test value in standard units;
d. geocoding match status for residential address at time of collection;
e. geocoded Statistical Area 1 (SA1) of residential address at time of collection;
f. 2011 Index of Relative Socioeconomic Advantage/Disadvantage for geocoded SA1 at time of collection;
g. 2011 Index of Relative Socioeconomic Disadvantage for geocoded SA1 at time of collection;
h. 2011 Index of Economic Resources for geocoded SA1 at time of collection;
i. 2011 Index of Education and Occupation for geocoded SA1 at time of collection; and
j. Diabetes status

The SMILR study uses an algorithm to identify the diabetes status of the individuals included in the database. Diabetes is indicated when HbA1c is ≥ 6.5%; or FBSL is ≥ 7.0 mmol/l within +/- 24 months of an HbA1c < 6.5%. This algorithm is consistent with the diagnostic guidelines published by the Royal Australian College of General Practitioners (RACGP) and Diabetes Australia[7]; and methods from the National Health Survey of the Australian Bureau of Statistics (ABS)[8].

The diabetes case definition status was set to “Yes” for patients the first time the algorithm criterion was met; and then propagated throughout the data set for subsequent testing records in the SIMLR data. Thus, the

extracted study data set included both prevalent (diabetes detected before the data period, i.e. 01 January 2012) and incident (diabetes detected within the data period, i.e. between 01 January 2012 and 31 December 2017) cases of diabetes mellitus.

3.3.3 Australian Census of Population and Housing data, 2011

2011 Australian Census of Population and Housing data as released by the Australian Bureau of Statistics (ABS) were accessed for this study. The publicly available data cubes pertaining to the population data in the study region were accessed from the ABS website and matched to the SA1s within the study region.[4] This data is mainly used to derive the population denominators of the total adult (>=18 years) population of the SA1s in the study region. The total adult population were used at different phases of the study for various statistical analyses. For example, study 1 had used the total adult population of the SA1s for the Empirical Bayes smoothing of the SA1 level proportions of the higher risk CMRF test results; study 3 had used the total population size at SA1 level to calculate the provider to population ratios.

3.3.4 Australian Statistical Geography Standard (ASGS) reference data, 2011

The Australian Statistical Geography Standard (ASGS) reference data of the year 2011 had been accessed for the geographic location specific reference details of the study region.[9] This data were used for building the base map and SA1 boundaries of the study region. Also, the IRSD scores of the SA1s in the study region as of 2011 census had been joined based on the SA1 codes from ASGS reference data.

3.3.5 Primary care provider data: 2016

The primary care provider data were manually extracted in 2016, from the contact details of the services listed in publicly available directories, and included publicly available data sources such as: Yellow pages; White pages; online booking services; and location specific Google search results. The database was constructed to temporally align within the CMRF data period (2012 – 2017). The primary care locations within 30 km distance of the study boundaries were also included in the list to facilitate buffer distance analyses. Geocoding of the primary care locations was performed by converting the service providers' addresses into geographic coordinates (latitude and longitude) using Google Map services.[10] The number of general practitioners practicing in each service had been retrieved to calculate the total number of providers. The total number of providers in each SA1 was used to calculate the provider to population ratios of the SA1s in the study region, which was primarily used in study 4 of this thesis.

3.4 Research design

The study adopted a cross-sectional design. A cross sectional study is a type of observational study that analyses data at a defined period of time.[11] The design seeks to identify associations and within cohort relative risk estimates, though not causality.[12] Cross-sectional studies are frequently conducted to estimate the rates and associations of the outcome of interest for a given population, commonly for the purposes of public health planning.

3.5 Analyses

The analyses in this thesis were performed in four stages. Study 1 focused on systematic review and thematic synthesis of the literature, which provided strong evidence based foundation for the subsequent stages of the study. Study 2 concentrated on geographic and spatial analyses of the CMRF study data across the study region, which was descriptive and hypothesis generating for the next stages. Study 3 applied multilevel analyses of associations between area-level disadvantage and CMRFs within the study region after adjusting for individual

level factors. Study 4 extended the multilevel analyses by analysing the association of area-level primary care access with CMRF, after adjusting for individual level factors and area-level disadvantage.

3.5.1 Study 1 - Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature

Study 1 was a systematic literature review and thematic synthesis of the existing literature on the geographic and area-level socioeconomic variation of CMRFs.

Systematic literature review is a methodical approach for appraising literature from a selection of studies that focuses on a central research question.[13] The objective was to synthesise results and key findings across studies to summarise the best available evidence relevant to the central research question. The systematic review of the literature required having transparency, i.e. the methods used to conduct the review were reproducible and follow commonly accepted guidelines, such as Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and the International Prospective Register of Systematic Reviews (PROSPERO) guidelines.

As detailed in chapter 2 of this thesis, a systematic search for relevant previous studies had been undertaken within four databases (MEDLINE (Ovid), PubMed, Scopus and Web of Science). All search outcomes were limited to: human studies; adult population (>=18 years); and availability in English language. All retrieved studies were screened by two independent reviewers in three stages to reduce the risk of bias. In stage 1, articles from all databases were combined and screened to remove duplicates. Titles and abstracts of remaining articles were screened for eligibility, in stage 2. The final stage of study selection was done after full text reading of the remaining studies. Quality of the individual studies were assessed using the STROBE checklist for cohort, case-control and cross-sectional studies (www.strobe-statement.org). The second coder repeated all three stages in parallel and selected studies were matched at the conclusion of each stage and any differences were resolved by consensus and arbitration. Supervisors served as additional reviewers when required.

Data extraction and coding of the chosen studies were carried out using two pilot-tested templates for consistency. Template 1 focused on the geographic variation in CMRFs and was used to extract information on: author, year, nation, study design, sample size and characteristics, geographic unit of reporting, studied CMRFs and the study outcome. Template 2 addressed the association of area-level socioeconomic disadvantage (ASED) and cardiometabolic risk distribution and extracted additional data on the reported measures of ASED and associations. An additional template was used for thematic mapping of the data from the included studies for further qualitative syntheses. Study origin, representation, nature of problem, ecological context and evidence strength were the mapped themes. The finalised templates are reported as tables in the published article based on study 1 (Chapter 2: Tables 2.1, 2.2 and 2.3).

Summary measures used in this review were descriptive and based on the frequency of relevant studies to its relevant denominator. Endnote software was used to keep track of the bibliographic details of the studies throughout the selection and data extraction process.

3.5.2 Study 2 – Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256, 525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia

Study 2 aimed to quantify the small-area geographic variation in the distribution of cardiometabolic risk factors within the Illawarra-Shoalhaven region of NSW, Australia.

The study 2 variables included de-identified laboratory data on eight CMRFs including fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI); and diabetes mellitus (DM) status of the adult residents in the study region. The CMRF test results were dichotomised into 'higher risk' and 'lower risk' values based on existing risk definitions. The risk definition values (as reported in chapter 4, Table 4.1) are as follows:

Table 3.2: Cardiometabolic risk value definitions.

'Higher risk' CMRFs	Value definition	Adopted from
High FBSL	FBSL ≥7.0 mmol/l	RACGP guidelines.[7]
High HbA1c	HbA1c > 7.5%	RACGP guidelines.[7]
High TC	TC ≥ 5.5 mmol/l	Australian Health Survey.[8]
Low HDL	HDL < 1 mmol/l	National heart foundation of Australia.[14]
High ACR	ACR ≥ 30 mcg/L to mg/l	Kidney Health Australia.[15]
Low eGFR	eGFR < 60 mL/min/1.73m^2	Kidney Health Australia.[15]
High BMI	BMI ≥ 30(Obese)	World Health Organization (WHO).[16]
DM Status	+ve DM test algorithm	RACGP guidelines[7], and Australian Health Survey[8].

Within-cohort proportion of 'higher risk' CMRF findings are calculated using the total number of tests within each SA1 as the denominator. The exception were DM cases, which uses an algorithm within the SIMLR database to identify DM cases. The DM cases identified prior to the study data period (year 2012 to 2017), were forward propagated into the study data extract labelled as 'existing' cases. Thus the study data include both the 'existing' and 'new' (identified within the study data period) cases of DM. This is likely to include most DM cases from the study area, considering the duration and population coverage of the pathology network in the study region from which the study data is sourced. Therefore SA1 adult populations aged 18 years and over were used as the denominators (accessed from ABS census 2011 data) of SA1's DM cases. In the year 2011, the NSW Health Statistics reported a prevalence of 9.5% DM (95% CI: 7.6 – 11.4) in 16 years or older from the Illawarra-Shoalhaven Local health district[17], which stands close to the study 2 identified 9.2% DM in the 18 years or older study cohort[1].

Australian Census Statistical Area Level 1 (SA1) was used as the geographic units of analysis. An Empirical Bayes (EB) approach was used to smooth all the CMRFs' raw rates to minimise extreme values arising from small sample sizes in SA1s. The EBest function in R was used to calculate the raw and Empirical Bayes smoothed proportions, using a 'binomial' family option as the higher risk events in the data were not rare.[18] The EB smoothed rates were then imported into GIS software for mapping and spatial statistical analyses. Choropleth maps demonstrating the distribution of CMRF rates at SA1 level were produced.

Spatial clustering of CMRFs was assessed using Global Moran's I test and Local Indicators of Spatial Autocorrelation (LISA). Global Moran's I test was used to identify spatial autocorrelation of CMRFs at a 0.05 level of significance. Global Moran's I tests if the geographic distribution of rates is clustered, dispersed or random.[19] The global Moran's I also indicate the general strength of spatial autocorrelation in the study area, which theoretically ranges between -1 to +1. Values of I significantly above -1/ (N-1) indicate positive spatial

autocorrelation, where N is the number of spatial units indexed.[20] When significant spatial autocorrelation was detected, Local Indicator of Spatial Autocorrelation (LISA) spatial statistics were used to identify any clustering of CMRFs.[[21] LISA was used to indicate spatial clustering of High-High (HH) or Low-Low (LL) CMRFs rates at SA1-level within the study region. HH refers to a statistically significant (0.05 level) cluster of high values (HH) and LL refers to a statistical significant cluster of low values (LL). False Discovery Rate (FDR) corrections were applied to LISA tests to correct p-values for multiple testing.

The main strength of this geospatial analytical approach was its ability to visually present area-level data, potentially rendering the data more accessible than traditional table and graph methods.[22] Usage of the smallest available area-level unit in the maps further enhanced the area-level specificity of the maps.[23] However, this approach was not without limitations. Spatial analyses are not free of ecological problems such as ecological fallacies and Modifiable Areal Unit Problem or (MAUP).[35] As ecological analyses, they focus on geographical variation only.[23] As these approaches were exclusively based on the analysis of area-level variance, a question considered was when to consider this area variance as large or important.[24] As statistical significance alone was not considered as a satisfying criterion, multilevel models of the individual CMRFs were planned for subsequent stages of the study.[25]

3.5.3 Study 3 – Geographic variation in the distribution of cardiometabolic risk factors explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia

Study 3 aimed to quantify the proportion of small-area geographic variation in CMRFs explained by area-level socioeconomic disadvantage in the study region.

The study 3 variables included de-identified laboratory data on seven CMRFs including fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); and body mass index (BMI) of the adult residents in the study region. The diabetes mellitus (DM) status of the individuals were not included in study 3 and 4 analyses as the covariate information of their denominator population were not available within study data. This is because the DM status in the study data set alone is synthesized from other variables (hba1c and FBSL results) in the SIMLR source system and included both ‘existing’ (prevalent) and ‘new’ (incident) cases, which is likely to include all the cases from the study region and hence the actual population of the study region was determined to be their appropriate denominator. However, it wasn’t possible to link the covariate data of the denominators to the DM cases in the study data. Hence, DM status codes were excluded from study 3 analyses onwards.

To achieve this aim, area-level analyses of the data were undertaken using multilevel logistic regression models. Multilevel models or hierarchical models are generally used for data nested in multiple levels.[25] While the units of analyses were individuals at level 1, contextual or area-level variables were analysed at level 2. Multilevel logistic regression analysis is proposed for studying data with a binary response and multilevel nesting.[21] Such analyses identify group-level associations between specific contextual level variables and individual health adjusting for area-level clustering.[26] They also have the ability to partition within-area effects from the between-area effects.[27] The 2011 ABS census based Index of Relative Socioeconomic Disadvantage (IRSD) expressed in quintiles was used as the study variables. The IRSD summarises a range of measures of relative socioeconomic disadvantage of people and households within SA1s and includes: level of income; education; employment; family structure; disability; housing; transportation; and internet connection.[26] Study

3 Analyses used IRSD reported as quintiles at SA1 level; the lowest quintile (Q1) indicating the most disadvantaged SA1s and the highest quintile (Q5) the least disadvantaged SA1s.[25]

Multilevel logistic regression models of the CMRFs data adjusted for individual-level age group and sex.

The general equations of the fully adjusted model are:

$$y_{ij} \sim Binomial(1, \pi_{ij}) \quad (3.1)$$

$$logit(\pi_{ij}) = \beta_0 + \beta_1 age_{ij} + \beta_2 sex_{ij} + \beta_3 IRSD_{ij} + u_j \quad (3.2)$$

$$u_j \sim N(0, \tau^2{}_u) \quad (3.3)$$

where: y_{ij} denote the binary response of CMRF test outcome (as 'higher risk' or 'lower risk', based on the adopted definitions) for individual i in the area (SA1) j; π_{ij} denotes the probability that individual i in area (SA1) j has a 'higher risk' CMRF test outcome given their individual-level age_{ij} and sex_{ij}; and their area-level IRSD index $IRSD_{ij}$. $\beta_1, \beta_2, \beta_3$ are the regression coefficients which measure the associations between the log-odds of the CMRF outcome and each covariate all else equal, and when exponentiated these are translated to ORs.[25] u_j is the random effect for the area (SA1) j and $\tau^2{}_u$ is the area level variance, which has to be estimated.

For each of the seven CMRFs analysed in this study, a hierarchy of four multilevel models at SA1 level were fit that included fixed effects for age, sex and IRSD and random effect (intercept) for SA1. Thus, a total of 28 logistic regression models were built to quantify the proportion of small-area geographic variation in CMRFs explained by the area level socioeconomic status in the study region. Model 1s (M1s) were null models; Model 2s (M2s) included the individual-level covariates (age and sex); Model 3s (M3s) included the area-level study variable (IRSD) only; and Model 4s (M4s) included both individual- and area-level variables (age sex and IRSD). The estimated regression coefficients of the derived models were exponentiated to calculate odds ratios (ORs) of the variables. The goodness of fit of the models were identified using Likelihood Ratio Tests at $p < 0.05$ level of significance. Figure 3.3 provides a schematic representation of the models pattern used in study 3.

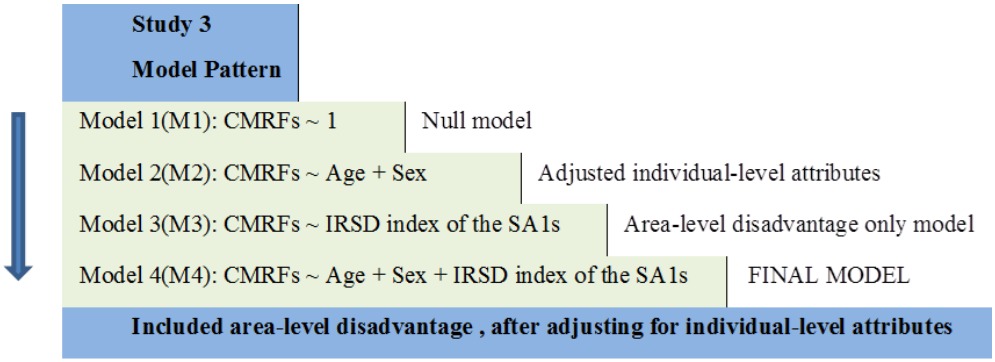

Figure 3.3: Schematic representation of the modelling patterns adopted in study 3

Model comparison

The Akaike Information Criterion (AIC) was used to evaluate model fit. The derived models were compared for: area level variance (τ^2) at SA1 level; proportional change in variance (PCV); Intra-cluster Correlation Coefficients (ICC); and Median Odds Ratios (MORs).

The area-level variances (τ^2 s) were initially identified from each model. PCVs were calculated for models M2 to M4 relative to M1. The ICCs of the fitted models were calculated using the latent variable approach.[28] This

approach assumes that a latent continuous outcome underlies the observed dichotomous outcomes and it is this latent outcome for which the ICC is calculated and interpreted. The ICC measured the expected correlation in CMRF outcomes between two individuals from the same SA1. The higher the ICC, the more relevant area-level context is for understanding individual latent outcome variation.[25] The MOR is calculated as an alternative way of interpreting the magnitude of area-level variance. The MOR translated the area-level variance which were estimated on the log-odds scale to the commonly used OR scale. The MOR result value is interpreted as the median increased odds of identifying the outcome if an individual move to another SA1 with higher risk. Thus, higher the MOR the greater the general area-level effect.[25] The unique contribution of the area-level study variable (IRSD) to the area-level variance of 'higher risk' CMRFs were assessed through the PCVs between M2s and M4s.

3.5.4 Study 4 - Does access to primary care reduce the geographic variation of cardiometabolic risk factor distribution? A multilevel analysis of the adult residents in the Illawarra-Shoalhaven region of Australia

Study 4 aimed to investigate area-level associations between access to primary care and various CMRFs, after adjusting for area-level disadvantage. To achieve this, study 4 analyses were done in two steps. Step 1 aimed to quantify the area-level access to primary care services in the study region, and step 2 analysed the area-level associations between access to primary care and various CMRFs, after adjusting for area-level disadvantage. Therefore, the analytical approaches adopted in this study are detailed under two sections as study 4a and study 4b as follows.

Study 4a – Deriving the primary care access index of the study region

Study 4a aimed to quantify the area-level primary care access of the study region. To achieve this aim, the two-step floating catchment area (2SFCA) method was used to derive an access index for each SA1 in the study region.

The 2SFCA method was created by Luo and Wang in 2003 to measure geographic accessibility of health care services.[29] Essentially, the 2SFCA method consists of two steps underpinned by gravity models.[29] The first step analysed the availability of primary care providers within 1 km, 16 km and 30 km distance of the geographic centroid of SA1s in the study region, to quantify the supply of services to the SA1. The 30 km buffer distance was observed to provide a better coverage of the population, and thus adopted for the further analyses. The total number of general practitioners (GPs) in the service provider locations within each SA1s had been the numerators for the provider to population ratio calculations. The second step considered the total population within 30 km of a primary care provider to determine the demand for services. Therefore, this gravity models considered the interaction between the supply and the demand of primary care in the study area.

Thus, step 1:

$$Rj = Sj \Big/ \sum_i Pi \quad , \qquad (3.4)$$

where *Sj* is the number of general practitioners at location *j*, pi is the number of adult residents in the SA1s (Those SA1's geographic centroids are located within the spatial buffer distance of the primary care locations) and Rj is the population-to-provider ratio for service *j*.[30]

In step 2, a population-to-provider ratio (access score) is computed for each geographic centroid of the SA1s by aggregating all primary care service population-to provider ratios of the primary care services that are located within the same spatial buffer distance.[30]

Thus, step 2:

$$Ai = \sum_j R_j , \qquad (3.5)$$

where *Ai* is the access index for population location *i*.

The resulting access indices had been retained as a continuous variable for the analyses. A higher score indicated better geographic access of the populations of SA1s to primary care services, which in practice is defined as an improved supply of primary care service locations, in balance with the population size of the small areas within the study region.

Study 4b – Multilevel analysis of the association between CMRFs and area-level access to primary care

Study 4b aimed to investigate area-level associations between access to primary care and various CMRFs, after adjusting for area-level disadvantage. To achieve this, multilevel logistic regression models were fit after adjusting for individual and area-level variables. The access index previously derived study was used as the study variable and the IRSD scores of the SA1s were used as measures of area-level disadvantage.

Multilevel logistic regression models were fit for the CMRF test data of individuals (Level 1) nested within SA1s (Level 2). For each of the seven CMRFs analysed in this study, a hierarchy of five multilevel models were fit. The models included fixed effects for sex, age, IRSD score and access index; and random effect (intercept) for SA1s. Thus, a total of 35 models were fit to achieve the study objectives. Model 1s (M1s) were null models; Model 2s (M2s) included the area-level study variable (access index) only; Model 3s (M3s) included individual-level covariates (age and sex) only; Model 4s (M4s) included both individual (age and sex) and area-level (IRSD score of the SA1s) covariates; and the final models (M5s) included the primary care access index of SA1s, adjusting for the individual (age and sex) and area- level (IRSD scores of SA1s) covariates. The goodness of fit of the models was estimated through Likelihood Ratio Tests at $p < 0.05$ level of significance. Figure 3.4 provides a schematic representation of the models pattern used in study 4.

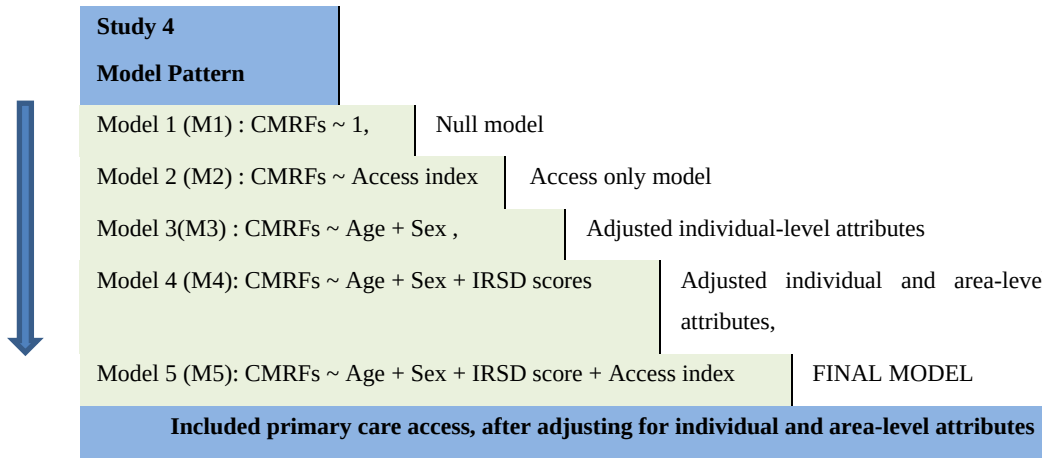

Figure 3.4: Schematic representation of the model pattern adopted in study 4

Model comparison

Model fit was compared using the Akaike Information Criterion (AIC). The models were also evaluated for: area-level variance (τ^2); proportional change in variance (PCV) in comparison with the null model; Intra-cluster Correlation Coefficient (ICC) of the model; and the Median Odds Ratios (MORs). The ICC and MOR of the models were used to index the between-area variability. A latent variable approach was used to derive the ICC of models.[28] The MOR translates the area-level variance into an easily interpretable OR and is assumed to be statistically independent of the test specific distribution of the CMRFs.[31] The unique contribution of the primary care access of the SA1s to the area-level variance of CMRF was estimated through the reduction in PCV between M4 and M5.

3.6 Analytical software

All spatial analyses and mapping were performed using ArcGIS version 10.4.1(ESRI Inc. Redlands, CA, USA).[32] Statistical analyses and multilevel models were performed using R version 3.4.4. (R Foundation for Statistical Computing, Vienna, Austria).[33] R is a language and environment for statistical computing and graphics.[34] R has multiple packages to meet a range of multilevel analytical requirements, for example: Linear mixed-effects models using S4 classes (lme4)[35]; Linear Mixed-Effects Models with Censored Responses (lmec)[36]; Multilevel Functions (multilevel)[37]; hierarchical generalised linear models (hglm)[38]; Generalised Linear Mixed Models and Spatial Models with BUGS (glmmBUGS)[39]; MCMC Generalised Linear Mixed Models (MCMCglmm)[40]; Linear and Nonlinear Mixed Effects Models (nlme)[41]; and Mixed models for discrete data in R (glmmADMB)[42]. [43, 44] The multi-level models were fit using the *glmer* function in the *lme4* package in R; [35] and the likelihood ratio tests were calculated using the *lrtest* function in the *lmtest* package in R.[45]

3.7 Ethics committee approval

The study was approved by the University of Wollongong (UOW) and Illawarra and Shoalhaven Local Health District (ISLHD) Human Research Ethics Committee (HREC 2017/124). The ethics committee determined informed consent was not required because individual-level data were non-identifiable.

3.8 Time line

The study had been initiated in Jan 2016 and extended over 4 years until Feb 2020. Figure 3.5 provides a schematic representation of the approximate time line required to complete this thesis.

	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**
2016	Registration of PhD		Continue literature review				Proposal writing			Ethics application	Proposal review by HREC*	Annual holiday
2017	**Study 1:** Systematic review of the literature (continual and ongoing…)						**Study 2:** Geographic and spatial analyses of CMRFs					Annual holiday
2018	**Study 3:** Multilevel analyses of area-level disadvantage & CMRFs						**Study 4a**: Geospatial analysis of the primary care access of the study region					Annual holiday
2019	**Study 4b:** Multilevel analyses of area-level access & CMRFs, after adjusting for area-level disadvantage									Chapter writing and submission of the thesis for examination.		
2020	**Thesis examination**		**Thesis revision and final examination**									

*HDRC – Higher degree research committee, UOW

Figure 3.5 Schematic representation of the thesis time line

References

1. Toms, R., et al., *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256.* 2019.
2. The NSW Government. *Regional NSW: Illawarra-Shoalhaven*. 2016; Available from: https://www.nsw.gov.au/our-regions/illawarra-shoalhaven#:~:text=The%20Illawarra%2DShoalhaven%20is%20the,are%20the%20region's%20largest%20employers.
3. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas:STATISTICAL AREA LEVEL 1 (SA1)*. 2016 [cited 2018 Sept 5]; Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by Subject/1270.0.55.001~July 2016~Main Features~Statistical Area Level 1 (SA1)~10013
4. Australian Bureau of Statistics. *2011 Census data* [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280.
5. Bonney, A., et al., *Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users–a cross-sectional study.* PLoS One, 2015. **10**(8): p. e0137261.
6. Mayne, D.J., et al., *A novel population health data source to inform local planning: the SIMLR Study.* 2014.
7. The Royal Australian College of General Practitioners & Diabetes Australia. General Practice Management of Type 2 Diabetes 2016-2018. The Royal Australian College of General Practitioners (2016). doi:10.1007/s00125-010-2011-6.
8. Australian Bureau of Statistics. *Australian Health Survey: Biomedical Results for Chronic Diseases,2011-12* . 2013; Available from: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0052011-12.
9. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Census Dictionary, 2011*. 2011; Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2901.0Chapter23102011.
10. Australian Bureau of Statistics. *Main Features - IRSD.*; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011.
11. *Guest G, Namey EE. Public health research methods. SAGE Publications Sage UK: London, England; 2014. 832 p.* .
12. Levin, K.A., *Study design III: Cross-sectional studies.* Evidence-based dentistry, 2006. **7**(1): p. 24-25.
13. Whiting, P., et al., *ROBIS: a new tool to assess risk of bias in systematic reviews was developed.* Recenti progressi in medicina, 2018. **109**(9): p. 421-431.
14. *National heart foundation of Australia. Lipid management profile for health professionals. Available at: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management.*
15. *National Kidney foundation(USA). Albumin creatinine Ratio (ACR). (2018). Available at: https://www.kidney.org/kidneydisease/siemens_hcp_acr.*
16. *WHO. Obesity : Preventing and managing the global epidemic. World Health Organization: Technical Report Series. WHO Technical Report Series, no. 894. (2000). doi:ISBN 92 4 120894 5.*
17. NSW Population Health Survey (SAPHaRI). *Diabetes prevalence in adults.* 2002 - 2019 [cited 2021 Jan 29]; Available from: http://www.healthstats.nsw.gov.au/Indicator/dia_prev_age/dia_prev_lhn_trend.
18. Marshall, R.J., *Mapping disease and mortality rates using empirical Bayes estimators.* Journal of the Royal Statistical Society: Series C (Applied Statistics), 1991. **40**(2): p. 283-294.

19. Li, H., C.A. Calder, and N. Cressie, *Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model.* Geographical Analysis, 2007. **39**(4): p. 357-375.
20. Moran, P.A., *Notes on continuous stochastic phenomena.* Biometrika, 1950. **37**(1/2): p. 17-23.
21. Anselin, L., *Local indicators of spatial association—LISA.* Geographical analysis, 1995. **27**(2): p. 93-115.
22. Jacquez, G.M., *Spatial analysis in epidemiology: Nascent science or a failure of GIS?* Journal of Geographical Systems, 2000. **2**(1): p. 91-97.
23. Rytkönen, M.J., *Not all maps are equal: GIS and spatial analysis in epidemiology.* International journal of circumpolar health, 2004. **63**(1): p. 9-24.
24. Merlo, J., et al., *The tyranny of the averages and the indiscriminate use of risk factors in public health: The case of coronary heart disease.* SSM-population health, 2017. **3**: p. 684-698.
25. Merlo, J., et al., *An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health.* PloS one, 2016. **11**(4): p. e0153778.
26. Sommet, N. and D. Morselli, *Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS.* International Review of Social Psychology, 2017. **30**: p. 203-218.
27. Wong, G.Y. and W.M. Mason, *The hierarchical logistic regression model for multilevel analysis.* Journal of the American Statistical Association, 1985. **80**(391): p. 513-524.
28. Goldstein, H., W. Browne, and J. Rasbash, *Partitioning variation in multilevel models.* Understanding statistics: statistical issues in psychology, education, and the social sciences, 2002. **1**(4): p. 223-231.
29. Luo, W. and F. Wang, *Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region.* Environment and Planning B: Planning and Design, 2003. **30**(6): p. 865-884.
30. McGrail, M.R., *Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements.* International journal of health geographics, 2012. **11**(1): p. 50.
31. Flowerdew, R., D.J. Manley, and C.E. Sabel, *Neighbourhood effects on health: does it matter where you draw the boundaries?* Social science & medicine, 2008. **66**(6): p. 1241-1255.
32. Environmental Systems Research Institute (ESRI). *ArcGIS 10.4.1, ESRI Inc. Redlands, CA, USA.*; Available from: https://www.esri.com/
33. R Core Team. *R: A language and environment for statistical computing.R Foundation for Statistical Computing,Vienna, Austria*. 2018; Available from: https://www.R-project.org/.
34. The R Foundation. *What is R? Introduction to R*. [cited 2020 Oct 2020]; Available from: https://www.r-project.org/about.html.
35. *Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi: 10.18637/jss.v067.i01.*
36. *Florin Vaida and Lin Liu (2012). lmec: Linear Mixed-Effects Models with Censored Responses. R package version 1.0. https://CRAN.R-project.org/package=lmec*
37. *Paul Bliese (2016). multilevel: Multilevel Functions. R package version 2.6. https://CRAN.R-project.org/package=multilevel.*
38. *Lars Ronnegard, Xia Shen and Moudud Alam (2010) hglm: A Package for Fitting Hierarchical Generalized Linear Models. The R Journal, 2(2): 20-28. URL https://journal.r-project.org/archive/2010-2/RJournal_2010-2_Roennegaard~et~al.pdf.*

39. *Brown PE and Zhou L (2016). glmmBUGS: Generalised Linear Mixed Models and Spatial Models with WinBUGS, Jags, and OpenBUGS. R package version 2.4.0, https://CRAN.R-project.org/package=glmmBUGS.*
40. *Jarrod D Hadfield (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. Journal of Statistical Software, 33(2), 1-22. URL http://www.jstatsoft.org/v33/i02/.*
41. *Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2020). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-149, https://CRAN.R-project.org/package=nlme.*
42. *Fournier DA, Skaug HJ, Ancheta J, Ianelli J, Magnusson A, Maunder M, Nielsen A and Sibert J (2012). "AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models." _Optim. Methods Softw.(27), pp. 233-249.*
43. CRAN-R PROJECT. *Available CRAN Packages By Name*. 2020 [cited 2020 Sept 5]; Available from: https://cran.r-project.org/web/packages/available_packages_by_name.html.
44. Harrison, X.A., et al., *A brief introduction to mixed effects modelling and multi-model inference in ecology.* PeerJ, 2018. **6**: p. e4794.
45. *Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/.*

Chapter 4

Geographic Variation in Cardiometabolic Risk Distribution: A Cross-sectional Study of 256, 525 Adult Residents in the Illawarra-Shoalhaven region of the NSW, Australia

Chapter 4: Geographic Variation in Cardiometabolic Risk Distribution: A Cross-sectional Study of 256, 525 Adult Residents in the Illawarra-Shoalhaven Region of the NSW, Australia

4.1 Publication profile

This chapter presents the substantive content of research published in: ***PLOS ONE***, on 1st October 2019. Parts of the methods and findings from this study were also presented in two abstract reviewed conferences, WONCA World Rural Health Conference (New Delhi, 2018) and Public Health Association of Australia (PHAA) Public Health Prevention Conference (Sydney, 2018).

Journal article

Toms R, Mayne DJ, Feng X, Bonney A. *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256, 525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia.* **PLOS ONE,** 2019. 14(10): e0223179.

Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223179

DOI: https://doi.org/10.1371/journal.pone.0223179

Published version of the article

The published version of the article is appended within the **'Supplementary Materials'** section of the thesis as **Supplementary Material 2**.

Peer reviewed conference abstracts

1. **World Rural Health Conference -2018 (Oral presentation)**

Toms, R., Bonney, A., Mayne, D. J. & Feng, X. *Geographic variance in distribution of cardiometabolic risk factors in the Illawarra-Shoalhaven region of the NSW, Australia*. **15th WONCA World Rural Health Conference**, 2018. New Delhi, India. (pp.1-11).

Available from: https://scholars.uow.edu.au/display/publication130220

2. **Public Health Association of Australia (PHAA) – 2018 (Poster presentation)**

Toms, R., Bonney, A., Mayne, D. J., &Feng, X. *Geographic distribution of cardiometabolic risk factors: a small area level approach. (Poster presentation) Public Health Prevention Conference – 2018,* **Public Health Association of Australia (PHAA)**, 4-5 May 2018. Sydney, Australia. Available from: https://scholars.uow.edu.au/display/publication130221

4.2 Abstract

Introduction: Metabolic risk factors for cardiovascular disease (CVD) warrant significant public health concern globally. This study aims to utilise the regional database of a major laboratory network to describe the geographic distribution pattern of eight different cardiometabolic risk factors (CMRFs), which in turn can potentially generate hypothesis for future research into locality specific preventive approaches.

Method: A cross-sectional design utilising de-identified laboratory data on eight CMRFs including fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI); and diabetes mellitus (DM) status was used to undertake descriptive and spatial analyses. CMRFs test results were dichotomised into 'higher risk' and 'lower risk' values based on existing risk definitions. Australian Census Statistical Area Level 1 (SA1) were used as the geographic units of analysis and an Empirical Bayes (EB) approach was used to smooth rates at SA1 level. Choropleth maps demonstrating the distribution of CMRFs rates at SA1 level were produced. Spatial clustering of CMRFs was assessed using Global Moran's I test and Local Indicators of Spatial Autocorrelation (LISA).

Results: A total of 1, 132, 016 test data derived from 256, 525 individuals revealed significant geographic variation in the distribution of 'higher risk' CMRFs findings. The populated eastern seaboard of the study region demonstrated the highest rates of CMRFs. Global Moran's I values were significant and positive at SA1 level for all CMRFs. The highest spatial autocorrelation strength was found among obesity rates (0.328) and the lowest for albuminuria (0.028). LISA tests identified significant High-High (HH) and Low-Low (LL) spatial clusters of CMRFs, with LL predominantly in the less populated northern, central and southern regions of the study area.

Conclusion: The study describes a range of CMRFs with different distributions in the study region. The results allow generation of hypotheses to test in future research concerning location specific population health approaches.

Keywords: *cardiometabolic risk factors, geographic variance, spatial autocorrelation, spatial clustering*

4.3 Introduction

Uncontrolled cardiometabolic risk factors (CMRFs) such as hyperglycaemia, dyslipidaemia, albuminuria, inadequate glomerular filtration, overweight and/or obesity and diabetes can predispose and heighten the risk for cardiovascular disease (CVD).[1-6] Cardiovascular diseases are the leading cause of death worldwide and the highest absorber of health care expenditure in many developed nations, including Australia.[7-9]

In Australia, CVDs remain the single leading cause of death; the largest health problem; and a major economic burden.[10, 11] Nine in 10 adult Australians have at least one CVD risk factor and one in four have three or more risk factors.[10] CVD kills one Australian every 12 minutes and one in six Australians (3.7 million people) are thought to be at risk.[12] In addition, the prevalence of CVD is projected to steeply increase in the coming decades.[10] A deceleration in the rapid growth of this major health care issue is possible only through the prevention and control of CMRFs. The role of CMRFs in the population, over and above individual level factors such as age, are being questioned in regard to discriminatory accuracy for development of CVD,[13] however identification of one or more CMRFs in a person at any age can initiate preventive lifestyle changes which may have significant benefits.[14-18] Similarly, identification of areas with higher rates of CMRFs can potentially trigger further area-level analyses investigating the potential for targeted health service commissioning.[19-21]

Advances in Geographic Information System (GIS) over the last quarter of a century have provided various tools to integrate epidemiological and geographical data.[22-24] Geocoding of risk parameters became feasible with such tools for its area-level analyses, which has facilitated area-level mapping of risk parameters, which has the potential to generate hypothesis for regional health care research.[23] Thus integrating risk parameters through GIS has the potential to facilitate area-level health research,[25-28] however, not without potential pitfalls.[29-31] A limitation of GIS-based mapping is that its outputs may be misleading, especially if maps are not smoothed using appropriate spatial or multilevel analyses.[32-34] However, it is well recognised in the literature that area level community interventions based on GIS approaches have been successful in a number of countries. [19-21, 35, 36]

There has been a significant increase in the number of epidemiological studies using spatial analytical methods in the last decade, including international studies reporting significant geographic variation in CMRFs at different spatial scales of measurements.[37-45] Hyperglycaemia was the most commonly reported CMRF displaying variation, followed by dyslipidaemia, overweight and/or obesity and inadequate glomerular filtration.[38] Multiple risk factors were rarely analysed in these studies, though most CMRFs are interrelated and often coexist.[46] In this study, we aim to demonstrate the feasibility of utilising laboratory based routine test data to generate basic distribution maps of eight different CMRFs in regional New South Wales (NSW), Australia. The research questions we address are: (1) what is the geographic distribution pattern of CMRFs in the study area; and (2) is there any significant spatial clustering of CMRFs rates? The research sought to identify area-level patterns in the distribution of CMRFs that could be used to generate hypotheses for future research with the goal of improving health service commissioning in the study region.

4.4 Methods

The study adopted a cross-sectional design and was approved by the University of Wollongong (UOW) and Illawarra and Shoalhaven Local Health District (ISLHD) Human Research Ethics Committee (HREC 2017/124).

Setting

The study was undertaken in the Illawarra-Shoalhaven region (ISR) of the NSW, Australia. The ISR region stretches from the immediate south of the metropolitan boarders of Sydney and extends along the south-eastern coastal belt of NSW - bordering Pacific Ocean in the east and the coastal escarpment of the Southern Tablelands in the West. This region encompasses multiple cities, towns and rural areas and includes the four local government areas of Wollongong, Shellharbour, Kiama and Shoalhaven. Overall, the ISR covers a land area of 5615 square kilometres and had an estimated residential population of 369, 469 persons at the 2011 Australian Census of Population and Housing, of which 285, 385 (77.24%) were adults (>=18 years).[47] De-identified data for this study were obtained from the Southern IML Research (SIMLR) Study, a large-scale community-derived cohort of internally-linked and geographically referenced pathology data collected in routine practice by the largest pathology provider servicing the study area. More details on this data source, its access and maintenance are published elsewhere.[48]

Statistical Area level 1 (SA1) was used as the geographic unit of analysis in this study, which was the smallest geographic unit for the release of Census data in 2011.[49] SA1s generally have a population of 200 to 800 persons (400 averages) and the ISR includes a total of 980 conterminous SA1s. Figure 4.1 shows the study area with SA1 units and the major landmarks of the region. Very small and crowded SA1s similar to the areas shown the inset map tend to be more densely populated. Figure 4.2 illustrates distribution of the adult population per SA1 in the study region.

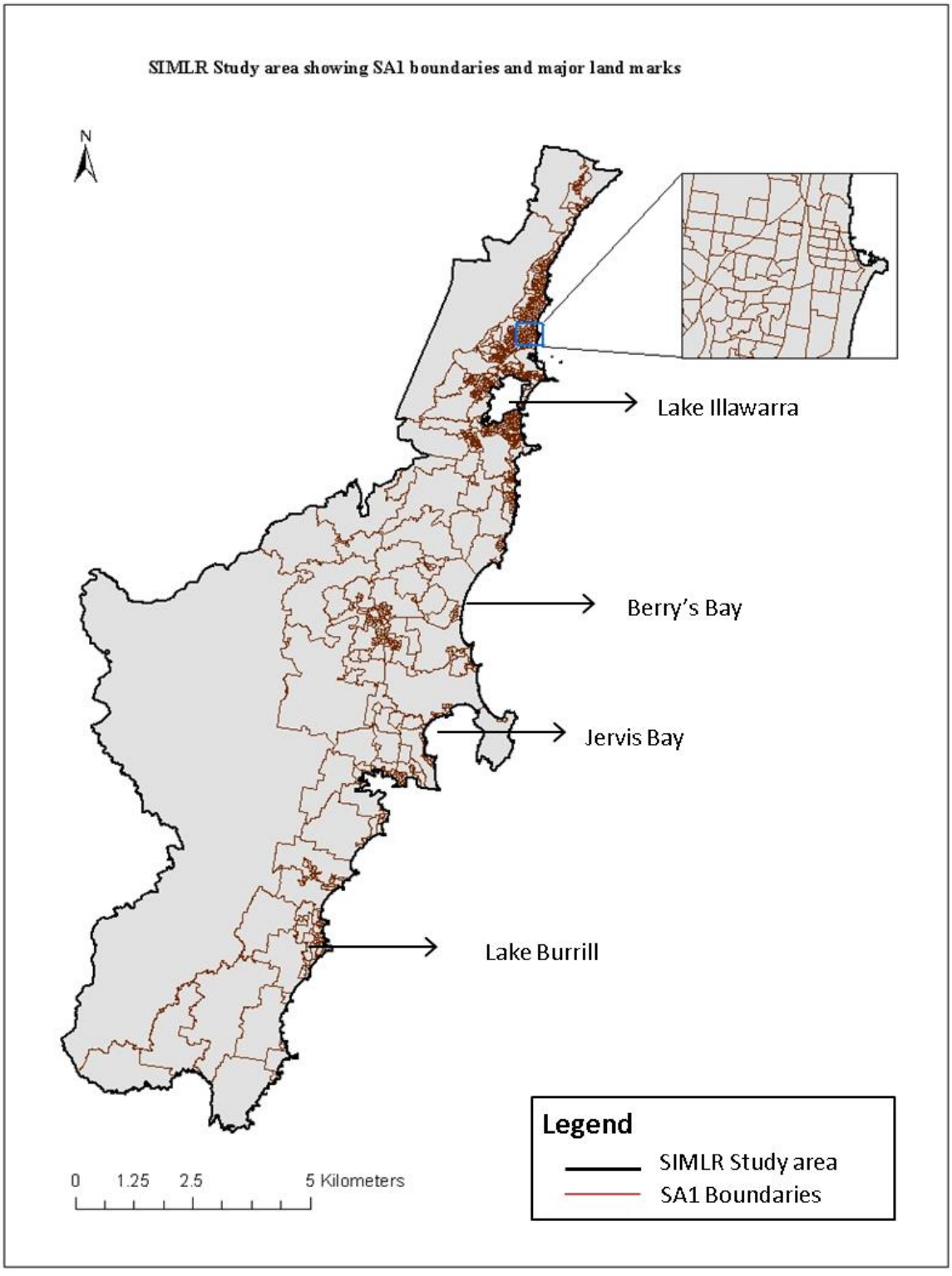

Figure 4.1 Map of the Illawarra-Shoalhaven region of NSW Australia showing SA_1 areas and major landmarks.

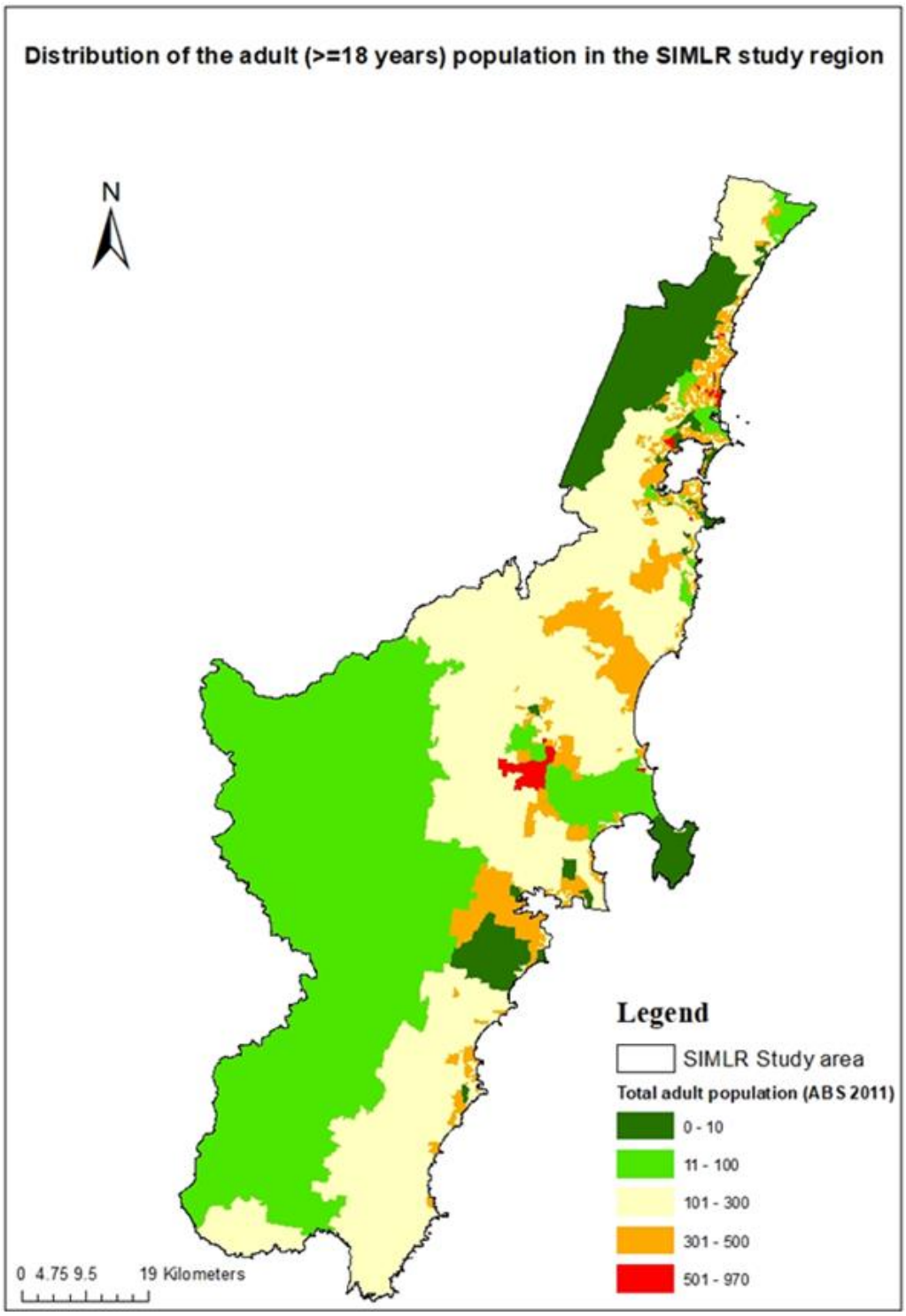

Figure 4.2 SIMLR study area showing distribution of the total adult population in SA1s.

Participants and Variables

The CMRFs test data of the adult residents of ISR between 1 Jan 2012 - 31 Dec 2017 (6 years) were extracted for analyses from the SIMLR database. Test data were extracted for eight CMRFs: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI) and diabetes mellitus (DM) status. The SIMLR database uses an algorithm to identify DM status based on diagnosis guidelines published by the Royal Australian College of General Practitioners (RACGP) and Diabetes Australia and methods from the National Health Survey of the Australian Bureau of Statistics (ABS).[50, 51] The algorithm identifies DM for HbA1c ≥ 6.5% or FBSL ≥ 7.0 mmol/l within +/- 24 months of HbA1c < 6.5%. The study data included both prevalent and incident DM cases.

Study data included only the most recent CMRF test result for each individual. We excluded extreme BMI values <12 and >80 based on cut-off points reported by Cheng (2016), Li (2009) and Littman (2012).[52-54] Table 4.1 lists the CMRFs value definitions adopted in this study and their source references.

Table 4.1 Cardiometabolic risk value definitions

CMRFs	Value definition	Adopted from
High FBSL	FBSL ≥7.0 mmol/l	RACGP guidelines.[50]
High HbA1c	HbA1c > 7.5%	RACGP guidelines.[50]
High TC	TC ≥ 5.5 mmol/l	Australian Health Survey.[51]
Low HDL	HDL < 1 mmol/l	National heart foundation of Australia.[55]
High ACR	ACR ≥ 30 mcg/L to mg/l	Kidney Health Australia.[56]
Low eGFR	eGFR < 60 mL/min/1.73m^2	Kidney Health Australia.[57]
High BMI	BMI ≥ 30(Obese)	World Health Organization (WHO).[58]
DM Status	+ve DM test algorithm	RACGP guidelines [50]; and Australian Health Survey.[51]

Statistical and spatial analyses

First, individual-level descriptive analyses of CMRFs were performed. The total number of each CMRFs tests and summary statistics of each tests' results are reported. The summary values for eGFR test results are calculated using the approach for grouped data as eGFR test result values are truncated at >90 in the SIMLR Study data. Test results were dichotomised into 'higher risk' and 'lower risk' categories based on the CMRFs definitions in Table 4.1.

Second, area-level analyses of CMRFs were undertaken. Within-cohort proportion of 'higher risk' CMRFs findings are calculated using the total number of tests within each SA1 as the denominator. The exception were DM cases, which are likely to include most prevalent cases in the study area, so SA1 adult populations aged 18 years and over were used as the denominators (accessed from ABS census 2011 data). Thereafter, an Empirical Bayes (EB) approach was used to smooth all the CMRF's raw rates to minimise extreme values arising from small sample sizes. The EB smoothed rates were then imported into GIS software for mapping and spatial statistical analyses.

As individuals with CMRFs are assumed randomly distributed within the study area, the geographic distribution of CMRFs is assumed spatially independent in this study. Global Moran's I test was used to identify spatial autocorrelation of CMRFs at a 0.05 level of significance. Global Moran's I tests if the geographic distribution of rates is clustered, dispersed or random based.[59] The global Moran's I also indicate the general strength of spatial autocorrelation in the study area, which theoretically ranges between -1 to +1. Values of I significantly above -1/(N-1) indicate positive spatial autocorrelation, where N is the number of spatial units indexed.[60] When significant spatial autocorrelation was detected, Local Indicator of Spatial Autocorrelation (LISA) spatial statistics were used to identify any clustering of CMRFs.[61] LISA was used to indicate spatial clustering of High-High (HH) or Low-Low (LL) CMRFs rates at SA1-level within the study region. False Discovery Rate (FDR) corrections were applied to LISA tests to correct p-values for multiple testing.

All descriptive statistics and EB smoothing were performed using R version 3.4.4. (R Foundation for Statistical Computing, Vienna, Austria).[62] The EBest function [63] for the 'binomial' family from the spdep package [64] in R is used for EB smoothing of the SA1's raw rates. Mapping and spatial analyses were performed using ArcGIS version 10.4.1(ESRI Inc. Redlands, CA, USA).[65]

4.5 Results

The study sample comprised 1, 132, 016 test results contributed by 256, 525 adult individuals residing in the study region. Of the 256, 525 individuals, 193, 679 (75.5%) had FBSL, 73, 885 (28.8%) had HbA1, 194, 816 (75.9%) had TC, 182, 237 had HDL (71.0%), 50, 790 had ACR (19.8%), 244, 166 had eGFR (95.2%) and 192, 443 had BMI (75.0%) test results. It was estimated 23, 704 (9.2%) of persons met the clinical criteria for diabetes. Table 4.2 provides the summary statistics of CMRFs test results. Table 4.3 presents the summary statistics of the total CMRF tests across the SA1s in the study region. Table 4.4 outlines the descriptive statistics of the individual CMRF tests across the SA1s in the study region.

Table 4.2 Summary statistics of CMRFs test results

CMRFs	Tests	Mean	SD	Min	1st Qu	Median	3rd Qu	Max
FBSL	193679	5.6	1.6	0.7	4.9	5.3	5.8	43.9
HbA1c	73885	6.0	1.3	2.6	5.3	5.6	6.4	17.8
TC	194816	5.0	1.1	1.1	4.2	4.9	5.7	39.4
HDL	182237	1.5	1.2	0.1	0.5	1.4	1.8	5.8
ACR	50790	7.4	40.3	0.1	0.4	0.8	2.3	1291.5
eGFR	244166	75.8	13.8	2.0	-	83.2	-	>90.0
BMI	192443	28.4	6.1	12.0	24.1	27.5	31.6	78.1

Table 4.3 Summary statistics of the overall CMRF tests across the SA1s in the study region.

		Summary statistics of the tests across SA1s							
		Min.	1st Qu.	Median	IQR	Mean	SD	3rd Qu.	Max.
1	Total CMRF tests/SA1	1	198	261	135	266.7	117.0	333	965
2	% Female tested /SA1	0	54.2	56.4	4.3	56.0	6.5	58.5	100
3	% Male tested /SA1	0	41.5	43.6	4.3	44.0	6.5	45.8	100
4	Total CMRF tests per person /SA1	1	4.3	4.4	0.4	4.4	0.3	4.6	6

Table 4.4 Summary statistics of the individual CMRF tests across the SA1s in the study region.

	Number of tests /SA1					Average proportion of tested people per age group (in years) /SA1						
CMRFs	**Total tests**	**Mean**	**SD**	**Median**	**IQR**	**18 - 29**	**30 - 39**	**40 - 49**	**50 - 59**	**60 - 69**	**70 - 79**	**80+**
FBSL	193679	202	87.8	199	107.5	10.2	12.1	15.4	19.5	19.8	14.8	8.2
HbA1c	73885	77.6	37.6	73.5	46	4.9	6.8	11.8	19.1	24.1	20.7	12.6
TC	194816	203.1	89.4	201	108.5	7.4	9.8	16.3	20.8	20.9	15.7	9.1
HDL	182237	190	84.2	187	101.5	6.3	9.3	16.5	21.2	21.5	16.2	9.1
ACR	50790	53.6	28.1	50	32.5	3.2	4.7	10.1	19.2	26.7	23.1	13.0
eGFR	244166	253.8	111.6	249.5	128	13.6	11.9	14.9	18.0	18.0	13.9	9.7
BMI	192455	200.5	87.6	197	103.3	11.7	11.5	15.9	19.8	20.3	15.7	5.2

The CMRFs test result values were dichotomised into 'higher risk' and 'lower risk' categories based on the CMRFs definitions in Table 4.1. The proportion of individuals with 'higher risk' CMRFs findings varied considerably between tests. The largest 'higher risk' proportions were found for BMI (33.74%) and TC (32.55%) and the lowest for ACR (4.03%). Table 4.5 provides details on the CMRFs test results classification and the identified proportions.

Table 4.5 Frequency and proportion of 'higher risk' results of CMRFs tests

Cardiometabolic risk	**Classification**	**Tests n (%)***
FBSL		193679 (100)
FBSL ≥7.0 mmol/L	Higher risk	16280(8.4)
FBG < 7.0 mmol/L	Lower risk	177399(91.6)
HbA1c		73885(100)
HbA1c > 7.5%	Higher risk	7927(10.7)
HbA1c ≤ 7.5%	Lower risk	65958(89.3)
TC		194816(100)
TC ≥ 5.5 mmol/L	Higher risk	63422(32.5)
TC < 5.5 mmol/L	Lower risk	131394(67.5)
HDL		182237 (100)
HDL < 1 mmol/l	Higher risk	21261(11.7)
HDL ≥ 1 mmol/l	Lower risk	160976(88.3)
ACR		50790(100)
ACR ≥30 mcg/L to mg/L	Higher risk	2047 (4.1)
ACR <30 mcg/L to mg/L	Lower risk	48743(95.9)
eGFR		244166(100)
eGFR < 60 mL/min/1.73m^2	Higher risk	27241(11.2)
eGFR20 ≥ 60 mL/min/1.73m^2	Lower risk	216925(88.8)
BMI		192455(100)
BMI ≥ 30 (Obesity)	Higher risk	64832(33.7)
BMI < 30	Lower risk	127511 (66.3)

Geographic distribution of cardiometabolic risk Factors

Figure 4.3 shows the geographic distribution of CMRFs at SA1-level in the ISR region with red indicating the highest and blue the lowest rates of risk. SA1s with no test data appear in white. Areas with higher rates of CMRFs were found to be clustering within the study region. The highest rates were found mainly along the populated eastern board of the study region; notably among SA1s around Lake Illawarra, south-east of Berry's bay and east of Lake Burill. However, the high TC rates showed a reversed pattern and higher rates were found in the relatively less populated central and westerly aspects of the study area. HDL rates did not follow this reversed pattern.

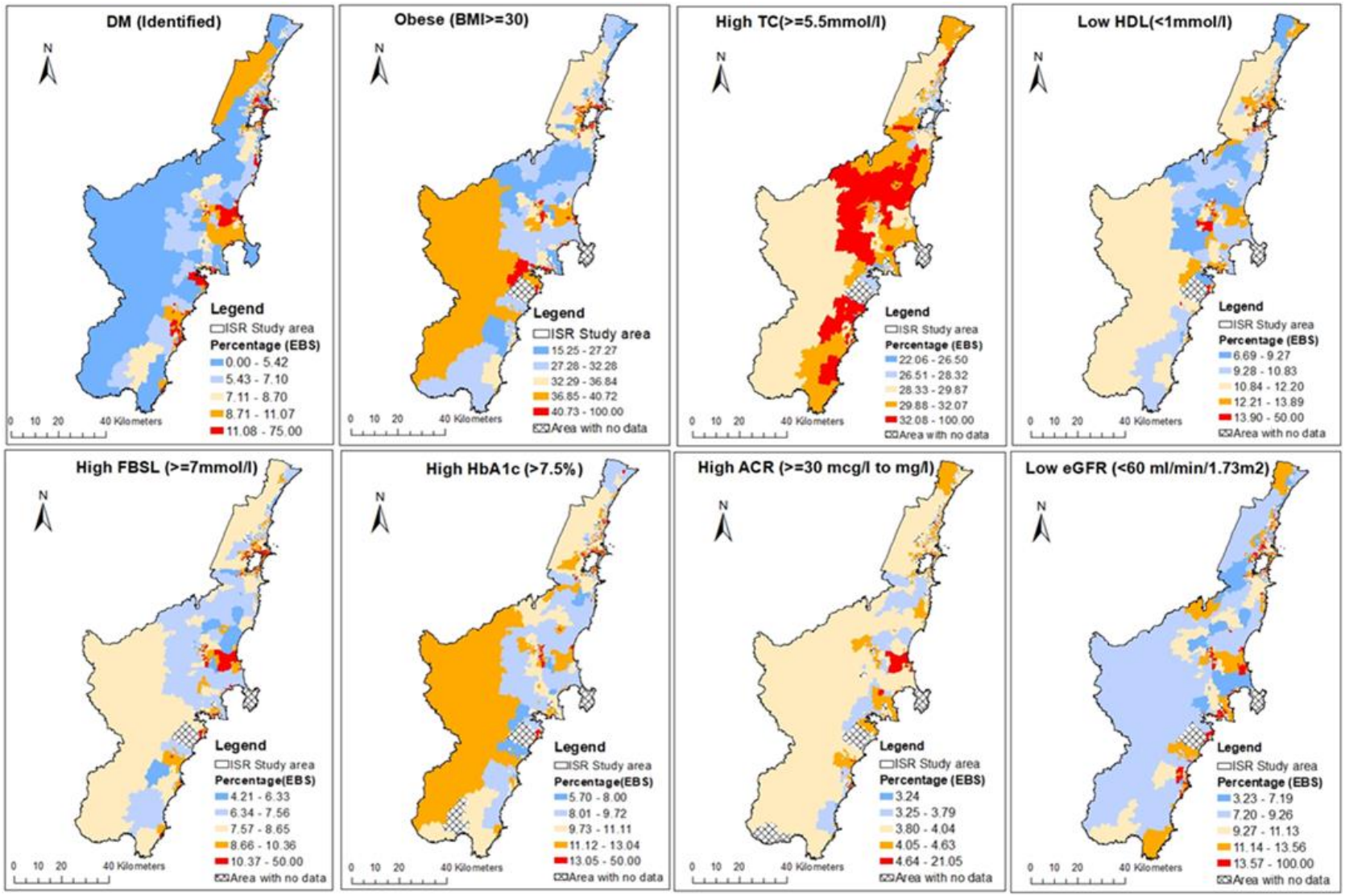

Figure 4.3: Geographic distribution of the proportion of CMRFs within the Illawarra Shoalhaven region of the NSW Australia.

Spatial Autocorrelation of CMRFs

The global Moran's I tests were significant and positive for all CMRFs (Table 4.6). The highest spatial autocorrelation strength was found among obesity rates (0.328), followed by high FBSL (0.184) and low HDL (0.174). The spatial autocorrelation strength was the lowest for albuminuria (0.028) and low eGFR (0.069).

Table 4.6: Spatial autocorrelation (Moran's I) of CMRFs

CMRFs	Moran's I	z-score	p-value
DM	0.097	27.952	<0.0001
Obesity	0.328	92.086	<0.0001
High FBSL	0.184	51.539	<0.0001
High HbA1c	0.101	28.030	<0.0001
High TC	0.146	41.154	<0.0001
Low HDL	0.174	48.733	<0.0001
Albuminuria	0.028	8.096	<0.0001
Low eGFR	0.069	19.699	<0.0001

LISA tests identified significant spatial clustering of CMRFs in the ISR region. The HH clusters were found mainly along the populated areas of the study region, except for TC. Areas around the immediate surroundings of Lake Illawarra had the most HH clusters, followed by the areas to the south-west of Berry's Bay and south of Jervis Bay. A few areas around Lake Burrill had HH clusters of DM, TC and eGFR. The LL clusters were mainly around the less populated north, central and south ends of the study area, except for TC. The TC clusters demonstrated a reverse pattern in comparison with all other CMRFs, where HH clusters were mainly around the less populated central and southern ends of the ISR and a few instances in the north-eastern end of the study area. LL clusters of TC were found around the immediate surroundings of Lake Illawarra. Figure 4.4 illustrates the spatial clustering of CMRFs in the study area.

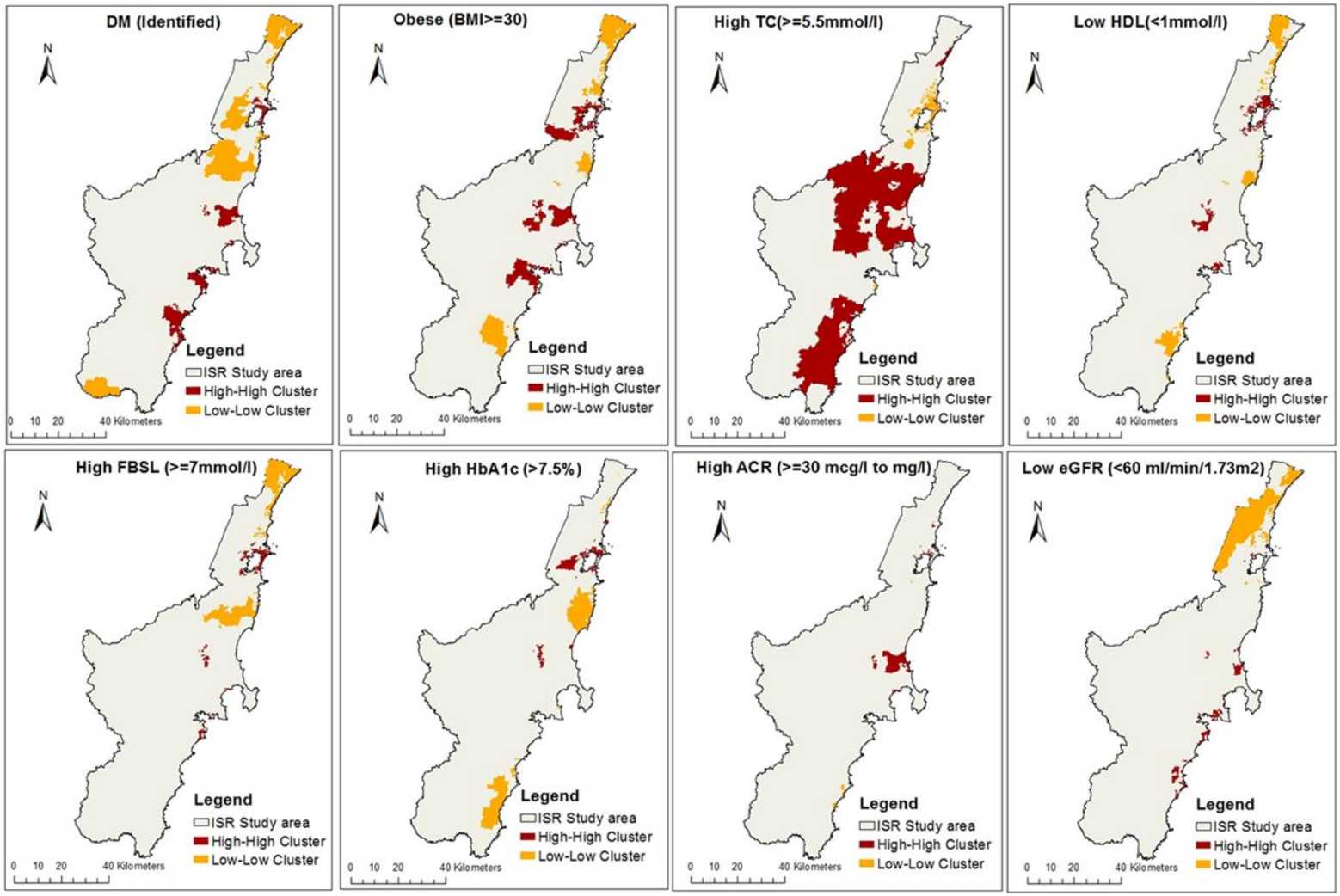

Figure 4.4: Local Moran's I cluster maps showing high-high and low-low spatial associations of CMRFs within the Illawarra Shoalhaven region of the NSW Australia.

4.6 Discussion

Place has always been a key element in human health and epidemiology. In the present study, we explored the geographic distribution of eight CMRFs in 980 SA1s in a regional area of NSW, Australia. The study is a first of its kind known to us in providing a comprehensive small-area-level profile of a wide range of CMRFs and provides an example of using population-derived routine laboratory data for area-level research.

Higher rates and clustering of CMRFs were mostly observed along the more densely populated eastern coast line of the study region. Also, some areas were common for multiple risk factors as their distribution pattern frequently converged in these areas, for example areas around Lake Illawarra, South of Jervis bay etc. However, not all populated areas were involved in this pattern and some less populated areas also had higher rates of risk. Spatial analyses revealed significant spatial autocorrelation for all eight CMRFs. Patterns of clustering were different for each CMRF at the small-area scale used in this study, which provides directions for future research using multilevel analytic methods.[66]

The distribution of high TC values were generally reversed to those distributions of other CMRFs described in this study. The reason for this observation is yet to be explored, but a possible treatment effect is suspected as the lower risk areas were often densely populated areas. It is possible that the people residing in these areas have better access to health care services and more frequently prescribed anti-cholesterol drugs.[67, 68] However, not all densely populated areas were involved in this 'higher risk' TC distribution pattern and further research is required.

The current study adds to the limited studies from Oceania reporting on geographic variation of CMRFs and the first from regional Australia. Previous studies from Australia have reported geographic variation of 42% in the odds of being diagnosed with DM among adults living in Sydney.[37] Another study reported geographic variation in glycated haemoglobin (HbA1c) values across 767 Census Collection Districts (CDs) in Adelaide.[44] The study builds on previous research by investigating the distribution of a wide range of CMRFs, which appears to be rare in the literature.

This study must be considered within its limitations. First, the cross-sectional design of the study precludes causal inference. Second, the descriptive analyses performed in this study indicate only significant variations in the geographic distribution of CMRFs, but do not differentiate the individual and/or area-level attributes which might be contributing to this variation.[13] Third, the maps include areas with no test data. Fourth, the study data were obtained from people attending health care services; therefore its point-estimates may not be representative of the general population. Fifth, we cannot exclude the possibility that a higher proportion of positive tests in an area could be due to greater access to pathology services; however exploring this possibility was beyond the scope of the current study.

Future research is required to understand the reasons for the geographic variation reported in this paper. The findings reported in this study suggest hypotheses that will be further explored using appropriate multilevel/hierarchical analyses to differentiate and quantify the individual and area-level contributions to this variation.[66, 69-71] Such hierarchical analyses will have the potential to inform development of appropriate area-level health care service policy initiatives. It is important to differentiate the contributions of individual (e.g. age, sex, etc.) and area-level (e.g. socioeconomic disadvantage, access or proximity to health care services, etc) attributes to the different patterns of clustering to inform targeted area-level preventive interventions and future health service commissioning decisions to these areas.

4.7 Conclusion

In conclusion, area-level descriptive analyses of CMRFs have the potential to highlight inequalities in the geographic distribution of CMRFs. Regional planning for the prevention and management of CMRFs requires information about its epidemiology within specific communities or areas. Centralised approaches of disease prevention and management may not suit regional requirements as the disease pattern in regional areas may differ to those in metropolitan areas and cities. Area specific evidence through regional health care research is important to inform health care service commissioning for area specific decisions and policy developments. This paper demonstrates an initial step in such regional health care research and a feasible method using population data derived from routine clinical practice.

References

1. D'agostino, R.B., et al., *General cardiovascular risk profile for use in primary care.* Circulation, 2008. **117**(6): p. 743-753.
2. Danaei, G., et al., *Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: a comparative risk assessment.* Lancet Diabetes & Endocrinology, 2014.
3. Gansevoort, R.T., et al., *Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention.* The Lancet, 2013. **382**(9889): p. 339-352.
4. Hubert, H.B., et al., *Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study.* Circulation, 1983. **67**(5): p. 968-977.
5. Liu, M., et al., *Cardiovascular disease and its relationship with chronic kidney disease.* Eur Rev Med Pharmacol Sci, 2014. **18**(19): p. 2918-26.
6. Wadwa, R.P., E.M. Urbina, and S.R. Daniels, *15 Cardiovascular Disease Risk Factors.* Epidemiology of Pediatric and Adolescent Diabetes, 2008: p. 235.
7. World Health Organization. *The top 10 causes of death*. 24 May 2018 [cited 2020 Sept 5]; Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
8. World Health Organization. *WHO | Noncommunicable diseases* 2017 [cited 2018 10 March]; Available from: : http://www.who.int/mediacentre/factsheets/fs355/en/.
9. World Health Organization; World Heart Federation and World Stroke Organization. *Global atlas on cardiovascular disease prevention and control*. 2011; Available from: https://www.cabdirect.org/cabdirect/abstract/20123402600%0Afile:///C:/Users/USER/Downloads/9789241564373_eng (2).pdf.
10. Australian Institute of Health and Welfare, *Cardiovascular disease, diabetes and chronic kidney disease—Australian facts: Morbidity–Hospital care in Cardiovascular, diabetes and chronic kidney disease* in *series no. 3*. : Canberra: AIHW; 2014.
11. Australian Institute of Health and Welfare. *Cardiovascular disease, diabetes and chronic kidney disease—Australian facts: Morbidity–Hospital care in Cardiovascular, diabetes and chronic kidney disease* series no. 3 2017; Available from: https://www.aihw.gov.au/reports/heart-stroke-vascular-disease/cardiovascular-diabetes-chronic-kidney-morbidity/contents/table-of-contents
12. Australian Institute of Health and Welfare. *Australia's health 2014*. 2014; Available from: http://www.aihw.gov.au/publication-detail/?id=60129547206.
13. Merlo, J., et al., *The tyranny of the averages and the indiscriminate use of risk factors in public health: The case of coronary heart disease.* SSM-population health, 2017. **3**: p. 684-698.
14. Ard, J.D., et al., *Effects of calorie restriction in obese older adults: the CROSSROADS randomized controlled trial.* The Journals of Gerontology: Series A, 2018. **73**(1): p. 73-80.
15. Dehghani, A., et al., *Influence of comprehensive life style intervention in patients of CHD.* Global Journal of Health Science, 2015. **7**(7): p. 6.
16. Ma, J., et al., *Evaluation of lifestyle interventions to treat elevated cardiometabolic risk in primary care (E-LITE): a randomized controlled trial.* BMC Family Practice, 2009. **10**(1): p. 71.

17. Tourlouki, E., A.-L. Matalas, and D.B. Panagiotakos, *Dietary habits and cardiovascular disease risk in middle-aged and elderly populations: a review of evidence.* Clinical interventions in aging, 2009. **4**: p. 319.
18. Weiss, E.P. and L. Fontana, *Caloric restriction: powerful protection for the aging heart and vasculature.* American Journal of Physiology-Heart and Circulatory Physiology, 2011. **301**(4): p. H1205-H1219.
19. Nissinen, A., X. Berrios, and P. Puska, *Community-based noncommunicable disease interventions: lessons from developed countries for developing ones.* Bulletin of the world Health Organization, 2001. **79**: p. 963-970.
20. Parker, D.R. and A.R. Assaf, *Community interventions for cardiovascular disease.* Primary care: Clinics in office practice, 2005. **32**(4): p. 865-881.
21. Duffany, K.O.C., et al., *Community Interventions for Health (CIH): a novel approach to tackling the worldwide epidemic of chronic diseases.* CVD Prevention and Control, 2011. **6**(2): p. 47-56.
22. Auchincloss, A.H., et al., *A review of spatial methods in epidemiology, 2000–2010.* Annual review of public health, 2012. **33**: p. 107-122.
23. Fradelos, E.C., et al., *Health based geographic information systems (GIS) and their applications.* Acta Informatica Medica, 2014. **22**(6): p. 402.
24. Lawson, A.B., *Statistical methods in spatial epidemiology*. 2013: John Wiley & Sons.
25. Craglia, M. and R. Maheswaran, *GIS in public health practice*. 2016: CRC press.
26. Cromley, E.K., *Using GIS to address epidemiologic research questions.* Current Epidemiology Reports, 2019. **6**(2): p. 162-173.
27. Jarrahi, A.M., M. Zare, and A. Sadeghi, *Geographic information systems (GIS), an informative start for challenging process of etiologic investigation of diseases and public health policy making.* Asian Pacific Journal of Cancer Care, 2017. **2**(1): p. 1-1.
28. Nykiforuk, C.I. and L.M. Flaman, *Geographic information systems (GIS) for health promotion and public health: a review.* Health promotion practice, 2011. **12**(1): p. 63-73.
29. Portnov, B.A., J. Dubnov, and M. Barchana, *On ecological fallacy, assessment errors stemming from misguided variable selection, and the effect of aggregation on the outcome of epidemiological study.* Journal of exposure science & environmental epidemiology, 2007. **17**(1): p. 106-121.
30. Rezaeian, M., et al., *Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary.* Journal of Epidemiology & Community Health, 2007. **61**(2): p. 98-102.
31. Shafran-Nathan, R., et al., *Ecological bias in environmental health studies: the problem of aggregation of multiple data sources.* Air Quality, Atmosphere & Health, 2017. **10**(4): p. 411-420.
32. Lawson, A.B., *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. 2018: CRC press.
33. Lawson, A.B., W.J. Browne, and C.L.V. Rodeiro, *Disease mapping with WinBUGS and MLwiN*. Vol. 11. 2003: John Wiley & Sons.
34. Lawson, A.B. and K. Kleinman, *Spatial and syndromic surveillance for public health*. 2005: John Wiley & Sons.
35. Barber, S., et al., *At the intersection of place, race, and health in Brazil: Residential segregation and cardio-metabolic risk factors in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil).* Social Science & Medicine, 2018. **199**: p. 67-76.
36. Gabert, R., et al., *Identifying high-risk neighborhoods using electronic medical records: a population-based approach for targeting diabetes prevention and treatment interventions.* PLoS One, 2016. **11**(7): p. e0159227.

37. Astell-Burt, T. and X. Feng, *Geographic inequity in healthy food environment and type 2 diabetes: can we please turn off the tap?* Medical Journal of Australia, 2015. **203**(6): p. 246-248.
38. Toms, R., et al., *Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature.* International journal of health geographics, 2019. **18**(1): p. 1.
39. Zhou, M., et al., *Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults.* Diabetes care, 2015. **38**(1): p. 72-81.
40. Alkerwi, A., et al., *Geographic variations in cardiometabolic risk factors in Luxembourg.* International journal of environmental research and public health, 2017. **14**(6): p. 648.
41. Barker, L.E., et al., *Geographic distribution of diagnosed diabetes in the US: a diabetes belt.* American journal of preventive medicine, 2011. **40**(4): p. 434-439.
42. Lawlor, D., et al., *Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study.* Journal of Epidemiology & Community Health, 2003. **57**(2): p. 134-140.
43. Ocaña-Riola, R., *Common errors in disease mapping.* Geospatial health, 2010. **4**(2): p. 139-154.
44. Paquet, C., et al., *Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia.* International journal of environmental research and public health, 2016. **13**(5): p. 519.
45. Valdés, S., et al., *Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@ bet. es study.* Revista Española de Cardiología (English Edition), 2014. **67**(6): p. 442-448.
46. Guh, D.P., et al., *The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis.* BMC public health, 2009. **9**(1): p. 88.
47. Australian Bureau of Statistics. *2011 Census data* [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280.
48. Bonney, A., et al., *Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users–a cross-sectional study.* PLoS One, 2015. **10**(8): p. e0137261.
49. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure and Greater Capital City Statistical Areas:STATISTICAL AREA LEVEL 1 (SA1)*. 2016 [cited 2018 Sept 5]; Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by Subject/1270.0.55.001~July 2016~Main Features~Statistical Area Level 1 (SA1)~10013
50. The Royal Australian College of General Practitioners & Diabetes Australia. General Practice Management of Type 2 Diabetes 2016-2018. The Royal Australian College of General Practitioners (2016). doi:10.1007/s00125-010-2011-6.
51. Australian Bureau of Statistics. *Australian Health Survey: Biomedical Results for Chronic Diseases, 2011-12.* 2013 [cited 2020 Sept 5]; Vol. 4364.0.55.:[Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.005main+features12011-12
52. Cheng, F.W., et al., *Body mass index and all-cause mortality among older adults.* Obesity, 2016. **24**(10): p. 2232-2239.
53. Li, W., et al., *Small-area estimation and prioritizing communities for tobacco control efforts in Massachusetts.* American journal of public health, 2009. **99**(3): p. 470-479.

54. Littman, A.J., et al., *Evaluation of a weight management program for veterans.* Preventing chronic disease, 2012. **9**.
55. *National heart foundation of Australia. Lipid management profile for health professionals. Available at: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management.*
56. *National Kidney foundation(USA). Albumin creatinine Ratio (ACR). (2018). Available at: https://www.kidney.org/kidneydisease/siemens_hcp_acr.*
57. *Kidney Health Australia. Fact sheet: Estimated Glomerular Filtration Rate (eGFR) [Internet]. Available: www.kidney.org.au.*
58. *WHO. Obesity : Preventing and managing the global epidemic. World Health Organization: Technical Report Series. WHO Technical Report Series, no. 894. (2000). doi:ISBN 92 4 120894 5.*
59. Li, H., C.A. Calder, and N. Cressie, *Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model.* Geographical Analysis, 2007. **39**(4): p. 357-375.
60. Moran, P.A., *Notes on continuous stochastic phenomena.* Biometrika, 1950. **37**(1/2): p. 17-23.
61. Anselin, L., *Local indicators of spatial association—LISA.* Geographical analysis, 1995. **27**(2): p. 93-115.
62. The R Core Team. *R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria*. 2018; Available from: https://www.R-project.org/.
63. Martuzzi, M. and P. Elliott, *Empirical Bayes estimation of small area prevalence of non-rare conditions.* Statistics in Medicine, 1996. **15**(17): p. 1867-1873.
64. Bivand, R.S. and D.W. Wong. *Comparing implementations of global and local indicators of spatial association*. TEST 2018 [cited 27 3]; 716-748]. Available from: https://doi.org/10.1007/s11749-018-0599-x.
65. Environmental Systems Research Institute (ESRI). *ArcGIS 10.4.1, ESRI Inc. Redlands, CA, USA.*; Available from: https://www.esri.com/
66. Merlo, J., et al., *An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health.* PloS one, 2016. **11**(4): p. e0153778.
67. Stocks, N., et al., *Gender, socioeconomic status, need or access? Differences in statin prescribing across urban, rural and remote Australia.* Australian Journal of Rural Health, 2009. **17**(2): p. 92-96.
68. Stocks, N.P., et al., *Statin prescribing in Australia: socioeconomic and sex differences.* Medical journal of Australia, 2004. **180**(5): p. 229-231.
69. Merlo, J., et al., *Population effects on individual systolic blood pressure: a multilevel analysis of the World Health Organization MONICA Project.* American Journal of Epidemiology, 2004. **159**(12): p. 1168-1179.
70. Merlo, J., et al., *Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia, Spain.* Social science & medicine, 2012. **75**(8): p. 1477-1487.
71. Merlo, J., P. Wagner, and G. Leckie, *A simple multilevel approach for analysing geographical inequalities in public health reports: The case of municipality differences in obesity.* Health & place, 2019. **58**: p. 102145.

1.

Chapter 5: Geographic Variation in Cardiometabolic Risk Factor Distribution Explained by Area-level Disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia

Chapter 5: Geographic Variation in Cardiometabolic Risk Factor Distribution Explained by Area-level Disadvantage in the Illawarra-Shoalhaven Region of the NSW, Australia

5.1 Publication profile

This chapter presents the substantive content of research submitted for publication in the journal **Nature - *Scientific Reports***. Parts of the methods and findings from this study were also presented in a national level conference held in Australia.

Journal article

Toms, R., Mayne, D. J., Feng, X., & Bonney, A. (2020). Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia. **Nature - Scientific Reports**, 10(1), 1-18.

Available from: https://www.nature.com/articles/s41598-020-69552-4

DOI: https://doi.org/10.1038/s41598-020-69552-4

Published version of the article

The published version of the article is appended within the **'Supplementary Materials'** section of the thesis as **Supplementary Material 3**.

Peer reviewed abstracts

National Health and Medical Research Council (NHMRC) Symposium

Toms, R., Bonney, A., Mayne, D. J., &Feng, X. (2019) Area-level socioeconomic disadvantage and cardiometabolic risk distribution: an analysis of 256, 565 adult residents in the Illawarra- Shoalhaven region of the NSW Australia. (Poster presentation) 8th Annual NHMRC Symposium on Research Translation **National Health and Medical Research Council** 19 - 20 November 2019, Melbourne, Australia.

Available from: https://scholars.uow.edu.au/display/publication140392

5.2 Abstract

Cardiometabolic risk factors (CMRFs) demonstrate significant geographic variation in their distribution. The study aims to quantify the general contextual effect of the areas on CMRFs; and the geographic variation explained by area-level socioeconomic disadvantage. A cross sectional design and multilevel logistic regression methods were adopted. Data included objectively measured routine pathology test data between years 2012 and 2017 on: fasting blood sugar level; glycated haemoglobin; total cholesterol; high density lipoprotein; urinary albumin creatinine ratio; estimated glomerular filtration rate; and body mass index. The 2011 Australian census based Index of Relative Socioeconomic Disadvantage (IRSD) were the area-level study variables, analysed at its smallest geographic unit of reporting. A total of 1,132,029 CMRF test results from 256,525 individuals were analysed. After adjusting for individual-level covariates, all CMRFs significantly associated with IRSD and the probability of higher risk CMRFs increases with greater area-level disadvantage. Though the specific contribution of IRSD in the geographic variation of CMRF ranged between 57.8 and 14.71%, the general contextual effect of areas were found minimal (ICCs 0.6–3.4%). The results support universal interventions proportional to the need and disadvantage level of populations for the prevention and control of CMRFs, rather than any area specific interventions as the contextual effects were found minimal in the study region.

Key words: cardiometabolic risk factor; area-level disadvantage; multilevel logistic regression models; area-level variance; geographic variation.

5.3 Introduction

The distribution of cardiometabolic risk factors (CMRFs) varies geographically.[1, 2] Previous research has reported higher prevalence and clustering of CMRFs in certain localities: typically in areas of higher socioeconomic disadvantage.[3-23] Quantifying the geographic variation in CMRFs contributed by area-level socioeconomic disadvantage can aid in designing appropriate area-level preventive approaches for CMRFs. Chronic and uncontrolled CMRFs predispose individuals to the development of cardiovascular disease (CVD), which continues to be the leading cause of health care expenditure and premature mortality worldwide.[24]

In Australia, a social gradient is observed in the distribution of many chronic conditions including various CMRFs (e.g. diabetes and chronic kidney disease).[25] Generally, Australians enjoy better health than people in many other countries in the world. However, within Australia this better health is not equally distributed.[26] It is well-recognised that socioeconomically disadvantaged individuals in Australia, on average, experience a greater disease burden than their less disadvantaged counterparts.[25-28]This tendency is also evident at a contextual level when studies have investigated association of CMRFs with area-level socioeconomic disadvantage in Australia[4, 17] and globally[5, 6, 8, 10, 11, 14-16, 19, 22, 29-33].

Consistent with this, men from highly urbanised environments have been reported to have higher incidence of coronary heart disease with increasing residential area socioeconomic disadvantage, after adjusting for individual characteristics.[16]Also, lower area-level disadvantage has been reported as being associated with lower rates of behavioural cardiac risk factors such as smoking, physical inactivity and obesity in some studies.[10, 11, 34] Most of the reported associations of CMRFs with area-level socioeconomic disadvantage were independent of individual-level characteristics such as age and educational attainment. Even though the area-level associations of CMRFs were significant in these studies, the results were often dependent on the CMRF analysed, the measures of area-level socioeconomic disadvantage and the geographic scale at which associations were examined.[35]

Multilevel analyses of CMRFs based on the average measures of association or variation alone are insufficient to report the geographical variance as similar associations were possible with very different scenarios of area variance.[36] Multilevel findings extending on the general contextual effects and reporting the proportion of the total area-level variance along with the measures of clustering and the average measures of association or variation are appropriate and informative in reporting area-level influences, but less common.[20, 36-38] To differentiate the relative importance of individual versus area-level interventions for the prevention and control of CMRFs, the geographical component of the total individual risk variance has to be identified in a multilevel approach.

Therefore, the aims of this study are to (1) quantify the general contextual or geographic effect of areas on CMRFs, over and above their individual-level compositions; and to (2) quantify the geographic variation across multiple CMRFs specifically explained by area-level socioeconomic disadvantage, within the Illawarra-Shoalhaven region of NSW Australia. Quantification of the general contextual effect and the variation specifically explained by area-level socioeconomic disadvantage will assist our understanding of the socioeconomic context of CMRFs in the study region and provide guidance for health service commissioning more generally nationally.

5.4 Methods

A cross-sectional multilevel design was adopted to account for the hierarchical nature of the data and analyses. Informed consent was not obtained for the individual-level data used in this study, as the study used existing data which were already de-identified. The study was approved by the University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical Human Research Ethics Committee (HREC protocol No: 2017/124). All the methods and analyses were performed meeting the relevent ethical guidelines and regulations of the committee.

5.4.1 Study area and data

The study was conducted in the Illawarra-Shoalhaven region of the New South Wales (NSW) state in Australia. The Illawarra-Shoalhaven region is geographically a coastal plain along the south-east of NSW; situated at the immediate south of the metropolitan boundaries of Sydney; and encompasses multiple regional cities, towns and rural areas. This region covers a land area of 5, 615 km^2, and had an estimated residential population of 369, 469 at the time of the 2011 Australian Census of Population and Housing conducted by the Australian Bureau of Statistics (ABS).[39] Statistical Area level 1 (SA1), the smallest geographical unit of the 2011 census data release, was the area-level unit of analysis in this study.[40]SA1s typically have a population size of 200 to 800 persons (average 400), and the Illawarra-Shoalhaven region covers a total of 980 conterminous SA1s.[40]

The Illawarra-Shoalhaven has a diverse socio-economic profile, making it a useful region for area-level population health studies.[41] The population profile of the Illawarra-Shoalhaven region is culturally and linguistically diverse, with a significant proportion of non-English speaking people (10.5%) residing in this region who have migrated from overseas. [42] In addition, at the 2011 Census the region is identified to have more than the NSW state and Australian national averages of: 1) Aboriginal and/or Torres Strait Islander peoples (3% versus 2.5% NSW and 2.5% Australia); 2) aged (>=65 years) population (17.6% versus 14.5% NSW and 13.8% Australia) ; 3) single-parent households (5.8% versus 5.3% NSW and 5.2% Australia); and 4) unemployment (7.1% versus 5.1% NSW and 5.1% Australia) and lower labour force participation rate (57.9% versus 64.6% NSW and 66.2% Australia).[42] ABS 2011 census data indicate that more than 31 % of people in the Illawarra-Shoalhaven reside in Inner Regional areas, and 9.1% of households within this region did not have a motor vehicle.[42] The Illawarra-Shoalhaven geography and a limited public transport system, especially in isolated communities, make it difficult for many people to access health services quickly.[42] These characteristics of the study region directly indicate the vulnerability of its population to poorer health outcomes.

The CMRF test data in this study were extracted from the Southern IML Research (SIMLR) Study database, which is comprised of de-identified and internally linked pathology results from a major network of pathology services in the study region. The individual-level data in SIMLR database are geocoded to their corresponding SA1 areas, but not to their residential address, for privacy and confidentiality concerns. More details on this data source, procurement and access are published elsewhere.[17] The CMRF test data were extracted for non-pregnant individuals aged 18 years or older presenting for testing between 01 January 2012 and December 2017. Only the most recent test result was included if an individual had undergone the same test multiple times in this data period. Test data with missing details on the individual and area-level factors analysed in this study were excluded from the analyses.

5.4.2 Variables

Outcome variable

Results of the CMRF tests were the individual-level outcome variables. Data on the seven CMRF tests analysed in this study included: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); urinary albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); and objectively-measured body mass index (BMI). These CMRF test results were dichotomised into higher risk and lower risk values based on the current national and international guidelines on risk definitions (Table 5.1).

Table 5.1: Definitions of CMRFs test results

	***'Higher risk'* CMRFs**	**Definition**
1.	High FBSL	FBSL ≥ 7.0 mmol/l.[43]
2.	High HbA1c	HbA1c > 7.5%.[43]
3.	High TC	TC ≥ 5.5 mmol/l.[44]
4.	Low HDL	HDL < 1 mmol/l.[45]
5.	High ACR	ACR ≥ 30 mcg/L to mg/l.[46]
6.	Low eGFR	eGFR < 60 mL/min/1.73m^2.[46]
7.	Obesity	BMI ≥ 30 kg/m^2.[47]

CMRFs - Cardiometabolic risk factors; **FBSL** - Fasting Blood Sugar Level; **HbA1c** - Glycated Haemoglobin; **TC** - Total Cholesterol; **HDL** - High Density Lipoprotein; **ACR** - Albumin Creatinine Ratio; **eGFR** - estimated Glomerular Filtration Rate; **BMI** - Body Mass Index.

Study variable

The 2011 ABS census based Index of Relative Socioeconomic Disadvantage (IRSD) of the SA1s was the study variable. IRSD summarises a range of measures of relative socioeconomic disadvantage of people and households within SA1s and includes: level of income; education; employment; family structure; disability; housing; transportation; and internet connection.[48] This study uses IRSD reported as quintiles; the lowest quintile (Q1) indicating the most disadvantaged SA1s and the highest quintile (Q5) the least disadvantaged SA1s.[48] The IRSD quintiles in the study were derived by ABS from the distribution of IRSD scores for the Illawarra-Shoalhaven region based on the 2011 census. The study region has a diverse IRSD profile with representation across IRSD scores in comparison with Australia as a whole, making the region useful for population-level studies.[41]

Covariates

Analyses were adjusted for sex (male and female) and age group (18–29y, 30–39y, 40–49y, 50–59y, 60–69y, 70–79y, 80+ years) of each individual at the time of the pathology collection of the CMRFs tests analysed in this study.

5.4.3 Statistical analyses

Initially, descriptive statistics of all individual and area-level variables were performed. Thereafter, single level and multilevel logistic regression models were fitted for the CMRF test data of individuals (Level 1), nested within SA1s (Level 2). For each of the seven CMRFs analysed in this study, a hierarchy of four multilevel models at SA1 level were fit that included fixed effects for age, sex and IRSD and random effect

(intercept) for SA1. Model 0 was a single level model adjusted for age and sex; Model 1 (M1) was null model at level 2; Model 2 (M2) adjusted for age and sex at level 2; Model 3 (M3) adjusted for the area-level study variable (IRSD) only at level 2; and the final model Model 4 (M4) included both M2 and M3 covariates (age, sex and IRSD) at level 2. The estimated regression coefficients of the derived models were exponentiated to calculate odds ratios (ORs). The goodness of fit of the models were identified using Likelihood Ratio Tests (LRT) at $p < 0.05$ level of significance. The general equation of the fully adjusted model is:

$$y_{ij} \sim Binomial(1, \pi_{ij}) \quad (5.1)$$

$$logit(\pi_{ij}) = \beta_0 + \beta_1 age_{ij} + \beta_2 sex_{ij} + \beta_3 IRSD_{ij} + u_j \quad (5.2)$$

$$u_j \sim N(0, \tau^2{}_u), \quad (5.3)$$

where: y_{ij} denote the binary response of CMRF test outcome (as 'higher risk' or 'lower risk', based on the adopted definitions) for individual i in the area (SA1) j; π_{ij} denotes the probability that individual i in area (SA1) j has a 'higher risk' CMRF test outcome given their individual-level age_{ij} and sex_{ij}; and their area-level IRSD index $IRSD_{ij}$. $\beta_1, \beta_2, \beta_3$ are the regression coefficients which measure the associations between the log-odds of the CMRF outcome and each covariate all else being equal, and when exponentiated these are translated to ORs.[36] u_j is the random effect for the area (SA1) j and $\tau^2{}_u$ is the area level variance, which has to be estimated.

Model comparison

The Akaike Information Criterion (AIC) was used to evaluate model fit. The derived models were compared for: area level variance (τ^2) at SA1 level; proportional change in variance (PCV); Intra-cluster Correlation Coefficients (ICC); Median Odds Ratios (MORs); area under the receiver operating characteristic (AUC) curve; and the change in AUC.

The τ^2 s was initially identified from each model. PCVs were calculated for models M2 to M4 relative to M1. The ICCs of the fitted models were calculated using the latent variable approach.[49] This approach assumes that a latent continuous outcome underlies the observed dichotomous outcomes and it is this latent outcome for which the ICC is calculated and interpreted. The ICC measured the expected correlation in CMRF outcomes between two individuals from the same SA1. The higher the ICC, the more relevant area-level context is for understanding individual latent outcome variation.[36] The MOR is calculated as an alternative way of interpreting the magnitude of area-level variance. The MOR translated the area-level variance which were estimated on the log-odds scale to the commonly used OR scale. The MOR result value is interpreted as the median increased odds of identifying the outcome if an individual move to another SA1 with higher risk. Thus, the higher the MOR the greater the general area-level effect and it will equal to 1 in the absence of area-level variance.[36] The general contextual effect of the geographic areas over and above their individual-level composition of the higher risk CMRFs, is obtained through the measure of clustering (ICC) in M2s. The geographic variance and ICC in the null models (M1s) of higher risk CMRFs may depend on both the contextual and individual-level variables. Therefore, M2s of the higher risk CMRFs which adjusted for individual-level attributes is better to provide information on the

'general contextual effect' of the areas. The unique contribution of the area-level study variable (IRSD) to the area-level variance of higher risk CMRFs were assessed through the PCVs between M2s and M4s.

The receiver operating characteristic (ROC) curves are created by plotting the true positive rates (TPR) i.e. sensitivity, against the false positive rates (FPR), i.e. 1 specificity for different binary classification thresholds of the predicted probabilities in all the models.[50] Post-estimation, predicted probabilities (πij) are calculated for each individual and are used to calculate the AUC for the model. The AUCs of the models measure the capacity of the models to correctly classify individuals with or without the outcome of a higher risk CMRFs analysed in this study, as a function of their predicted probabilities.[36] The AUC values range from 1 and 0.5, where 1 is the perfect predictive discrimination and 0.5 have no predictive power.[51] The AUCs also indicate the general contextual effects and can be compared it to the ICC and the MOR values.[36] The added value of knowing an individual's area of residence besides individual-level information (age and sex) can be obtained through the AUC change in Model 2 in reference to Model 0, where a higher AUC change would indicate higher relevance of areas in relation to CMRFs.

Statistical package

All analyses were performed using R version 3.4.4. (R Foundation for Statistical Computing, Vienna, Austria).[52] Multi-level models were fit using the glmer function in the lme4 package[53]; and likelihood ratio tests were calculated using the lrtest function in the lmtest package[54]; and ROC curves using the roc function in the pROC package[55].

5.5 Results

A total of 1, 132, 029 CMRFs test data which belong to 256, 525 individuals were extracted for the analyses. Figure 5.1 provides a flow chart of the individual tests in CMRF test data. The mean number of tests per person was 4.4. After removing 1162 (1.0%) test results data with missing details, a total of 1, 130, 894 tests were included in the analytic data set.

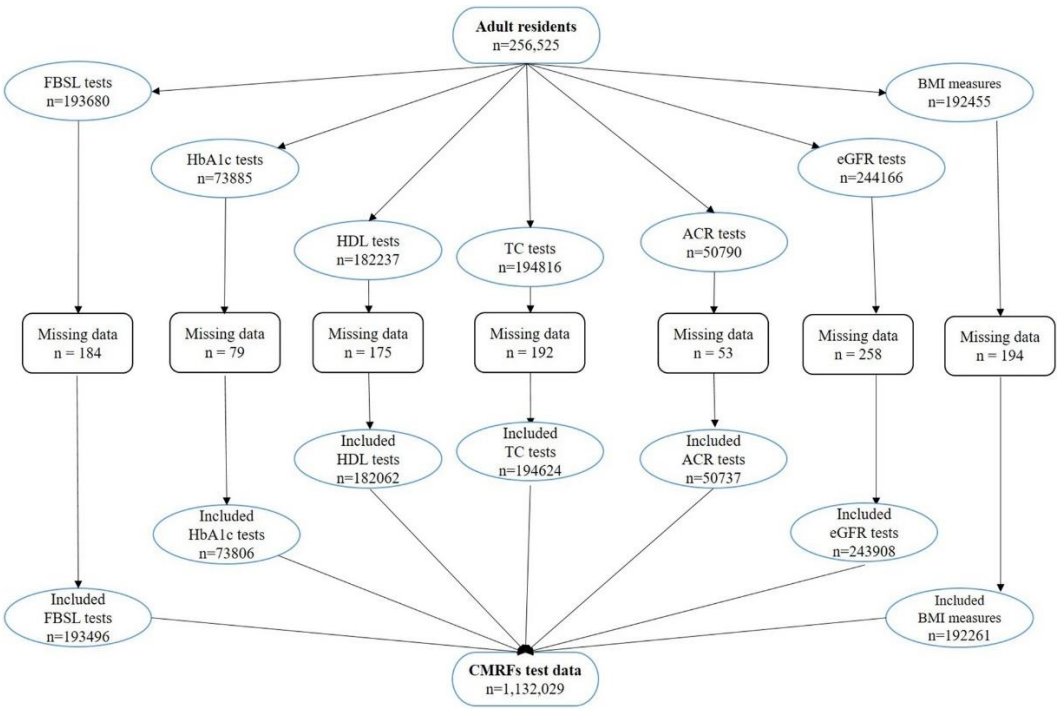

Figure 5.1: Flow chart of the included/excluded tests in the CMRFs test data.

Table 5.2 provides details of the missing data and test data distribution of each CMRF tests. Most frequently missing data were the IRSD indices from SA1s in the study area for which an IRSD index was not available from ABS 2011 census either due to low populations or poor data quality.[56]

Table 5.2: Table of excluded test data which had missing details

	FBSL	HbA1c	TC	HDL	ACR	eGFR	BMI	Total
Extracted	193680	73885	194816	182237	50790	244166	192455	**1132029**
Missing data:								
Test value	1	0	0	0	0	0	0	1
Age	1	1	1	1	0	2	1	7
Sex	0	0	0	0	0	0	0	0
IRSD	182	78	191	174	53	256	193	1154
SA1 coding	0	0	0	0	0	0	0	0
Excluded tests:	184	79	192	175	53	258	194	1162
Included tests: Total n	**193496**	**73806**	**194624**	**182062**	**50737**	**243908**	**192261**	**1130894**

FBSL - Fasting Blood Sugar Level; **HbA1c** - Glycated Haemoglobin; **TC** - Total Cholesterol; **HDL** - High Density Lipoprotein; ACR - Albumin Creatinine Ratio; **eGFR** - estimated Glomerular Filtration Rate; **BMI** - Body Mass Index; **IRSD –** Index of the Relative Socioeconomic Disadvantage**; SA1**- Statistical Area level 1**.**

Table 5.3 presents the summary statistics of the total CMRF tests across the SA1s in the study region and Table 5.4 outlines the descriptive statistics of the individual CMRF tests across the SA1s in the study region.

Table 5.3: Summary statistics of the overall CMRF tests across the SA1s in the study region.

		Summary statistics of the tests across SA1s							
		Min.	1st Qu.	Median	IQR	Mean	SD	3rd Qu.	Max.
1	Total CMRF tests/SA1	1	198	261	135	266.7	117.0	333	965
2	% Female tested /SA1	0	54.2	56.4	4.3	56.0	6.5	58.5	100
3	% Male tested /SA1	0	41.5	43.6	4.3	44.0	6.5	45.8	100
4	Total CMRF tests per person /SA1	1	4.3	4.4	0.4	4.4	0.3	4.6	6

Table 5.4: Summary statistics of the individual CMRF tests across the SA1s in the study region.

	Number of tests /SA1					Average proportion of tested people per age group (in years) /SA1						
CMRFs	**Total tests**	**Mean**	**SD**	**Median**	**IQR**	**18 - 29**	**30 - 39**	**40 - 49**	**50 - 59**	**60 - 69**	**70 - 79**	**80+**
FBSL	193679	202	87.8	199	107.5	10.2	12.1	15.4	19.5	19.8	14.8	8.2
HbA1c	73885	77.6	37.6	73.5	46	4.9	6.8	11.8	19.1	24.1	20.7	12.6
TC	194816	203.1	89.4	201	108.5	7.4	9.8	16.3	20.8	20.9	15.7	9.1
HDL	182237	190	84.2	187	101.5	6.3	9.3	16.5	21.2	21.5	16.2	9.1
ACR	50790	53.6	28.1	50	32.5	3.2	4.7	10.1	19.2	26.7	23.1	13.0
eGFR	244166	253.8	111.6	249.5	128	13.6	11.9	14.9	18.0	18.0	13.9	9.7
BMI	192455	200.5	87.6	197	103.3	11.7	11.5	15.9	19.8	20.3	15.7	5.2

Table 5.5and 5.6 shows the frequencies and relative frequencies of CMRF tests results. Overall, the higher risk frequencies of all CMRFs increased with increasing area-level socioeconomic disadvantage, except for TC which demonstrated an inverse trend.

Table 5.5: Cross-tabulation of individual CMRFs (FBSL, HbA1c, TC and HDL) with the variables in study

CMRFs	FBSL		HbA1c		TC		HDL	
	Total tests	***Higher risk**** **results, n (%)**	**Total tests**	***Higher risk**** **results, n (%)**	**Total tests**	***Higher risk**** **results, n (%)**	**Total tests**	***Higher risk**** **results, n (%)**
Rates	193496	16259(8.4)	73806	7920(10.73)	194624	57506(29.55)	182062	21238(11.67)
Sex								
Male	83603	9279(4.8)	35757	4, 444(6.02)	90950	23503(12.0)	85266	15872(8.72)
Female	109893	6980(3.6)	38049	3, 476(4.71)	103674	34003(17.47)	96796	5366(2.95)
Age(years)								
18 - 29	19747	238(0.1)	3480	250(0.34)	14247	2127(1.09)	11435	1377(0.76)
30 - 39	23515	459(0.2)	4889	293(0.40)	18960	4889(2.51)	16787	2301(1.26)
40 - 49	29424	1265(0.65)	8447	760(1.03)	31395	10719(5.51)	29339	3585(1.97)
50 - 59	37085	2948(1.52)	13510	1507(2.04)	39663	16316(8.38)	37824	4283(2.35)
60 - 69	37962	4670(2.41)	17665	2064(2.80)	40471	13620(7.00)	39134	4227(2.32)
70 - 79	29009	4396(2.27)	15715	1860(2.52)	31186	6748(3.47)	30114	3419(1.88)
80+	16754	2283(1.18)	10100	1186(1.61)	18702	3087(1.59)	17429	2046(1.12)
IRSD								
Most D Q-1	38885	4495(2.32)	17024	2429(3.29)	39347	10631(5.46)	36625	5520(3.03)
Q-2	41545	3757(1.94)	16680	1875(2.54)	41937	12015(6.17)	39050	4901(2.69)
Q-3	39828	3386(1.75)	15376	1585(2.15)	40401	12045(6.19)	37794	4201(2.31)
Q-4	37137	2594(1.34)	13101	1138(1.54)	36865	11163(5.74)	34566	3581(1.97)
Least D Q-5	36101	2027(1.05)	11625	893(1.21)	36074	11652(5.99)	34027	3035(1.67)

FBSL - Fasting Blood Sugar Level; **HbA1c** - Glycated Haemoglobin; **TC** - Total Cholesterol; **HDL** - High Density Lipoprotein; **Most D** – Most Disadvantaged; **Least D** – Least Disadvantaged. *** -** Refer to Table 1 for *high risk* threshold levels of CMRFs.

Table 5.6: Cross-tabulation of individual CMRFs (ACR, eGFR, and Obesity) with the variables in study

CMRFs	ACR		eGFR		Obesity	
	Total tests	***Higher risk*** results, n (%)*	**Total tests**	***Higher risk**** **results, n (%)**	**Total tests**	***Higher risk**** **results, n (%)**
Rates	50737	2046(4.03)	243908	27205(11.15)	192261	64875(33.7)
Sex						
Male	25043	1265(2.49)	108140	12441(5.1)	86853	29585(15.3)
Female	25694	781(1.54)	135768	14764(6.05)	105408	35290(18.3)
Age (years)						
18 - 29	1546	47(0.09)	32961	72(0.03)	23277	4582(2.38)
30 - 39	2278	71(0.14)	29047	105(0.04)	22799	6535(3.40)
40 - 49	4870	108(0.21)	35778	330(0.14)	30401	10595(5.51)
50 - 59	9272	230(0.45)	42695	1112(0.46)	37285	13825(7.19)
60 - 69	13388	412(0.81)	43423	3626(1.49)	38370	15310(7.96)
70 - 79	12337	605(1.19)	34406	8507(3.49)	30074	11324(5.89)
80+	7046	573(1.13)	25598	13453(5.52)	10055	2704(1.41)
IRSD						
Most D Q-1	11915	638(1.26)	49288	7061(2.89)	37476	15365(7.99)
Q-2	11350	485(0.96)	52947	6354(2.61)	40172	14334(7.46)
Q-3	10494	391(0.77)	50816	5917(2.43)	39133	13007(6.77)
Q-4	8732	308(0.61)	46440	4406(1.81)	37370	11766(6.12)
Least D Q-5	8246	224(0.44)	44417	3467(1.42)	38110	10403(5.41)

ACR - Albumin Creatinine Ratio; **eGFR** - estimated Glomerular Filtration Rate; **BMI** - Body Mass Index; Most **D** – Most Disadvantaged; **Least D** – Least Disadvantaged. *** -** Refer to Table 1 for *higher risk* threshold levels.

Single and multilevel models for each of the CMRFs analysed in this study are presented in Table 5.7-5.13. After adjusting for the covariates, all seven CMRFs were found to be significantly associated with area-level IRSD in the study region. For all but one variable the associations were positive (i.e. increased with area-level disadvantage). TC was the exception; being inversely associated with area-level disadvantage, with the most disadvantaged quintile (Q1) displaying the lowest odds for higher risk test results. Among the covariates, there was no significant association between gender and higher risk test results of eGFR or BMI. It was also noted that the odds of higher risk eGFR tests results accelerated with increasing age group, and the 80+ age group demonstrated a very high odds of being identified with a higher risk eGFR tests result in the study region.

Table 5.7: Single and multilevel logistic regression model summaries for high FBSL (FBSL ≥7.0 mmol/l)

	Single level model	Multilevel models			
	Model 0	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Significance (LRT)	***	***	***	***	***
High FBSL	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.01 (0.01 - 0.01)	0.09 (0.09 - 0.09)	0.01 (0.01-0.01)	0.06 (0.06-0.06)	0.01 (0.01-0.01)
Sex: Female	Reference		-		
Male	1.62 (1.56 - 1.67)		1.63 (1.56-1.7)		1.63 (1.58-1.7)
Age:18—29	Reference		-		1.64 (1.41-1.9)
30—39	1.60 (1.36 - 1.87)		1.63 (1.39-1.9)		3.58 (3.12-4.1)
40—49	3.41 (2.97 - 3.93)		3.53 (3.07-4.1)		6.81 (5.98-7.8)
50—59	6.48 (5.68 - 7.42)		6.77 (5.93-7.7)		11.05 (9.71-12.6)
60—69	10.48 (9.21 -11.98)		11.07 (9.72-12.6)		13.74 (12.07-15.6)
70—79	13.35 (11.73 -15.27)		13.93 (12.22-15.9)		12.02 (10.52-13.7)
80+	12.01 (10.51 - 13.78)		12.33 (10.79-14.1)		
IRSD: Q-5				Reference	-
Q-4				1.27 (1.18-1.36)	1.27 (1.18-1.37)
Q-3				1.58 (1.47-1.69)	1.49 (1.39-1.61)
Q-2				1.68 (1.57-1.80)	1.62 (1.50-1.74)
Most D Q-1				2.20 (2.06-2.36)	2.11 (1.96-2.26)
AIC	103645	111022.8	103066.2	110552.5	102689.6
Variance		0.101	0.103	0.034	0.044
PCV		-	+1.88%	-66.41%	-56.33%
ICC(%)		3.00	3.00	1.0	1.3
MOR		1.35	1.36	1.19	1.22
AUC	0.70	0.61	0.73	0.60	0.72
AUC change†			+0.03		
Proportional variance explained by IRSD : 57.14%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0; Model 0— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

Table 5.8: Single and multilevel logistic regression model summaries for high HbA1c (HbA1c > 7.5%)

	Single level model	Multilevel models			
	Model 0	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Significance (LRT)	***	***	***	***	***
High HbA1c	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.07 (0.06 - 0.08)	0.12 (0.12-0.12)	0.07 (0.06-0.07)	0.08(0.08-0.09)	0.05 (0.04-0.05)
Sex: Female	Reference		-		-
Male	1.37 (1.31 - 1.43)		1.38 (1.32-1.45)		1.39 (1.32-1.45)
Age:18—29	Reference		-		-
30—39	0.81 (0.68 - 0.96)		0.81(0.68-0.96)		0.81(0.68-0.96)
40—49	1.22 (1.06 - 1.42)		1.24(1.07-1.44)		1.26(1.08-1.46)
50—59	1.53 (1.34 - 1.77)		1.56(1.36-1.80)		1.57(1.36-1.81)
60—69	1.61 (1.40 - 1.85)		1.64(1.43-1.88)		1.64(1.43-1.88)
70—79	1.63 (1.42 - 1.87)		1.64(1.42-1.88)		1.62(1.41-1.86)
80+	1.64 (1.43 - 1.90)		1.63(1.41-1.88)		1.60(1.39-1.85)
IRSD : Q-5				Reference	-
Q-4				0.08(0.08-0.09)	1.15(1.04-1.28)
Q-3				1.14(1.03-1.27)	1.39(1.26-1.54)
Q-2				1.40(1.27-1.54)	1.55(1.41-1.71)
Most D Q-1				1.54(1.40-1.69)	2.02(1.84-2.22)
AIC	49897	50114.5	49690.2	49875.3	49453.3
Variance		0.103	0.106	0.047	0.049
PCV		-	+3.02 %	-54.82%	-51.91%
ICC (%)		3.0	3.1	1.4	1.5
MOR		1.36	1.36	1.23	1.24
AUC	0.56	0.63	0.64	0.61	0.63
AUC change†			+0.08		
Proportional variance explained by IRSD : 53.31%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0; Model 0— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—Model 1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

Table 5.9: Single and multilevel logistic regression model summaries for high TC (TC ≥ 5.5 mmol/l)

	Single level model	Multilevel models			
	Model 0	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Significance (LRT)	***	***	***	***	***
High TC	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.20 (0.19-0.21)	0.42 (0.42-0.42)	0.20 (0.19-0.21)	0.08(0.08-0.09)	0.22 (0.21-0.23)
Sex: Female	Reference		-		-
Male	0.69 (0.68-0.71)		0.69 (0.68-0.71)		0.69 (0.68-0.71)
Age:18—29	Reference		-		-
30—39	2.01 (1.90-2.13)		2.02 (1.91-2.14)		2.01 (1.90-2.13)
40—49	3.01 (2.86-3.17)		3.01 (2.86-3.17)		3.00 (2.85-3.16)
50—59	4.09 (3.89-4.30)		4.08 (3.88-4.29)		4.07 (3.87-4.28)
60—69	2.97 (2.83-3.13)		2.95 (2.80-3.10)		2.95 (2.80-3.10)
70—79	1.61 (1.53-1.70)		1.60 (1.52-1.69)		1.61 (1.52-1.70)
80+	1.14 (1.07-1.21)		1.13 (1.07-1.20)		1.14 (1.07-1.21)
IRSD : Q-5				Reference	-
Q-4				0.91 (0.87-0.95)	0.94 (0.90-0.98)
Q-3				0.88 (0.85-0.92)	0.94 (0.90-0.98)
Q-2				0.84 (0.80-0.87)	0.90 (0.87-0.94)
Most D Q-1				0.77 (0.74-0.81)	0.84 (0.81-0.88)
AIC	227464	235931.6	227254.6	235795.4	227199.2
Variance		0.026	0.020	0.018	0.017
PCV		-	⁻ 21.76%	⁻27.81%	⁻33.27%
ICC (%)		0.8	0.6	0.6	0.5
MOR		1.16	1.14	1.14	1.13
AUC	0.63	0.56	0.64	0.56	0.64
AUC change†			⁺0.01		
Proportional variance explained by IRSD : 14.71%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0; Model 0— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

Table 5.10: Single and multilevel logistic regression model summaries for low HDL (HDL < 1 mmol/l)

	Single level model	Multilevel models			
	Model 0	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Significance (LRT)	***	***	***	***	***
Low HDL	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.06 (0.06-0.07)	0.13(0.13-0.13)	0.06 (0.06-0.07)	0.10(0.09-0.10)	0.05(0.04-0.05)
Sex: Female	Reference		-		-
Male	3.92 (3.80-4.05)		3.98(3.85-4.11)		3.98(3.85-4.11)
Age:18—29	Reference		-		-
30—39	1.11 (1.03-1.20)		1.11(1.03-1.20)		1.12(1.04-1.21)
40—49	0.97 (0.91-1.04)		0.99(0.92-1.05)		1.00(0.93-1.07)
50—59	0.87 (0.81-0.93)		0.88(0.82-0.94)		0.89(0.83-0.95)
60—69	0.81 (0.76-0.87)		0.82(0.77-0.88)		0.82(0.77-0.88)
70—79	0.85 (0.80-0.91)		0.86(0.80-0.92)		0.85(0.79-0.91)
80+	0.94 (0.87-1.01)		0.93(0.86-1.00)		0.91(0.85-0.98)
IRSD : Q-5				Reference	-
Q-4				1.18(1.11-1.26)	1.20(1.13-1.28)
Q-3				1.29(1.21-1.37)	1.32(1.24-1.41)
Q-2				1.48(1.39-1.57)	1.51(1.42-1.61)
Most D Q-1				1.81(1.71-1.92)	1.90(1.78-2.02)
AIC	123277	130649.7	122700.0	130294.3	122328.3
Variance		0.071	0.081	0.030	0.034
PCV		-	$^{+}$15.25%	$^{-}$58.05%	$^{-}$51.37%
ICC (%)		2.1	2.4	0.9	1.0
MOR		1.29	1.31	1.18	1.19
AUC	0.67	0.60	0.71	0.59	0.70
AUC change†			$^{+}$0.04		
Proportional variance explained by IRSD : 57.8%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0; Model 0— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

Table 5.11: Single and multilevel logistic regression model summaries for high ACR (ACR ≥ 30 mcg/L to mg/l)

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
High ACR	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.02 (0.02-0.03)	0.04 (0.04-0.04)	0.02 (0.02-0.03)	0.03(0.02-0.03)	0.02(0.01-0.02)
Sex: Female	Reference	-	-		-
Male	1.74 (1.59-1.91)		1.75 (1.59-1.92)		1.76(1.60-1.93)
Age:18—29	Reference		-		-
30—39	0.99 (0.69-1.45)		1.00 (0.69-1.45)		0.99(0.68-1.44)
40—49	0.68 (0.49-0.98)		0.69 (0.49-0.97)		0.70(0.49-0.98)
50—59	0.76 (0.56-1.06)		0.77 (0.56-1.05)		0.77(0.56-1.05)
60—69	0.95 (0.70-1.30)		0.95 (0.70-1.30)		0.95(0.70-1.29)
70—79	1.54 (1.15-2.11)		1.55 (1.14-2.09)		1.52(1.12-2.05)
80+	2.73 (2.04-3.74)		2.74 (2.02-3.71)		2.65(1.96-3.59)
IRSD : Q-5				Reference	-
Q-4				1.31(1.10-1.57)	1.25(1.04-1.50)
Q-3				1.39(1.16-1.65)	1.27(1.06-1.51)
Q-2				1.61(1.36-1.90)	1.45(1.23-1.72)
Most D Q-1				2.02(1.72-2.38)	1.84(1.56-2.16)
AIC	16596	17130.0	16585.2	17053.0	16527.2
Variance		0.092	0.073	0.044	0.036
PCV		-	-20.53%	-52.88%	-61.19%
ICC (%)		2.7	2.2	1.3	1.1
MOR		1.34	1.30	1.22	1.20
AUC	0.65	0.70	0.69	0.62	0.67
AUC change†			+0.04		
Proportional variance explained by IRSD : 51.17%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0; Model 0— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

Table 5.12: Single and multilevel logistic regression model summaries for low eGFR (eGFR < 60 mL / min / 1.73m^2)

	Single level model	Multilevel models			
	Model 0	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Significance (LRT)	***	***	***	***	***
Low eGFR	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.00 (0.00-0.00)	0.13(0.12-0.13)	0.00(0.00-0.00)	0.08(0.07-0.08)	0.00(0.00-0.00)
Sex: Female	Reference		-		-
Male	0.98 (0.95-1.01)		0.98(0.95-1.01)		0.98(0.95-1.01)
Age:18—29	Reference		-		-
30—39	1.66 (1.23-2.25)		1.66(1.24-2.20)		1.66(1.24-2.22)
40—49	4.26 (3.32-5.54)		4.26(3.34-5.42)		4.30(3.36-5.49)
50—59	12.24 (9.72-15.68)		12.26(9.78-15.35)		12.32(9.80-15.47)
60—69	41.72 (33.30-53.19)		41.81(33.55-52.1)		41.86(33.49-52.31)
70—79	150.44 (120.24-191.57)		150.66(121-187.6)		149.53(120-186.6)
80+	506.80 (405.05-645.35)		509.18(409-633.9)		501.47(401.7-626)
IRSD : Q-5				Reference	-
Q-4				1.23(1.12-1.35)	1.09(1.03-1.16)
Q-3				1.59(1.45-1.74)	1.19(1.13-1.26)
Q-2				1.65(1.51-1.81)	1.22(1.15-1.29)
Most D Q-1				1.97(1.80-2.15)	1.38(1.31-1.46)
AIC	115340	167164.8	115257.1	166930.0	115125.7
Variance		0.189	0.024	0.138	0.014
PCV		-	ˉ 87.26%	ˉ26.84%	ˉ92.79%
ICC (%)		5.4	0.7	4.0	0.4
MOR		1.51	1.16	1.43	1.12
AUC	0.88	0.64	0.89	0.63	0.88
AUC change†			ˉ0.01		
Proportional variance explained by IRSD : 41.75%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0; Model 0— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

Table 5.13: Single and multilevel logistic regression model summaries for obesity (BMI ≥ 30 kg/m²)

	Single level model	Multilevel models			
	Model 0	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Significance (LRT)	***	***	***	***	***
Obesity	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.25 (0.24-0.25)	0.51(0.50-0.51)	0.25(0.24-0.26)	0.37(0.35-0.39)	0.18(0.17-0.19)
Sex: Female	Reference		-		-
Male	0.99 (0.97-1.00)		0.99(0.97-1.01)		0.99(0.97-1.01)
Age:18—29	Reference		-		-
30—39	1.64 (1.57-1.71)		1.63(1.56-1.71)		1.64(1.57-1.71)
40—49	2.18 (2.10-2.27)		2.20(2.11-2.29)		2.21(2.12-2.30)
50—59	2.41 (2.32-2.50)		2.44(2.34-2.53)		2.44(2.35-2.54)
60—69	2.71 (2.61-2.82)		2.73(2.63-2.84)		2.74(2.63-2.84)
70—79	2.47 (2.37-2.57)		2.44(2.34-2.54)		2.43(2.33-2.53)
80+	1.50 (1.42-1.59)		1.46(1.38-1.55)		1.45(1.37-1.53)
IRSD : Q-5				Reference	-
Q-4				1.25(1.17-1.33)	1.26(1.18-1.34)
Q-3				1.37(1.29-1.46)	1.38(1.30-1.47)
Q-2				1.51(1.42-1.61)	1.54(1.44-1.64)
Most D Q-1				1.90(1.79-2.03)	1.94(1.83-2.07)
AIC	242064	242793.2	239122.6	242443.7	238748.4
Variance		0.115	0.117	0.071	0.069
PCV		-	⁺1.48%	⁻38.76%	⁻40.30%
ICC (%)		3.4	3.4	2.1	2.0
MOR		1.38	1.39	1.29	1.28
AUC	0.56	0.60	0.63	0.60	0.62
AUC change†			⁺0.07	⁻0.03	⁻0.01
Proportional variance explained by IRSD : 41.06%					

*** - p<0.001; † - Change in Model 2 in relation to Model 0 and the rest in relation to Model 2; Model 0— Single level model adjusted for age + sex; Model 1— Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3.

The overall comparisons of model random effects are presented in Table 5.14. Reductions in the AIC values were observed among all CMRFs from the null model (M1) to the final model (M4) indicating a better fit for the final models. In the unadjusted null models, higher risk test results of eGFR demonstrated the most area-level variance (0.189) and TC the least (0.026). Adjusting the CMRFs for age and sex initially increased the τ^2 of M2 for FBSL (PCV = +1.88%), HbA1c (PCV = +3.02%), HDL (PCV = +15.25%) and BMI (PCV = +1.48%). The τ^2 was reduced in the final model among all CMRFs compared with the null models.

Table 5.14: Summary model fit values and comparison of the models

		FBSL	HbA1c	TC	HDL	ACR	eGFR	Obesity
Model 1	**Null Model**							
	AIC	111022.8	50114.5	235931.6	130649.7	17130.0	167164.8	242793.2
	τ^2	0.101	0.103	0.026	0.071	0.092	0.189	0.115
	ICC (%)	3.0	3.0	0.8	2.1	2.7	5.4	3.4
	MOR	1.35	1.36	1.16	1.29	1.34	1.51	1.38
Model 2	**Sex + Age Adjusted Model**							
	AIC	103066.2	49690.2	227254.6	122700.0	16585.2	115257.1	239122.6
	τ^2	0.103	0.106	0.020	0.081	0.073	0.024	0.117
	ICC (%)	3.0	3.1	0.6	2.4	2.2	0.7	3.4
	MOR	1.36	1.36	1.14	1.31	1.30	1.16	1.39
	PCV	+ 1.88 %	+ 3.02 %	- 21.76%	+15.25%	-20.53%	- 87.26%	+1.48%
Model 3	**IRSD Adjusted Model**							
	AIC	110552.5	49875.3	235795.4	130294.3	17053.0	166930.0	242443.7
	τ^2	0.034	0.047	0.018	0.030	0.044	0.138	0.071
	ICC (%)	1.0	1.4	0.6	0.9	1.3	4.0	2.1
	MOR	1.19	1.23	1.14	1.18	1.22	1.43	1.29
	PCV	-66.41%	-54.82%	-27.81%	-58.05%	-52.88%	-26.84%	-38.76%
Model 4	**Sex + Age + IRSD Adjusted Model**							
	AIC	102689.6	49453.3	227199.2	122328.3	16527.2	115125.7	238748.4
	τ^2	0.044	0.049	0.017	0.034	0.036	0.014	0.069
	ICC (%)	1.3	1.5	0.5	1.0	1.1	0.4	2.0
	MOR	1.22	1.24	1.13	1.19	1.20	1.12	1.28
	PCV	-56.33%	-51.91%	-33.27%	-51.37%	-61.19%	-92.79%	-40.30%

AIC - Akaike Information Criterion; τ^2 – variance; ICC - Intra-cluster Correlation Coefficients; MOR - Median Odds Ratio; PCV - Proportional Change in Variance; FBSL - Fasting Blood Sugar Level; HbA1c - Glycated Haemoglobin ; TC - Total Cholesterol; HDL - High Density Lipoprotein; ACR - Albumin Creatinine Ratio; eGFR - estimated Glomerular Filtration Rate.

The Akaike Information Criterion (AIC) was used to evaluate model fit. The derived multilevel models were compared for: area-level variance (τ^2) at SA1 (level 2) level; proportional change in variance (PCV); Intra-cluster Correlation Coefficients (ICC); Median Odds Ratios (MORs); Area under the receiver operating characteristic (AUC) curve; and the change in AUC.

The ICCs of the unadjusted models ranged between 0.8% in high TC to 5.4% in low eGFR. Inclusion of IRSD after adjusting for age and sex had reduced the ICCs of all CMRFs in the final models, which ranged between 0.4% in low eGFR to 2.0% in obesity test results. The ICCs of the final models were low and suggest very limited area-level contextual effects. The AUC changes in model 2 and MORs of the final model support these findings.

Figure 5.2 provides a comparison of the ROC curves of the fitted models. Model 4s (age + sex + IRSD adjusted models) and models 3s (IRSD adjusted models) were chosen for the ROC curve plotting for comparative purpose. The predicted outcomes in the CMRFs plots are for the reference individual, i.e., individuals residing in the least disadvantaged areas (model 3) + female + age group 18–29 years (Model4). A model curve closer to the top left corner of the subfigures indicate a better predictive accuracy of the model. The single measure summary of the ROC curves, AUCs of the final models ranged 0.62 -0.88. The highest AUC value was observed for the final model of low eGFR. The AUC changes of model 2s in relation to M0s ranged 0.01 – 0.08, which reconfirm the contextual findings of ICCs that the general contextual effects observed in the models were minimal.

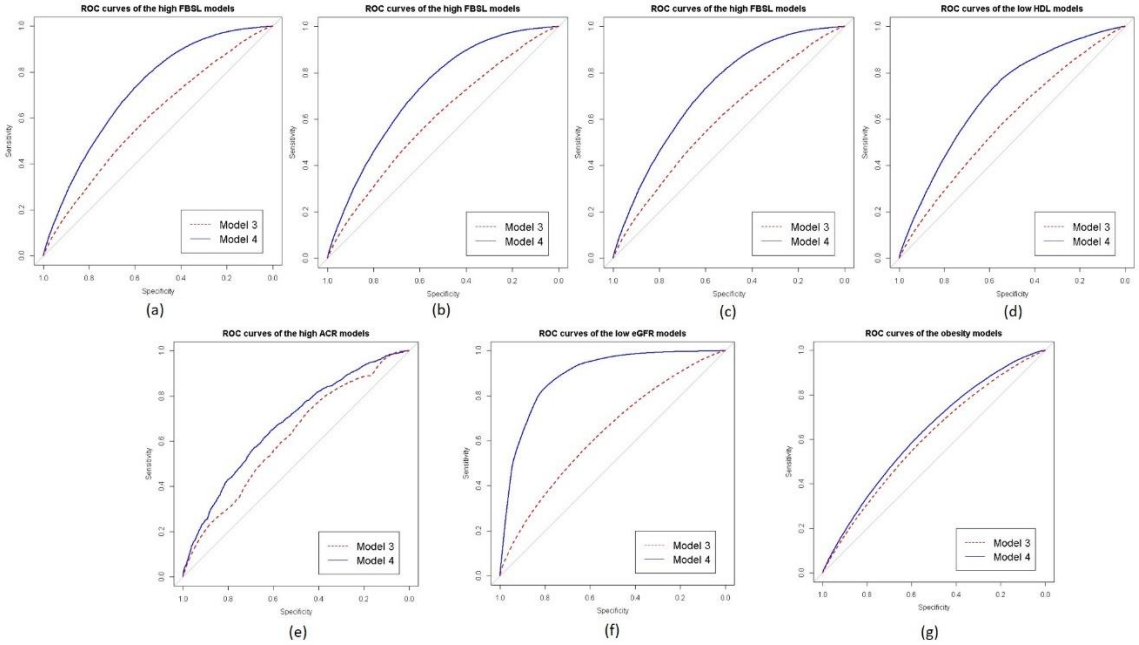

Figure 5.2: ROC curves of the fitted models (Model 3s and Model4s) of CMRFs for comparison: (a) FBSL Models; (b) HbA1c Models; (c) TC Models; (d) HDL Models; (e) ACR Models; (f) FBSL Models; (g) Obesity Models; Model 3s— CMRFs adjusted only for IRSD quintiles of areas; Model 4s—Final models of CMRFs adjusted for age + sex + IRSD quintiles of area.

The proportions of the geographic variance in CMRFs contributed by IRSD were estimated through the PCV between M2 and M4. Adjusting the models for IRSD and individual-level variables explained a maximum 92.79% of the variance expressed by the null model of eGFR, reducing the ICC from 5.4% to 0.4%. The changes were least among the adjusted models of TC, with a marginal reduction of ICC from 0.8% to 0.5%. Thus, in the final models, the proportional reduction in variance was the largest for eGFR (PCV = 92.79%) and the least for TC (PCV = 33.27%).

The identified specific contribution of IRSD in the geographic variation of CMRF was the highest among the geographic variance of *higher risk* findings of HDL tests (57.8%), which was closely followed by FBSL (57.14%); HbA1c (53.31%); and ACR (51.17%) test results. The contribution of IRSD was comparatively lower among the geographic variance of the *higher risk* findings of eGFR (41.75%); BMI (41.06%); and

TC (14.71%) test results, though not the least. Even though these specific proportions are large, it should be noted that it actually explained a lot of very little (i.e, variance of 0.01 to 0.07)

5.6 Discussion

The study reports on the contextual influence of areas on *higher risk* CMRFs distribution and quantifies the specific proportion of geographic variance explained by IRSD. The work adds to the very few studies which consider multiple CMRF variables from the same region, or which are based on population derived data over extended years; [6, 15, 22, 29-31] and reports on both single and multilevel analyses.[38, 57] The results present both the measures of association and area-level variance based on multilevel logistic regression analyses.[36] The findings of the study add to the existing evidence and discussion regarding the relevance of individual versus area-level interventions for the prevention and control of CMRFs.

We found consistent evidence for the association between area-level disadvantage and seven CMRFs among adult health service using residents of the Illawarra-Shoalhaven region in NSW Australia. In adjusted models, the odds of a higher risk finding increased with increase in area-level disadvantage among all CMRFs excepting TC, which showed an inverse pattern of association with increase in area-level disadvantage. Thus, in the final models we observed that, over and above individual age and sex, living in a disadvantaged neighbourhood proportionally increased the individual-level probability of being identified with a higher risk CMRF. The findings highlight the importance of including of area-level variables into health risk analyses.

The ICCs of CMRFs in all the models were comparatively small in all the models (Table 5.14). In the fully adjusted models, the ICCs were further reduced and ranged between 0.4% and 2.0% in low eGFR and BMI respectively. As per the interpretation framework proposed by Merlo et al (2019), an ICC value less than 10% is indicative of very little geographic difference.[58]However, this has to be interpreted along with the traditional geographic comparisons such as the proportion of the individuals who are affected with higher risk CMRF outcomes. Therefore, a small geographic difference with uniformly higher, medium, or lower proportion of affected individuals indicates homogeneity of the higher risk CMRF findings within their geographic units.[58] Such a situation would call for balanced universal approaches to prevent and control the higher risk CMRFs, with a proportional focus to the need and disadvantage level of affected populations.[59, 60] However, it is also worth noting that when the exposure to an agent is homogenic in a community, the traditional epidemiological methods are not very helpful in identifying markers of susceptibility.[61]

Our results confirm, and are comparable with, associations between area-level disadvantage and CMRFs reported in previous studies, [3-23, 29-31] and extends their findings. The results primarily confirm the geographic variation of CMRFs and associations with area level disadvantage, as reported in previous studies. Further, the study provides means to compare the observed associations of a range of CMRFs concurrently. The study extends on previous reports by differentiating the individual and area-level contributors to the exhibited geographic variance of CMRFs. And most importantly, the specific contributions of IRSD on the geographic variance of multiple CMRFs were identified, which is unique in the literature and informative for health care service commissioning.

The TC test results often stood apart from the major findings of this study, demonstrating inverse associations with IRSD. This was in contract to the HDL associations, even though both are components of the lipid profile in an individual. This raises the possibility of a medication effect on TC in these areas, where the lipid lowering drugs have a less consistent effect in raising HDL than in lowering TC.[62] Other factors associated with the higher risk HDL test results may include uncontrolled diabetes[63], smoking[64], sedentary life style[65-67], obesity[65], and poor diet quality[68, 69]. However the reason for the inverse association demonstrated by TC test results are not clearly established within the current study results and requires further research to explore possible individual and area-level contributions.

The study has to be considered within its limitations. Primarily, the cross sectional nature of analyses adopted in this study do not yield support for any causal relationships. In addition, the non-linear and time varying effects of covariates analysed in this study restrict generalisability of their findings. Secondly, the IRSD quintiles included as the key explanatory variable represent relative disadvantage in an area and have limitations intrinsic to aggregate measures. Thirdly, it should be noted that the data used in this study are extracted from people already utilising the health care service facilities in the area. This is likely to create an overestimation of the odds ratios in the study findings than if it would have been in the actual general population. Fourthly, the readers should be mindful that the variance reported in this study are attributable to 1) individual-level factors (age, sex), 2) area-level contextual influences (IRSD), and 3) other individual and area-level characteristics not considered in this study. However, further individual-level data analyses are not possible with this study's dataset as the de-identification process precludes the inclusion of any further individual level data. Other individual and area-level factors not considered in this study could include individual-level SES [70], smoking, hypertension, diet, and physical activity; type of neighbourhood food outlets [71-74], poor physical activity resources [75-77], and service availability. Accounting for these factors might have further reduced the estimates of the relationship between area-level disadvantage and CMRFs. However it should also be noted that acquisition of additional individual-level data wasn't possible with the de-identified study data set; and more importantly, unraveling the underlying causative structure of the derived estimates were not the primary intention of this study than obtaining age and sex standardised effect of ASED on CMRFs to inform evidence for targeted area-level health care service and resource allocation planning. Finally, the standard multilevel logistic regression modelling methods adopted in this study would not be able to account for the autocorrelation of the area-level residuals (if any) of the models. Expected shortcomings due to this could be an overestimation of random effects in our models.[78] However, any such effects are not expected to be substantial in our results as the random effect estimates are already quite small. While acknowledging this limitation, we believe the effects of this are not critical in our results. Hybrid models which provide more precise estimates of random effects are becoming increasingly available with advances in computational technologies.[79] However, they would not be directly applicable to our data sets, mainly due to the non-availability of location specific data at individual-level in our study data.

Notwithstanding these limitations, the study is unique in that it analysed a range of CMRFs across a widely dispersed population and included both rural and urban residents. In addition, the study used six years (year 2012 –2017) of CMRF tests data from the region in the hierarchical multilevel analyses. The findings of the study indicate that those residing in the most disadvantaged areas are more likely to be identified

with higher risk CMRFs than those in lower disadvantage areas. However, the low ICC and MOR values of the area-level models do not provide support for contextual approaches. Rather, the findings of the study support a proportionate universalism approach in which health resources are made universally available but proportional to the need and disadvantage level of the affected population.[59, 60]

5.7 Conclusion

The study demonstrates that in the Illawarra Shoalhaven region of Australia, people residing in socioeconomically disadvantaged areas have a higher probability of being identified with higher risk CMRFs across a range of factors. The low general contextual effects of the areas suggest for universal intervention for the prevention and control of CMRFs in this study region, but proportional to the need and disadvantage level. The patterns were consistent across the six CMRFs analysed in this study; and comparable with similar studies reported nationally and globally. Based on our findings, we recommend further area-level research to discern the role of other contextual factors not analysed in this study especially the area-level access to health care services to determine its existing role and adequacy; and evidence based universal interventions for the prevention and control of CMRFs but proportionate to the priority level of the populations based on area-level disadvantage.

References

1. Toms, R., et al., *Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature.* International journal of health geographics, 2019. **18**(1): p. 1.
2. Toms, R., et al., *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia.* PloS one, 2019. **14**(10): p. e0223179.
3. Alkerwi, A., et al., *Geographic variations in cardiometabolic risk factors in Luxembourg.* International journal of environmental research and public health, 2017. **14**(6): p. 648.
4. Astell-Burt, T., et al., *Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia.* Diabetes research and clinical practice, 2014. **106**(3): p. e68-e73.
5. Maier, W., et al., *Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update'GEDA 2009 and 2010.* PloS one, 2014. **9**(2).
6. Roux, A.V.D., D.R. Jacobs, and C.I. Kiefe, *Neighborhood characteristics and components of the insulin resistance syndrome in young adults: the coronary artery risk development in young adults (CARDIA) study.* Diabetes care, 2002. **25**(11): p. 1976-1982.
7. Valdés, S., et al., *Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@ bet. es study.* Revista Española de Cardiología (English Edition), 2014. **67**(6): p. 442-448.
8. Andersen, A., et al., *Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study.* Diabetic medicine, 2008. **25**(12): p. 1462-1468.
9. Barker, L.E., et al., *Geographic distribution of diagnosed diabetes in the US: a diabetes belt.* American journal of preventive medicine, 2011. **40**(4): p. 434-439.
10. Cubbin, C., et al., *Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects.* Scandinavian journal of public health, 2006: p. 228-237.
11. Dragano, N., et al., *Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany.* BMC Public Health, 2007. **7**(1): p. 255.
12. Inoue, Y., et al., *Neighborhood characteristics and cardiovascular risk among older people in Japan: findings from the JAGES project.* PloS one, 2016. **11**(10).
13. Lawlor, D., et al., *Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study.* Journal of Epidemiology & Community Health, 2003. **57**(2): p. 134-140.
14. Lawlor, D.A., et al., *Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study.* American journal of public health, 2005. **95**(1): p. 91-97.
15. Naimi, A.I., et al., *Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area.* International journal of environmental research and public health, 2009. **6**(12): p. 3082-3096.

16. Silhol, R., et al., *Investigating the spatial variability in incidence of coronary heart disease in the Gazel cohort: the impact of area socioeconomic position and mediating role of risk factors.* Journal of Epidemiology & Community Health, 2011. **65**(2): p. 137-143.
17. Bonney, A., et al., *Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users–a cross-sectional study.* PloS one, 2015. **10**(8).
18. Congdon, P., *Estimating diabetes prevalence by small area in England.* Journal of Public Health, 2006. **28**(1): p. 71-81.
19. Mujahid, M.S., et al., *Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics.* Obesity research, 2005. **13**(8): p. 1412-1421.
20. Paquet, C., et al., *Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia.* International journal of environmental research and public health, 2016. **13**(5): p. 519.
21. Sundquist, K., et al., *Neighborhood walkability, deprivation and incidence of type 2 diabetes: a population-based study on 512,061 Swedish adults.* Health & place, 2015. **31**: p. 24-30.
22. Unger, E., et al., *Association of neighborhood characteristics with cardiovascular health in the multi-ethnic study of atherosclerosis.* Circulation: Cardiovascular Quality and Outcomes, 2014. **7**(4): p. 524-531.
23. Zhou, M., et al., *Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults.* Diabetes care, 2015. **38**(1): p. 72-81.
24. World Health Organisation. *The top 10 causes of death* 2017 [cited 2017; Available from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.
25. Australian Institute of Health and Welfare. *Indicators of socioeconomic inequalities in cardiovascular disease, diabetes and chronic kidney disease*. 2019; Available from: https://www.aihw.gov.au/getmedia/01c5bb07-592e-432e-9fba-d242e0f7e27e/aihw-cdk-12.pdf.aspx?inline=true.
26. Australian Institute of Health and Welfare. *Australian burden of disease study : impact and causes of illness and death in Australia 2011*. 2016; Available from: https://www.aihw.gov.au/getmedia/d4df9251-c4b6-452f-a877-8370b6124219/19663.pdf.aspx?inline=true.
27. Australian Institute of Health and Welfare. *Australia's Health 2014*. 2014; Available from: https://www.aihw.gov.au/reports/australias-health/australias-health-2014/contents/table-of-contents.
28. Australian Institute of Health and Welfare. *Cardiovascular disease, diabetes and chronic kidney disease: Australian facts: morbidity—hospital care.* 2017; Available from: https://www.aihw.gov.au/reports/heart-stroke-vascular-disease/cardiovascular-diabetes-chronic-kidney-morbidity/contents/table-of-contents.
29. Clark, C.R., et al., *Neighborhood disadvantage, neighborhood safety and cardiometabolic risk factors in African Americans: biosocial associations in the Jackson Heart study.* PloS one, 2013. **8**(5).
30. Gabert, R., et al., *Identifying high-risk neighborhoods using electronic medical records: a population-based approach for targeting diabetes prevention and treatment interventions.* PLoS One, 2016. **11**(7).
31. Keita, A.D., et al., *Associations of neighborhood area level deprivation with the metabolic syndrome and inflammation among middle-and older-age adults.* BMC Public Health, 2014. **14**(1): p. 1319.
32. Laraia, B.A., et al., *Place matters: neighborhood deprivation and cardiometabolic risk factors in the Diabetes Study of Northern California (DISTANCE).* Social science & medicine, 2012. **74**(7): p. 1082-1090.

33. Cox, M., et al., *Locality deprivation and Type 2 diabetes incidence: a local test of relative inequalities.* Social science & medicine, 2007. **65**(9): p. 1953-1964.
34. Mobley, L.R., et al., *Environment, obesity, and cardiovascular disease risk in low-income women.* American journal of preventive medicine, 2006. **30**(4): p. 327-332. e1.
35. Riva, M., L. Gauvin, and T.A. Barnett, *Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998.* Journal of Epidemiology & Community Health, 2007. **61**(10): p. 853-861.
36. Merlo, J., et al., *An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health.* PloS one, 2016. **11**(4).
37. Merlo, J., et al., *Diastolic blood pressure and area of residence: multilevel versus ecological analysis of social inequity.* Journal of Epidemiology & Community Health, 2001. **55**(11): p. 791-798.
38. Merlo, J., et al., *Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia, Spain.* Social science & medicine, 2012. **75**(8): p. 1477-1487.
39. Australian Bureau of Statistics. *2011 Census data* [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280.
40. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure :STATISTICAL AREA LEVEL 1 (SA1).* 2016; Available from: https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+(ASGS).
41. Ghosh, A., et al., *Using data from patient interactions in primary care for population level chronic disease surveillance: The Sentinel Practices Data Sourcing (SPDS) project.* BMC Public Health, 2014. **14**(1): p. 557.
42. Ghosh A, M.K., Marshall K. . *Illawarra-Shoalhaven Medicare Local - Population Health Profile: 2013.* 2013; Available from: https://www.gph.org.au/assets/Main-Site/Uploads/Resources/Improving-population-health/ISML-Population-Health-Profile-2013-FINAL.pdf.
43. The Royal Australian College of General Practitioners & Diabetes Australia. *General Practice Management of Type 2 Diabetes 2016-2018.* . The Royal Australian College of General Practitioners 2016; Available from: Doi: 10.1007/s00125-010-2011-6.
44. Australian Bureau of Statistics. *Australian Health Survey: Biomedical Results for Chronic Diseases,2011-12* . 2013; Available from: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0052011-12.
45. National heart foundation of Australia. *Lipid management profile for health professionals.* . Available from: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management.
46. National Kidney foundation(USA). *Albumin creatinine Ratio (ACR).* 2018; Available from: https://www.kidney.org/kidneydisease/siemens_hcp_acr. .
47. World Health Organization. *Obesity : Preventing and managing the global epidemic: Technical Report Series. WHO Technical Report Series, no. 894.* 2000; Available from: doi:ISBN 92 4 120894 5.
48. Australian Bureau of Statistics. *Main Features - IRSD.*; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011.

49. Goldstein, H., W. Browne, and J. Rasbash, *Partitioning variation in multilevel models.* Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences, 2002. **1**(4): p. 223-231.
50. Wagner, P. and J. Merlo, *Measures of discriminatory accuracy in multilevel analysis.* Eur J Epidemiol, 2013. **28**(1): p. 135.
51. Pepe, M.S., et al., *Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker.* American journal of epidemiology, 2004. **159**(9): p. 882-890.
52. R Core Team. *R: A language and environment for statistical computing.R Foundation for Statistical Computing,Vienna, Austria*. 2018; Available from: https://www.R-project.org/.
53. Bates, D., et al., *Fitting linear mixed-effects models using lme4.* arXiv preprint arXiv:1406.5823, 2014.
54. Zeileis, A. and T. Hothorn, *Diagnostic checking in regression relationships. R News 2: 7–10.* Available at (accessed August 2011). http://CRAN. R-project. org/doc/Rnews/(http://CRAN. R-project. org/doc/Rnews/), 2002.
55. Robin, X., et al., *pROC: an open-source package for R and S+ to analyze and compare ROC curves.* BMC bioinformatics, 2011. **12**(1): p. 1-8.
56. AUSTRALIAN BUREAU OF STATISTICS. *Technical Paper: Socio-Economic Indexes for Areas (SEIFA)*. 2011; Available from: https://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/22CEDA8038AF7A0DCA257B3B00116E34/$File/2033.0.55.001%20seifa%202011%20technical%20paper.pdf.
57. Merlo, J., et al., *Individual and collective bodies: using measures of variance and association in contextual epidemiology.* Journal of Epidemiology & Community Health, 2009. **63**(12): p. 1043-1048.
58. Merlo, J., P. Wagner, and G. Leckie, *A simple multilevel approach for analysing geographical inequalities in public health reports: The case of municipality differences in obesity.* Health & place, 2019. **58**: p. 102145.
59. Lu, D. and I. Tyler, *Focus on: a proportionate approach to priority populations.* Public Health Ontario. https://www. publichealthontario. ca/en/eRepository/Focus_On_Priority_Populations. pdf. Accessed, 2016. **29**.
60. Marmot, M. and R. Bell, *Fair society, healthy lives.* Public health, 2012. **126**: p. S4-S10.
61. Rose, G., *Sick individuals and sick populations.* International journal of epidemiology, 2001. **30**(3): p. 427-432.
62. Barter, P.J., et al., *Effect of statins on HDL-C: a complex process unrelated to changes in LDL-C: analysis of the VOYAGER Database.* Journal of lipid research, 2010. **51**(6): p. 1546-1553.
63. Mooradian, A.D., *Dyslipidemia in type 2 diabetes mellitus.* Nature Reviews Endocrinology, 2009. **5**(3): p. 150-159.
64. Frei, B., et al., *Gas phase oxidants of cigarette smoke induce lipid peroxidation and changes in lipoprotein properties in human blood plasma. Protective effects of ascorbic acid.* Biochemical Journal, 1991. **277**(1): p. 133-138.
65. Arai, T., et al., *Increased plasma cholesteryl ester transfer protein in obese subjects. A possible mechanism for the reduction of serum HDL cholesterol levels in obesity.* Arteriosclerosis and thrombosis: a journal of vascular biology, 1994. **14**(7): p. 1129-1136.
66. Hu, F.B., *Sedentary lifestyle and risk of obesity and type 2 diabetes.* Lipids, 2003. **38**(2): p. 103-108.

67. Thorp, A.A., et al., *Deleterious associations of sitting time and television viewing time with cardiometabolic risk biomarkers: Australian Diabetes, Obesity and Lifestyle (AusDiab) study 2004–2005.* Diabetes care, 2010. **33**(2): p. 327-334.
68. McNaughton, S.A., et al., *Dietary quality is associated with diabetes and cardio-metabolic risk factors.* The Journal of nutrition, 2009. **139**(4): p. 734-742.
69. Williams, E.D., et al., *Health behaviours, socioeconomic status and diabetes incidence: the Australian Diabetes Obesity and Lifestyle Study (AusDiab).* Diabetologia, 2010. **53**(12): p. 2538-2545.
70. Sodjinou, R., et al., *Obesity and cardio-metabolic risk factors in urban adults of Benin: relationship with socio-economic status, urbanisation, and lifestyle patterns.* BMC public health, 2008. **8**(1): p. 84.
71. Cummins, S. and S. Macintyre, *Food environments and obesity—neighbourhood or nation?* International journal of epidemiology, 2006. **35**(1): p. 100-104.
72. Fraser, L.K., et al., *The geography of fast food outlets: a review.* International journal of environmental research and public health, 2010. **7**(5): p. 2290-2308.
73. Macdonald, L., S. Cummins, and S. Macintyre, *Neighbourhood fast food environment and area deprivation—substitution or concentration?* Appetite, 2007. **49**(1): p. 251-254.
74. Pearce, J., et al., *Neighborhood deprivation and access to fast-food retailing: a national study.* American journal of preventive medicine, 2007. **32**(5): p. 375-382.
75. Buttar, H.S., T. Li, and N. Ravi, *Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation.* Experimental & clinical cardiology, 2005. **10**(4): p. 229.
76. Chomistek, A.K., et al., *Healthy lifestyle in the primordial prevention of cardiovascular disease among young women.* Journal of the American College of Cardiology, 2015. **65**(1): p. 43-51.
77. Fiuza-Luces, C., et al., *Exercise benefits in cardiovascular disease: beyond attenuation of traditional risk factors.* Nature Reviews Cardiology, 2018. **15**(12): p. 731-743.
78. Xu, H., *Comparing spatial and multilevel regression models for binary outcomes in neighborhood studies.* Sociological methodology, 2014. **44**(1): p. 229-272.
79. Chaix, B., J. Merlo, and P. Chauvin, *Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France.* Journal of Epidemiology & Community Health, 2005. **59**(6): p. 517-526.

Chapter 6

Role of Area-level Access to Primary Care on the Geographic Variation of Cardiometabolic Risk Factor Distribution: A Multilevel Analysis of the Adult Residents in the Illawarra-Shoalhaven Region of NSW, Australia

Chapter 6: Role of Area-level Access to Primary Care on the Geographic Variation of Cardiometabolic Risk Factor Distribution: A Multilevel Analysis of the Adult Residents in the Illawarra- Shoalhaven Region of NSW, Australia

6.1. Publication profile

This chapter presents the substantive content of research published in the ***International Journal of Environmental Research and Public Health***, on 14th June 2020.

Journal article

Toms, R., Feng, X., Mayne, D. J., & Bonney, A. *Role of area-level access to primary care on the geographic variation of cardiometabolic risk factor distribution: a multilevel analysis of the adult residents in the Illawarra—Shoalhaven Region of NSW, Australia***. International journal of environmental research and public health**, 2020. 17(12), 4297.

Available from: https://www.mdpi.com/1660-4601/17/12/4297

DOI: https://doi.org/10.3390/ijerph17124297

Published version of the article

The published version of the article is appended within the **'Supplementary Materials'** section of the thesis as **Supplementary Material 4**.

6.2 Abstract

Introduction: Previous research reports geographic variation in cardiometabolic risk factors (CMRFs), with higher prevalence often in disadvantaged areas. Access to primary care is important for the identification, control and management of CMRFs. This study investigated whether geographic access to primary care contributed to the area-level variation in CMRFs.

Methods: A cross-sectional study design was used to analyse data on seven CMRFs collected from 2012 to 2017: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); urinary albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); and body mass index (BMI). Multilevel logistic regression models were used to derive the association between area-level access to primary care and geographic variation of CMRFs after adjusting for individual and area-level co-variates. Two-step floating catchment area method was used to calculate primary care access for small areas within the study region. Primary care provider data were retrieved from publicly available sources current in 2016.

Results: Multilevel logistic regression models indicated that after adjusting for age, sex and area-level disadvantage, primary care access was inversely associated with low HDL (OR 0.94, CI 0.91-0.96) and obesity (OR 0.91, CI 0.88-0.93), but was not associated with five of the remaining CMRFs. The area-level variation in CMRFs explained by primary care access was ≤ 10.5% and didn't demonstrate any attenuating effect on the association between area-level disadvantage and CMRFs shown in previous models. The ICCs of the fully adjusted models ranged between 0.4 -1.8% in low eGFR and BMI respectively.

Conclusion: The amount of geographic variation in CMRFs in the study region specifically explained by geographic access to primary care was small and did not have an attenuating effect on the association between area-level disadvantage and CMRFs. Thus, the observed geographic variation in CMRFs in the study region could be better explained by the area-level disadvantage rather than their area-level geographic access to primary care services. The findings are consistent with the previous findings from Australia but suggest both the complexity of defining access and desirability of future studies to gain more understanding of performance measures for health outcomes associated with primary care access.

Keywords: geographic access; cardiometabolic risk factor; geographic variation; multilevel logistic regression; primary care access.

6.3 Introduction

Cardiometabolic risk factors (CMRFs) demonstrate significant variation in geographic distribution within countries globally.[1-10] Higher prevalence and clustering of CMRFs is often reported for socioeconomically disadvantaged areas.[11-20] Area-level access to primary care is essential for the identification and management of CMRFs, especially when considering their chronic nature after detection.[21-23] Therefore, access to primary care may have an associating role in the geographic variation of CMRFs.[24]

Previous studies have reported that access to primary care can play a role in the control and management of certain CMRFs.[21, 25-28] Research indicates that access to primary care varies across areas, as the locations of primary care physicians and services is tends to be positively correlated with population distributions.[29, 30]There is evidence that medical consultations were reported less likely to happen when physical access to health care services is lower.[21] Also, access to adequate treatment facilities were reported to have an inverse association with certain CMRFs, such as hypertension [31, 32], end stage renal disease[33], and diabetes mellitus[34]. However, these reports are based on individual risk factors and consistent evidence across a range of CMRFs may provide a stronger evidence base for healthcare service commissioning across areas.

Evidence regarding the association of CMRFs with primary care access over and above area-level disadvantage may inform area-level resource allocation of primary care services in disadvantaged areas.[24, 35] The aim of this study was to quantify the amount of geographic variation in CMRFs within the Illawarra-Shoalhaven region of Australia explained by differences in area-level primary care access after adjusting for area-level disadvantage.

6.4 Methods

A retrospective cross-sectional design was adopted to analyse the association between the area-level access to primary care and the geographic distribution of CMRFs among the residents of Illawara-Sholhaven region and to quantify the extent to which primary care access accounts for geographic variation in CMRFs. The study was approved by the University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical Human Research Ethics Committee (HREC protocol No: 2017/124).

The study focused on the Illawarra-Shoalhaven region of New South Wales (NSW), Australia. The study area consists of multiple regional cities, smaller towns and rural areas; including the local government areas of Kiama, Shellharbour, Shoalhaven and Wollongong. The region covers a geographical area of 5,615 square kilometres and had a population of 369,469 people at the 2011 Australian Census of Population and Housing.[36] The geographic unit of analysis used in this study was the Statistical Area 1 (SA1), which is the smallest statistical output unit of the 2011 Census and have an average population of 400 people (range: 200 to 800).[37] The study area encompasses 980 conterminous SA1s.[38] Figure 6.1 shows the study area showing SA1s and major landmarks of the region.

The Illawarra-Shoalhaven has a diverse socio-economic profile, making it a useful region for area-level population health studies.[39] The population profile of the Illawarra-Shoalhaven region is culturally and linguistically diverse, with a significant proportion of non-English speaking people (10.5%) residing in this region who have migrated from overseas. [40] In addition, at the 2011 Census the region is identified to have more than the NSW state and Australian national averages of: 1) Aboriginal and/or Torres Strait

Islander peoples (3% versus 2.5% NSW and 2.5% Australia); 2) aged (>=65 years) population (17.6% versus 14.5% NSW and 13.8% Australia) ; 3) single-parent households (5.8% versus 5.3% NSW and 5.2% Australia); and 4) unemployment (7.1% versus 5.1% NSW and 5.1% Australia) and lower labour force participation rate (57.9% versus 64.6% NSW and 66.2% Australia).[40] ABS 2011 census data indicate that more than 31 % of people in the Illawarra-Shoalhaven reside in Inner Regional areas, and 9.1% of households within this region did not have a motor vehicle.[40] The Illawarra-Shoalhaven geography and a limited public transport system, especially in isolated communities, make it difficult for many people to access health services quickly.[40] These characteristics of the study region directly indicate the vulnerability of its population to poorer health outcomes.

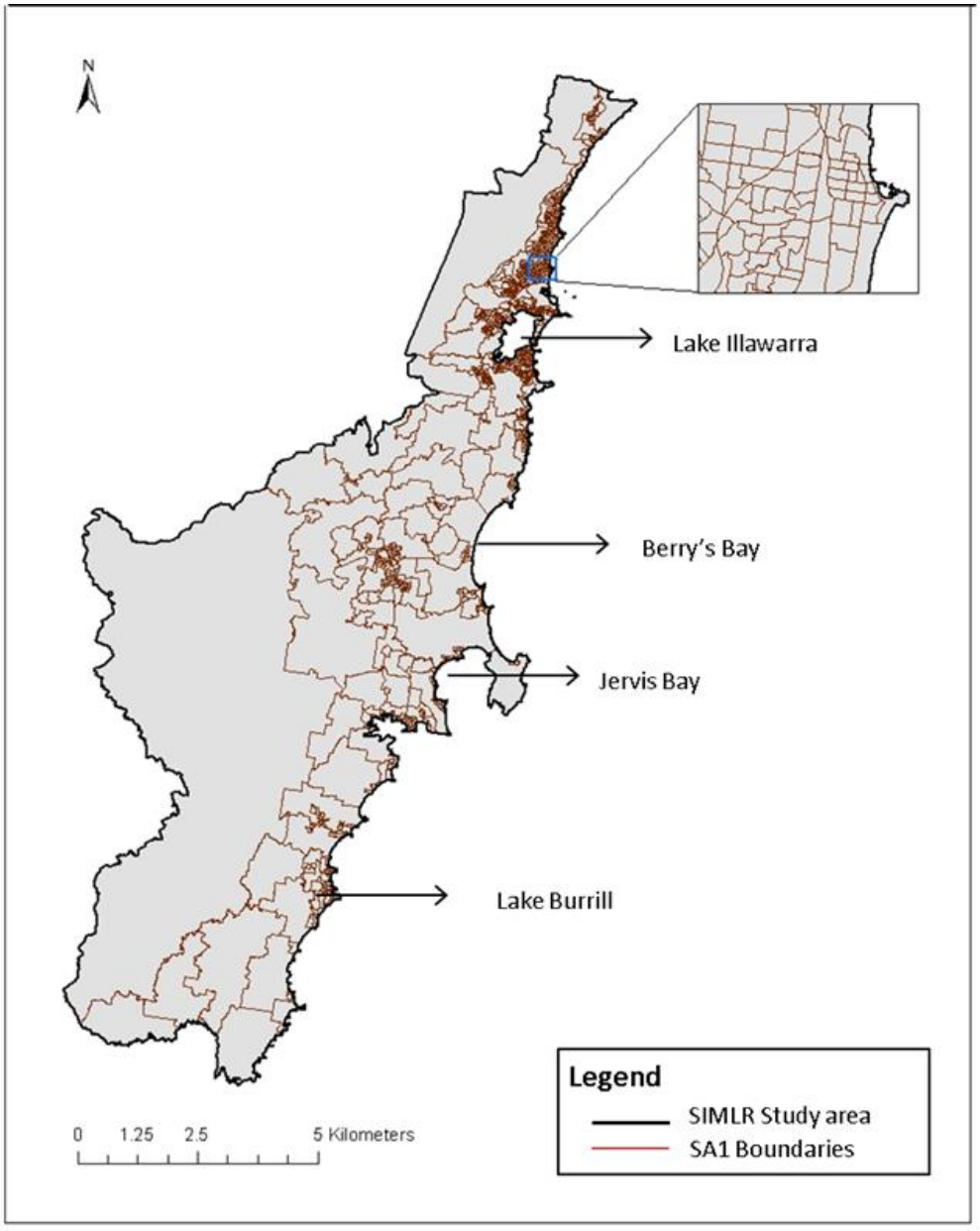

Figure 6.1: Map of the Illawarra-Shoalhaven region of NSW, Australia, showing SA1 areas and major landmarks.

Data

The study used three different databases: the CMRF pathology test data; primary care provider data; and the estimated resident populations from the 2011 Australian Census of Population and Housing. The CMRF test data were extracted from the Southern IML Research (SIMLR) Study database. The SIMLR Study database comprises de-identified and internally linked pathology results from a major pathology provider in the study region and provides near-census coverage of the study population.[41] The CMRF test data were extracted for multiple risk factors on the most recent test results, of non-pregnant adults aged 18 years and over, undergoing a laboratory test between 01 January 2012 and 31 December 2017.

The primary care provider data were manually extracted in 2016 from publicly available data sources, including Yellow Pages, White Pages, online general practitioner (GP) appointment booking services and Google search results. The 2011 Australian Census of Population and Housing data was accessed to extract the population denominator data of the study region at SA1 level.[42]

Variables

Dependent variable

Dichotomised results of the CMRF tests were the individual-level dependent variables. The test results analysed in this study included: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); urinary albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); and objectively-measured body mass index (BMI). The CMRF test results were dichotomised into higher risk and lower risk values based on established risk classification guideline values used previously.[10] Table 6.1 shows the CMRF definitions used in this study.

Table 6.1: Definitions of CMRFs test results

	'Higher risk' CMRFs	**Definition**
1.	High FBSL	FBSL ≥ 7.0 mmol/l.[43]
2.	High HbA1c	HbA1c > 7.5% .[43]
3.	High TC	TC ≥ 5.5 mmol/l.[44]
4.	Low HDL	HDL < 1 mmol/l.[44]
5.	High ACR	ACR ≥ 30 mcg/L to mg/l.[45]
6.	Low eGFR	eGFR < 60 mL/min/1.73m^2[45]
7.	Obesity	BMI ≥ 30 kg/m^2.[46]

CMRFs - Cardiometabolic risk factors; FBSL - Fasting Blood Sugar Level; HbA1c - Glycated Haemoglobin; TC - Total Cholesterol; HDL - High Density Lipoprotein; ACR - Albumin Creatinine Ratio; eGFR - estimated Glomerular Filtration Rate; BMI - Body Mass Index.

Independent variable

Primary care access calculated at the SA1-level was the independent study variable. An access index score was calculated for each SA1 using a two-step floating catchment area (2SFCA) method, which balanced both supply and demand of primary care services in the study region.

The 2SFCA method was developed by Luo and Wang in 2003 to measure geographic accessibility of health care services.[47] The method has undergone several enhancements since its inception but essentially consists of two steps underpinned by a gravity model.[47, 48] The first step computes a

population-to-provider ratio for each primary care service location by aggregating the population size of the SA1s whose centroids are located within a defined spatial buffer distance.[49] The total number of general practitioners working in the primary care service locations within this buffer distance were the numerators for the provider to population ratio calculations.

Thus, step 1:

$$Rj = Sj / \sum_i Pi , \quad (6.1)$$

where *Sj* is the number of general practitioners at location *j*, pi is the number of adult residents in the SA1s (Those SA1's geographic centroids are located within the spatial buffer distance of the primary care locations) and *Rj* is the population-to-provider ratio for service *j*.[49]

In step 2, a population-to-provider ratio (access score) is computed for each geographic centroid of the SA1s by aggregating all primary care service population-to provider ratios of the primary care services that are located within the same spatial buffer distance.[49]

Thus, step 2:

$$Ai = \sum_j R_j , \quad (6.2)$$

where *Ai* is the access index for population location *i*.

The resulting access indices were retained as a continuous variable for the analyses. A higher score indicated better geographic access of the SA1s to primary care services.

A spatial radial buffer distance of 30 km was chosen to compute primary care access for SA1s in the study region. In the preliminary stage, sensitivity analyses were performed using 1 km, 16 km and 30 km spatial radial buffer distances. In step 1 2SFCA analyses, the 1 km distance covered only 545 (56%) SA1 centroids in the study region in relation to the primary care provider locations, whereas a 16 km radial buffer distance covered 973 (~99%) and a 30 km radial buffer distance covered 978 (~100%) SA1s' geographic centroids. Therefore, a radial buffer distance of 30 km was chosen to determine the access which was observed to cover the mixed rural, semi-rural and urban distribution of the population in the study region well.

Covariates

The individual-level variables adjusted for in the study were: sex (male and female) and age group (18–29, 30–39, 40–49, 50–59, 60–69, 70–79 and 80+ years). The area-level covariate was the area-level socioeconomic disadvantage of the SA1s. The Index of Relative Socioeconomic Disadvantage (IRSD) score of the SA1s in the study region was used as the measurement variable for the area-level socioeconomic disadvantage of the SA1s. The IRSD summarises a range of measures of relative socioeconomic disadvantage of people and/or households within SA1s and includes: level of income; education; employment; family structure; disability; housing; transportation; and internet connection.[50] A higher IRSD score indicated lower levels of disadvantage.[50]

Statistical analyses

Multilevel logistic regression models were fitted to individual CMRF test data (Level 1) nested within SA1s (Level 2). For each of the seven CMRFs analysed in this study, five multilevel models were fit that included fixed effects for sex, age, IRSD score and access index; and random effect intercepts for SA1s.

Model 1 (M1) was a null model; Model 2 (M2) included the area-level study variable (access index) only; Model 3 (M3) included individual-level factors (age and sex) only; Model 4 (M4) included individual and area-level factors (age, sex and IRSD score); and Model 5 (M5) included M4 variables plus access index. Thus, the final model (M5) estimated the effect of primary care access after adjusting for individual and area-level factors. Odds ratios (ORs) were derived from the exponentials of regression coefficients from fitted models. As the IRSD scores and access index of the SA1s were fitted as mean-centred continuous variables, ORs were expressed per standard deviation unit change in these variables. Statistical significance of the models was evaluated using likelihood ratio tests and a type I error rate of 0.05.

Model comparison

Model fit was compared using the Akaike Information Criterion (AIC). The models were also evaluated for: area-level variance (τ^2); proportional change in variance (PCV) in comparison with the null model; Intra-cluster Correlation Coefficient (ICC) of the model; and the Median Odds Ratios (MORs). The ICC and MOR of the models were used to index the between-area variability. A latent variable approach was used to derive the ICC of models.[51] The MOR translates the area-level variance into an easily interpretable OR and is assumed to be statistically independent of the test specific prevalence of the CMRFs.[52] The unique contribution of the primary care access of the SA1s to the area-level variance of CMRF was estimated through the reduction in PCV between M4 and M5.

Statistical package

All mapping and geospatial measurements were performed using ArcGIS version 10.4.1(ESRI Inc. Redlands, CA, USA).[53] All statistical analyses were performed using R version 3.4.4. (R Foundation for Statistical Computing, Vienna, Austria).[54] Multilevel models were fit using the glmer function in the lme4 package[55]; and likelihood ratio tests were calculated using the lrtest function in the lmtest package[56].

6.5 Results

A total of 1, 132, 029 CMRF test results for 256, 525 individual residents in the Illawarra-Shoalhaven region between 2012 and 2017 were extracted for analysis. The mean number of tests undertaken per person was 4.4 (SD = 1.8, range = 1-7). After excluding 1, 162 (1.0%) test results with incomplete details, a total of 1, 130, 894 tests were retained in the final data set. IRSD score of the SA1s were the most frequent missing variable, as this was not available for some SA1s in the study region.[57] Available IRSD scores ranged between 446.7 and 1143.7 (mean = 976.7, SD = 98.6) for SA1s, with a higher score indicating lower area-level disadvantage. Table 6.2 details the individual-level CMRFs risk proportions of the final data set.

Table 6.2: Frequency and proportion of CMRFs risk classification with gender

	Cardiometabolic risk	Test n	Higher risk n (%)*	Male n (%)*	Female n (%)*
1	High FBSL	193679	16280(8.4)	9289 (4.8)	6991 (3.6)
2	High HbA1c	73885	7927(10.7)	4448(6.0)	3479(4.7)
3	High TC	194816	63422(32.6)	26139(13.4)	37283(19.1)
4	Low HDL	182237	21261(11.7)	15885(8.7)	5376(3.0)
5	High ACR	50790	2047(4.0)	1266(2.5)	781(1.5)
6	Low eGFR	244166	27241(11.2)	12456(5.1)	14785(6.1)
7	Obesity	192455	64832(33.7)	29613(15.4)	35319(18.4)

*The denominators of the percentages are the total number of each CMRFs tests.

Table 6.3 presents the summary statistics of the overall CMRF tests across the SA1s in the study region. Table 6.4 outlines the descriptive statistics of the individual CMRF tests across the SA1s in the study region.

Table 6.3: Summary statistics of the overall CMRF tests across the SA1s in the study region.

		Summary statistics of the tests across SA1s							
		Min.	1st Qu.	Median	IQR	Mean	SD	3rd Qu.	Max.
1	Total CMRF tests/SA1	1	198	261	135	266.7	117.0	333	965
2	% Female tested /SA1	0	54.2	56.4	4.3	56.0	6.5	58.5	100
3	% Male tested /SA1	0	41.5	43.6	4.3	44.0	6.5	45.8	100
4	Total CMRF tests per person /SA1	1	4.3	4.4	0.4	4.4	0.3	4.6	6

Table 6.4: Summary statistics of the individual CMRF tests across the SA1s in the study region.

		Number of tests /SA1				Average proportion of tested people per age group (in years) /SA1						
CMRFs	Total tests	Mean	SD	Median	IQR	18 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80+
FBSL	193679	202	87.8	199	107.5	10.2	12.1	15.4	19.5	19.8	14.8	8.2
HbA1c	73885	77.6	37.6	73.5	46	4.9	6.8	11.8	19.1	24.1	20.7	12.6
TC	194816	203.1	89.4	201	108.5	7.4	9.8	16.3	20.8	20.9	15.7	9.1
HDL	182237	190	84.2	187	101.5	6.3	9.3	16.5	21.2	21.5	16.2	9.1
ACR	50790	53.6	28.1	50	32.5	3.2	4.7	10.1	19.2	26.7	23.1	13.0
eGFR	244166	253.8	111.6	249.5	128	13.6	11.9	14.9	18.0	18.0	13.9	9.7
BMI	192455	200.5	87.6	197	103.3	11.7	11.5	15.9	19.8	20.3	15.7	5.2

A total of 165 primary care service locations with 611 general practitioners were identified in the study area in 2016. The primary care access index of the SA1s in the study region ranged between 0 and 5.41 general partitioners per 1000 people (mean = 2.1, SD = 0.77). Figure 6.2 illustrates the distribution of the primary care access index within the within the study region. Multilevel logistic regression models for each CMRF are presented in Table 6.5(a-g) and comparisons of the random effects of the models are presented in Table 6.6.

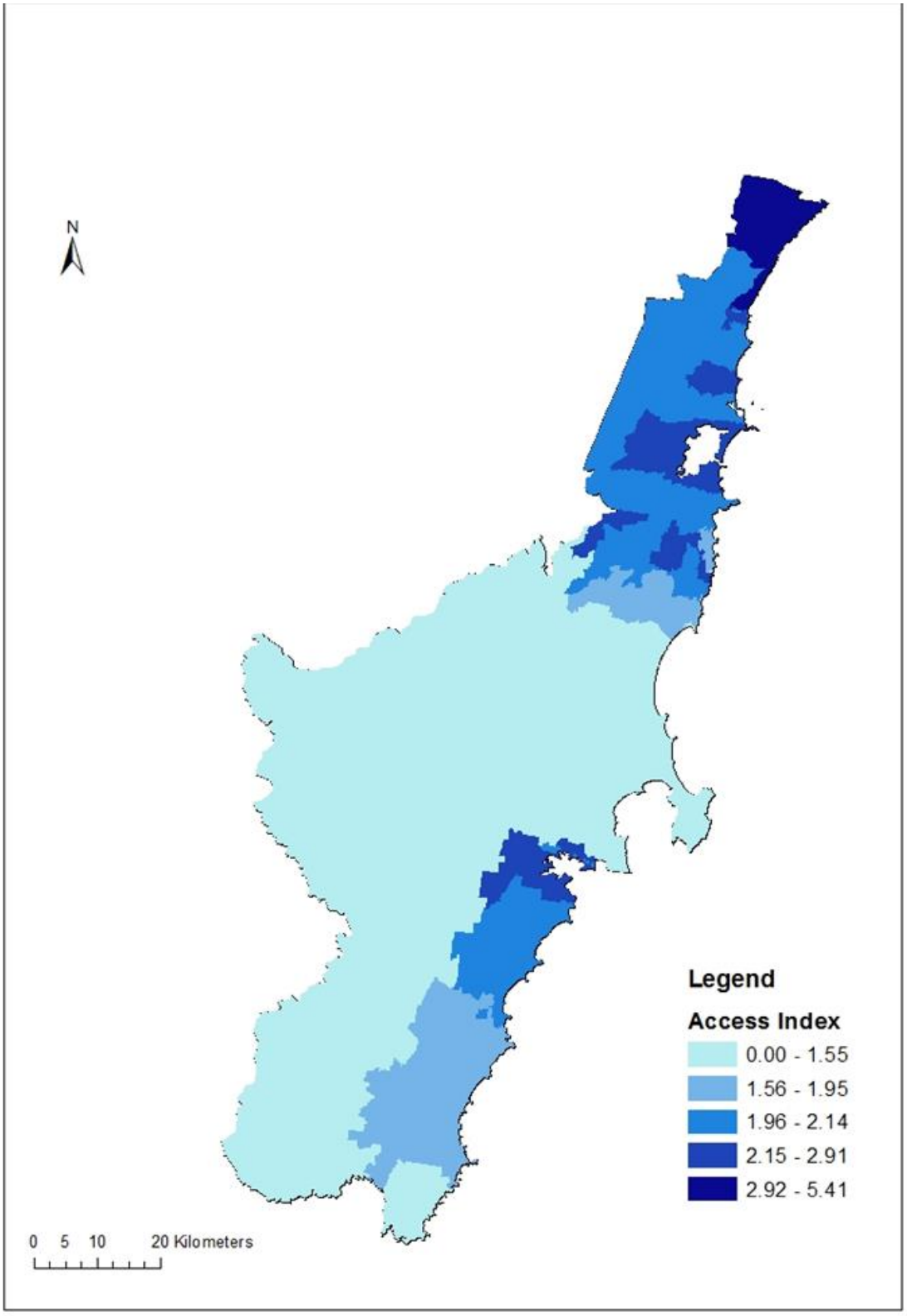

Figure 6.2: Geographic access to primary care services in the Illawarra-Shoalhaven region of the NSW, Australia

Table 6.5a: Multilevel logistic regression model summaries of high FBSL (FBSL ≥ 7.0 mmol/l).

Variables	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
Significance (LRT)	p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
High FBSL										
Intercept	0.09(0.09 - 0.09)	p < 0.001	0.09(0.09 - 0.09)	p < 0.001	0.01(0.01-0.01)	p<0.001	0.01(0.01 - 0.01)	p<0.001	0.01(0.01 - 0.01)	p<0.001
Access			0.89(0.87 - 0.92)	p < 0.001						
Sex: Female					Reference					
Male					1.63 (1.58-1.69)	p<0.001	1.63(1.58 - 1.69)	p<0.001	1.63(1.58 - 1.69)	p<0.001
Age:18—29					Reference					
30—39					1.63(1.40-1.90)	p<0.001	1.65(1.41 - 1.92)	p<0.001	1.65(1.41 - 1.92)	p<0.001
40—49					3.53(3.08-4.05)	p<0.001	3.57(3.11 - 4.10)	p<0.001	3.57(3.11 - 4.10)	p<0.001
50—59					6.77(5.93-7.72)	p<0.001	6.81(5.97 - 7.77)	p<0.001	6.80(5.97 - 7.75)	p<0.001
60—69					11.07 (9.72-12.6)	p<0.001	11.07(9.7 - 12.6)	p<0.001	11.05(9.7 - 12.6)	p<0.001
70—79					13.93 (12.2-15.9)	p<0.001	13.8(12.1 - 15.7)	p<0.001	13.8(12.1 - 15.7)	p<0.001
80+					12.33 (10.8-14.1)	p<0.001	12.1(10.6 - 13.9)	p<0.001	12.1(10.6 - 13.8)	p<0.001
IRSD							0.79(0.77 - 0.80)	p<0.001	0.79(0.77 - 0.81)	p<0.001
Access									0.98(0.96 - 1.00)	0.111
AIC	111022.8		110962.3		103066.2		102652.6		102652.0	
Variance	0.101		0.091		0.103		0.040		0.039	
PCV	-		- 9.98%		+ 1.88 %		-60.90%		-61.05%	
ICC (%)	3.0		2.7		3.0		1.2		1.2	
MOR	1.36		1.334		1.36		1.209		1.209	
Proportional variance explained by Access to primary care: - 0.38%										

FBSL – Fasting blood sugar level; IRSD - Index of relative socioeconomic disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1.

Table 6.5b: Multilevel logistic regression model summaries of high HbA1c (HbA1c > 7.5%).

Variables	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
Significance (LRT)	p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
High HbA1c										
Intercept	0.12(0.11 - 0.12)	p < 0.001	0.12(0.11 - 0.12)	p < 0.001	0.07 (0.06 - 0.07)	p<0.001	0.07(0.06 - 0.08)	p<0.001	0.07(0.06 - 0.08)	p<0.01
Access			0.95(0.92 - 0.98)	p < 0.001						
Sex: Female					Reference					
Male					1.38 (1.3 - 1.45)	p<0.001	1.39(1.32 - 1.45)	p<0.001	1.39(1.32 -1. 45)	p<0.001
Age:18—29					Reference					
30—39					0.81(0.68 - 0.96)	p<0.01	0.81(0.68 - 0.96)	p<0.01	0.81(0.68 -0. 97)	p<0.01
40—49					1.24(1.07 - 1.44)	p<0.001	1.25(1.08 - 1.45)	p<0.001	1.26(1.08 -1. 46)	p<0.001
50—59					1.56(1.36 - 1.80)	p<0.001	1.56(1.36 - 1.80)	p<0.001	1.57(1.36 -1. 81)	p<0.001
60—69					1.64(1.43 - 1.88)	p<0.001	1.64(1.43 - 1.88)	p<0.001	1.64(1.43 -1. 89)	p<0.001
70—79					1.64(1.42 - 1.88)	p<0.001	1.62(1.41 - 1.86)	p<0.001	1.63(1.42 -1. 87)	p<0.001
80+					1.63(1.41 - 1.88)	p<0.001	1.62(1.40 - 1.87)	p<0.001	1.62(1.41 -1. 87)	p<0.001
IRSD							0.79(0.77 - 0.81)	p<0.001	0.79(0.77 -0. 81)	p<0.001
Access									1.00(0.97-1. 03)	0.750
AIC	50114.5		50105.9		49690.2		49438.2		49440.0	
Variance	0.103		0.100		0.106		0.048		0.047	
PCV	-		- 2.430%		+ 3.02 %		- 53.78%		- 53.80%	
ICC (%)	3.0		3.0		3.1		1.4		1.4	
MOR	1.36		1.353		1.358		1.231		1.231	
Proportional variance explained by Access to primary care: - 0.04%										

HbA1c - Glycated Haemoglobin; IRSD - Index of Relative Socioeconomic Disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1.

Table 6.5c: Multilevel logistic regression model summaries of high TC (TC ≥ 5.5 mmol/l).

Variables	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
Significance (LRT)	p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
high TC										
Intercept	0.42(0.41- 0.43)	p < 0.001	0.42(0.41- 0.43)	p < 0.001	0.20 (0.19 - 0.21)	p<0.001	0.20(0.19 - 0.21)	p<0.001	0.20(0.19 - 0.21)	p<0.01
Access			1.02 (1.00 - 1.03)	p < 0.01						
Sex: Female					Reference					
Male					0.69 (0.68 - 0.71)	p<0.001	0.69(0.68 - 0.71)	p<0.001	0.69(0.68 - 0.71)	p<0.001
Age:18—29					Reference					
30—39					2.02 (1.91 - 2.14)	p<0.001	2.01(1.90 - 2.13)	p<0.01	2.01(1.90 - 2.13)	p<0.001
40—49					3.01 (2.86 - 3.17)	p<0.001	3.00(2.85 - 3.16)	p<0.001	3.00(2.85 - 3.16)	p<0.001
50—59					4.08 (3.88 - 4.29)	p<0.001	4.07(3.87 - 4.28)	p<0.001	4.07(3.87 - 4.28)	p<0.001
60—69					2.95 (2.80 - 3.10)	p<0.001	2.95(2.80 - 3.10)	p<0.001	2.95(2.80 - 3.10)	p<0.001
70—79					1.60 (1.52 - 1.69)	p<0.001	1.61(1.52 - 1.69)	p<0.001	1.61(1.52 - 1.69)	p<0.001
80+					1.13 (1.07 - 1.20)	p<0.001	1.14(1.07 - 1.21)	p<0.001	1.14(1.07 - 1.21)	p<0.001
IRSD							1.06(1.04 - 1.07)	p<0.001	1.06(1.04 - 1.07)	p<0.001
Access									1.00(0.98 - 1.01)	0.616
AIC	235931.6		235927.9		227254.6		227193.8		227195.5	
Variance	0.0255		0.0250		0.020		0.01703		0.01705	
PCV	-		-1.69%		-21.76%		-33.11%		-33.07%	
ICC (%)	0.8		0.8		0.6		0.5		0.5	
MOR	1.16		1.163		1.14		1.133		1.133	
Proportional variance explained by Access to primary care: + 0.12%										

TC - Total Cholesterol; IRSD - Index of Relative Socioeconomic Disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1.

Table 6.5d: Multilevel logistic regression model summaries of Low HDL (< 1 mmol/l).

Variables Significance (LRT)	**Model 1** p < 0.0001		**Model 2** p < 0.0001		**Model 3** p < 0.0001		**Model 4** p < 0.0001		**Model 5** p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
low HDL										
Intercept	0.13(0.13—0.13)	p < 0.001	0.13(0.13—0.18)	p < 0.001	0.06 (0.06—0.07)	p<0.001	0.06(0.06—0.07)	p<0.001	0.06(0.06—0.07)	p<0.001
Access			0.92(0.90—0.94)	p < 0.001						
Sex: Female					Reference					
Male					3.98(3.85—4.11)	p<0.001	3.98 (3.85—4.11)	p<0.001	3.98(3.85—4.11)	p<0.001
Age:18—29					Reference					
30—39					1.11(1.03—1.20)	p<0.001	1.12 (1.04—1.21)	p<0.001	1.12(1.04—1.21)	p<0.001
40—49					0.99(0.92—1.05)	0.658	1.00 (0.93—1.07)	0.957	1.00(0.93—1.07)	0.947
50—59					0.88(0.82—0.94)	p<0.001	0.89 (0.83—0.95)	p<0.001	0.88(0.83—0.95)	p<0.001
60—69					0.82(0.77—0.88)	p<0.001	0.83 (0.77—0.88)	p<0.001	0.82(0.77—0.88)	p<0.001
70—79					0.86(0.80—0.92)	p<0.001	0.85 (0.80—0.91)	p<0.001	0.85(0.79—0.91)	p<0.001
80+					0.93(0.86—1.00)	p<0.010	0.92 (0.85—0.99)	p<0.010	0.91(0.85—0.99)	p<0.010
IRSD							0.81 (0.80—0.82)	p<0.001	0.82(0.80—0.83)	p<0.001
Access									0.95(0.93—0.97)	p<0.001
AIC	130649.70		130601.4		122700.0		122291.9		122271.4	
Variance	0.07		0.064		0.081		0.031		0.029	
PCV	-		-9.48%		+15.25%		-55.90%		-59.05%	
ICC (%)	2.1		1.9		2.4		0.9		0.9	
MOR	1.289		1.273		1.313		1.183		1.183	
Proportional variance explained by Access to primary care: - 6.61%									6.61%	

HDL - High Density Lipoprotein; IRSD - Index of Relative Socioeconomic Disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1.

Table 6.5e: Multilevel logistic regression model summaries of High ACR (≥ 30 mcg/L to mg/l)

Variables	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
Significance (LRT)	p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
High ACR										
Intercept	0.04(0.04 - 0.04)	p < 0.001	0.04(0.04 - 0.04)	p < 0.001	0.02(0.02 - 0.03)	p < 0.001	0.02(0.02 - 0.03)	p < 0.001	0.02(0.02 - 0.03)	p < 0.001
Access			0.91(0.86 - 0.96)	p < 0.001						
Sex: Female					Reference					
Male					1.75(1.60 - 1.92)	p < 0.001	1.76(1.60 - 1.93)	p<0.001	1.75(1.60 - 1.92)	p < 0.001
Age:18—29					Reference	p < 0.001				
30—39					1.00(0.69 - 1.45)	0.985	1.01(0.69 - 1.46)	0.978	1.00(0.69 - 1.46)	0.982
40—49					0.69(0.47 - 0.97)	p<0.01	0.70(0.50 - 1.00)	p<0.01	0.70(0.50 - 1.00)	p<0.01
50—59					0.77(0.56 - 1.05)	0.101	0.77(0.56 - 1.07)	0.115	0.77(0.56 - 1.06)	0.115
60—69					0.95(0.70 - 1.30)	0.762	0.96(0.71 - 1.31)	0.794	0.96(0.70 - 1.30)	0.777
70—79					1.55(1.15 - 2.10)	p<0.001	1.55(1.14 - 2.09)	p<0.001	1.54(1.14 - 2.08)	p<0.001
80+					2.74(2.02 - 3.71)	p<0.001	2.71(2.00 - 3.67)	p<0.001	2.70(1.99 - 3.66)	p<0.001
IRSD							0.82(0.78 - 0.85)	p<0.001	0.82(0.79 - 0.86)	p<0.001
Access									0.97(0.91 - 1.02)	0.206
AIC	17130.0		17119.9		16585.2		16510.8		16511.2	
Variance	0.092		0.085		0.073		0.028		0.025	
PCV	-		-7.92%		-20.53%		-69.14%		-72.39%	
ICC (%)	2.7		2.5		2.2		0.9		0.8	
MOR	1.34		1.321		1.30		1.175		1.165	
Proportional variance explained by Access to primary care: -10.53%										

ACR - Albumin Creatinine Ratio; IRSD - Index of Relative Socioeconomic Disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1.

Table 6.5f: Multilevel logistic regression model summaries of low eGFR (< 60 mL/min/1.73m2)

Variables	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
Significance (LRT)	p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Low eGFR										
Intercept	0.11(0.11 - 0.12)	p < 0.001	0.11(0.11 - 0.12)	p < 0.001	0.00(0.00 - 0.00)	p < 0.001	0.00(0.00 - 0.00)	p < 0.001	0.00(0.00 - 0.00)	p < 0.001
Access			0.89(0.86 - 0.92)	p < 0.001						
Sex: Female					Reference					
Male					0.98(0.95 - 1.01)	0.208	0.98(0.95 - 1.01)	0.258	0.98(0.95 - 1.01)	0.248
Age:18—29					Reference	p < 0.001				
30—39					1.66(1.25 - 2.20)	p < 0.001	1.66(1.24 - 2.23)	p < 0.001	1.65(1.22- 2.24)	p < 0.001
40—49					4.26(3.35 - 5.41)	p < 0.001	4.27(3.34 - 5.50)	p < 0.001	4.30(3.32- 5.58)	p < 0.001
50—59					12.26(9.8 - 15.3)	p < 0.001	12.29(9.73 - 15.52)	p < 0.001	12.28(9.63- 15.66)	p < 0.001
60—69					41.8(33.6 - 51.8)	p < 0.001	41.84(33.29 - 52.57)	p < 0.001	41.83(32.97- 53.06)	p < 0.001
70—79					150.7(121.3 - 187.1)	p < 0.001	149.69(119.3 - 187.9)	p < 0.001	149.6(118.1- 189.5)	p < 0.001
80+					509.3(410.1 - 632.4)	p < 0.001	503.19(400.9 - 631.6)	p < 0.001	503.0(396.9- 637.4)	p < 0.001
IRSD							0.90(0.88 - 0.91)	p < 0.001	0.90(0.88- 0.91)	p < 0.001
Access									1.00(0.98- 1.02)	0.925
AIC	167164.8		167113.4		115257.1		115109.2		115111.2	
Variance	0.189		0.176		0.024		0.013		0.013	
PCV	-		-6.53%		-87.26%		-93.31%		-93.26%	
ICC (%)	5.4		5.1		0.7		0.4		0.4	
MOR	1.51		1.492		1.16		1.113		1.113	
Proportional variance explained by Access to primary care: (+) 0.63%										

eGFR - estimated Glomerular Filtration Rate; IRSD - Index of Relative Socioeconomic Disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1.

Table 6.5g: Multilevel logistic regression model summaries of obesity (BMI ≥ 30 kg/m^2)

Variables	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
Significance (LRT)	p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001		p < 0.0001	
	OR (95% CI)	P value	OR (95% CI	P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)	P value
Obesity										
Intercept	0.51(0.50 - 0.52)	p < 0.001	0.51(0.50 - 0.52)	p < 0.001	0.25(0.24 - 0.26)	p < 0.001	0.25(0.24 - 0.25)	p < 0.001	0.25(0.24 - 0.26)	p < 0.001
Access			0.88(0.86 - 0.90)	p < 0.001						
Sex: Female					Reference					
Male					0.99(0.97 - 1.01)	0.214	0.99(0.97 - 1.01)	0.195	0.99(0.97 - 1.01)	0.193
Age:18—29					Reference	p < 0.001				
30—39					1.63(1.56 - 1.71)	p < 0.001	1.64(1.57 - 1.71)	p < 0.001	1.64(1.57 - 1.71)	p < 0.001
40—49					2.20(2.11 - 2.29)	p < 0.001	2.21(2.12 - 2.30)	p < 0.001	2.20(2.12 - 2.30)	p < 0.001
50—59					2.44(2.34 - 2.53)	p < 0.001	2.45(2.35 - 2.54)	p < 0.001	2.44(2.34 - 2.53)	p < 0.001
60—69					2.73(2.63 - 2.84)	p < 0.001	2.74(2.63 - 2.85)	p < 0.001	2.72(2.62 - 2.83)	p < 0.001
70—79					2.44(2.34 - 2.54)	p < 0.001	2.44(2.34 - 2.54)	p < 0.001	2.42(2.33 - 2.52)	p < 0.001
80+					1.46(1.39 - 1.55)	p < 0.001	1.45(1.38 - 1.54)	p < 0.001	1.45(1.37 - 1.53)	p < 0.001
IRSD							0.81(0.79 - 0.82)	p < 0.001	0.82(0.80 - 0.83)	p < 0.001
Access									0.93(0.91 - 0.95)	p < 0.001
AIC	242793.2		242686.2		239122.6		238731.8		238680.6	
Variance	0.115		0.099		0.117		0.068		0.062	
PCV	-		-14.20%		+1.48%		-41.21%		-46.19%	
ICC (%)	3.4		2.9		3.4		2.0		1.8	
MOR	1.38		1.350		1.39		1.282		1.268	
Proportional variance explained by Access to primary care: - 8.47%										

BMI - Body Mass Index; IRSD - Index of Relative Socioeconomic Disadvantage; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1 — Statistical area-level 1

All the null models indicated geographic variation in the distribution of all CMRFs at the SA1 level. Model 2s showed significant inverse associations between access index and all CMRFs except TC, which displayed a positive association with the access index. Model 3s adjusted CMRF models for individual-level age and sex, which accounted for 1.5% (obesity) to 87.3% (eGFR) of unexplained variation in the null model. Model 4s demonstrated significant inverse associations between area-level IRSD and all CMRFs except for TC after adjusting for individual-level factors. Total cholesterol again showed a positive association with IRSD scores. In the final models (M5s), the access index was found to be inversely associated with low HDL (HDL< 1 mmol/l) and obesity (BMI≥ 30 kg/m^2), after adjusting for individual and area-level factors. Including the access index in the final models did not attenuate associations between area-level disadvantage and CMRFs observed in M4s. In the final models, no significant association was found between primary care access and CMRFs except low HDL and obesity.

Reductions in the AIC values were observed for all CMRFs except in TC and eGFR models from the null model (M1) to the final model (M5), indicating a better fit of the final models. The AIC for TC and eGFR models indicated M4 was the best fitting model for these CMRFs. In the unadjusted null models (M1s), low eGFR demonstrated the most area-level variance and high TC showed the least. The access only models (M2s) showed a reduction in the residual variance of all CMRFs from those of null models. The proportional change in variance was highest for obesity (PCV =-14.2%) and lowest for TC (PCV = -1.7%). In Model 3s, adjusting for age and sex initially increased the residual variance of FBSL (PCV = +1.9%), HbA1c (PCV = +3.0%), HDL (PCV = +15.3%) and BMI (PCV = +1.5%). In model 4s, adjusting the CMRFs for individual-level age and sex and area-level disadvantage resulted in major reductions of variance from -33.1% (in TC) to -93.3% (in eGFR). In the final models (M5s), including access index in the models after adjusting for the covariates, had extended the reduction in variance in all CMRFs, except on TC and eGFR. Including the access index had been observed to increase the variance in the TC and eGFR final models, compared with the lower level model. The proportional variance specifically explained by access to primary care was minimal in the final models. Area-level primary care access explained 10.5% of the geographic variation in high ACR, followed by 8.5% variation of obesity and 6.6% variation of low HDL. The geographic variation explained by primary care access was close to zero for high FBSL (0.4%), high HbA1c (0.0%), high TC (-0.1%) and eGFR (-0.6%).

Similarly, in the unadjusted models, the MORs, which indicate the odds of having a higher risk CMRFs test result for a person from the most, compared to the least, area-level disadvantage, were the highest among eGFR ($\tau 2$=0.189; ICC= 5.4%; MOR = 1.51) and the least among TC ($\tau 2$ =0.025; ICC= 0.8%; MOR = 1.16). The ICCs of CMRFs in all the models were comparatively small (Table 6.6) in all the models. In the fully adjusted models, the ICCs further reduced and ranged between 0.4% and 1.8% in low eGFR and BMI respectively. Table 6.6 presents a summary and comparison of the model fit.

Table 6.6: Summary of model fit values and comparison of the models

		FBSL	HbA1c	TC	HDL	ACR	eGFR	Obesity
Model 1	**Null Model**							
	AIC	111022.8	50114.5	235931.6	130649.7	17130.0	167164.8	242793.2
	τ^2	0.101	0.103	0.025	0.071	0.092	0.189	0.115
	ICC(%)	3.0	3.0	0.8	2.1	2.7	5.4	3.4
	MOR	1.36	1.36	1.16	1.29	1.34	1.51	1.38
Model 2	**Access Model**							
	AIC	110962.3	50105.9	235927.9	130601.4	17119.9	167113.4	242686.2
	τ^2	0.091	0.100	0.025	0.064	0.085	0.176	0.099
	ICC(%)	2.7	3.0	0.8	1.9	2.5	5.1	2.9
	MOR	1.334	1.353	1.163	1.273	1.321	1.492	1.350
	PCV	-9.98%	-2.430%	-1.69%	-9.48%	-7.92%	-6.53%	-14.20%
Model 3	**Sex + Age Adjusted Model**							
	AIC	103066.2	49690.2	227254.6	122700.0	16585.2	115257.1	239122.6
	τ^2	0.103	0.106	0.020	0.081	0.073	0.024	0.117
	ICC(%)	3.0	3.1	0.6	2.4	2.2	0.7	3.4
	MOR	1.36	1.358	1.14	1.31	1.30	1.16	1.39
	PCV	+ 1.88 %	+ 3.02 %	-21.76%	+15.25%	-20.53%	- 87.26%	+1.48%
Model 4	**Sex + Age + IRSD Adjusted Model**							
	AIC	102652.6	49438.2	227193.8	122291.9	16510.8	115109.2	238731.8
	τ^2	0.040	0.048	0.017	0.031	0.028	0.013	0.068
	ICC(%)	1.2	1.4	0.5	0.9	0.9	0.4	2.0
	MOR	1.209	1.231	1.133	1.183	1.175	1.113	1.282
	PCV	-60.90%	-53.78%	-33.11%	-55.90%	-69.14%	-93.31%	-41.21%
Model 5	**Sex + Age + IRSD Adjusted and Access included Model**							
	AIC	102652.0	49440.0	227195.5	122271.4	16511.2	115111.2	238680.6
	τ^2	0.039	0.047	0.017	0.029	0.025	0.013	0.062
	ICC(%)	1.2	1.4	0.5	0.9	0.8	0.4	1.8
	MOR	1.209	1.231	1.133	1.183	1.165	1.113	1.268
	PCV	-61.05%	-53.80%	-33.07%	-59.05%	-72.39%	-93.26%	-46.19%

AIC - Akaike Information Criterion; τ^2 – residual variance; ICC - Intra-cluster Correlation Coefficients; MOR - Median Odds Ratio; PCV - Proportional Change in Variance; FBSL - Fasting Blood Sugar Level; HbA1c - Glycated Haemoglobin ; TC - Total Cholesterol; HDL - High Density Lipoprotein; ACR - Albumin Creatinine Ratio; eGFR - estimated Glomerular Filtration Rate.

6.6 Discussion

The study reports the area-level association between access to primary care service and distribution of CMRFs in the Illawarra-Shoalhaven region of Australia after adjusting for the area-level and individual-level covariates. Access to primary care was inversely associated with low HDL and obesity but was not associated with high FBSL, high HbA1C, high TC, high ACR and low eGFR. The geographic variation of CMRFs explained by primary care access was small and did not demonstrate any attenuating effect on the contribution of area-level disadvantage on the variation of CMRFs in the study region.

Primary care access was only associated with low HDL and obesity in models fully adjusted for individual- and area-level covariates. These findings question previously reported associations between primary care access and improved health.[24-28] However, it should be noted that the current findings pertain only to the geographical/spatial accessibility of the primary health care services within 30 km distance of an SA1 centroid, rather than their actual usage and affordability.

Primary care access was not associated with five out of the seven CMRFs analysed in this study. This was unexpected given the previous research findings, although it is worth noting that most of the previous research indicating inverse associations originated from the United States [24-28], though not all[23, 32, 34]. Related studies reported from Australia did not find any association between the GP supply and preventable hospitalisations after adjusting for sociodemographic and health characteristics of areas in the state of Victoria [58], and found that GP supply was not a significant predictor of preventable hospitalisation in a population sample from NSW[59]. GP supply explained only a small proportion (2.9%) of the geographic variation in hospitalisation rates in individuals aged ≥45 years in that study.[59] Importantly, it was reported in Australia that reductions in preventable hospitalisations are not necessarily associated with improved clinical outcomes[60], but the crude rates may reflect the existing morbidity burden on primary care services [60, 61].

In keeping with previous reports from Australia, the current findings from an adult (≥18 years) sample from the study region in NSW demonstrated only 6.6 % to 10.5% of geographic variation in CMRFs could be attributed to geographic access to primary care services. This finding does not suggest that access to primary is unimportant in the study region, rather highlights the context of general primary health care service in Australia. Australia has universal health insurance and targeted patient benefit schemes such as 'safety net thresholds' to improve low-income and vulnerable populations groups' access to health care services and the annual physician visits per capita of Australia (6.5) is much higher than that of United states (3.9) and the UK (5).[59, 62] It is possible that such strategies have been more effective in Australia in enhancing access to primary care service than in countries such as the United States, however their definitive implications are beyond the scope of the current study.[59] The higher risk CMRFs outcomes observed in the region could be attributable to a range of individual and area-level factors which are not analysed in this study which include individual-level SES, behavioural risk factors, area-level resources, service availability and performance. The observed geographic variation in CMRFs in the study region could also be an actual reflection of the socio-economic and health gradient characteristics intrinsic to this population beyond their access status to primary care services. Previous research suggests that up to 57.8% of geographic variation in CMRFs is attributable to the area-level disadvantage in this study region.[63] Certainly health system performance and the social distribution of wealth could play important roles in explaining the geographic variation of risk factors for the development of CVD.[64] However, the current findings suggest both the complexity of defining primary care access and desirability of future studies to gain more understanding of their performance measures.

Primary care access was inversely associated with low HDL and obesity in the study region. While not disregarding the role of geographic access to primary care and the broader health system in reducing the rates of low HDL and obesity, it should also be noted that the areas in this study region with higher geographic access to primary care service are also the population hubs of the study region, with potentially better access to healthy food sources and exercise facilities which are not analysed in this study. The research literature

suggests that the availability of healthy food sources and exercise facilities can have a direct impact low HDL and obesity in individuals.[65-72] Therefore, the current findings call for future research directions to include more area-level and service-level variables and their interactions with geographic access to primary care in this study region to gain an in-depth understanding on the context of the observed inverse association of primary care access with low HDL and obesity.

The primary care access index, derived from the study region ranged from 0 to 5.41 general partitioners per 1000 people (mean = 2.1, SD = 0.77). Multiple previous studies had reported inequalities in the geographic access to primary care services, using different enhanced versions of 2SFCA method (55, 56, 65, 57–64). For example, the spatial accessibility index derived from rural Otago in New Zealand, using the travel time distance, ranged between 1 to 10; where higher the score indicated a better access.[73] The accessibility index reported from Thimphu district in Bhutan ranged between 0 and 1, where 1 was the maximum access.[74] The spatial accessibility index of GP accessibility in England had been reported to range between 7.2 (South of England) and 13.3 (in London).[75] The access map of the study region (Figure 2) clearly shows a polarisation of the higher access indices along the northern and southern ends of the study region, thus a visible inequality in their distribution. The WHO recommends universal access to primary care for all populations, where geographic access is one part of physical access to primary care.[76]

Area-level disadvantage explained more geographic variation in CMRFs than area-level access to primary care. Inclusion of the access index in the final model did not demonstrate any reduction in the variance explained by area-level disadvantage on the geographic variation of CMRFs. This finding supports the importance of overall socioeconomic development of areas to reduce CMRF risk. Moreover the ICC values of the final models were too small to suggest any meaningful area-level difference in the modelled CMRF variables. This would support the call for universal approaches for the prevention and control of CMRFs rather than any targeted area-level approaches; however with a proportional priority to disadvantaged populations in the study region.[24, 64, 77]

The study has to be considered within its limitations. First, the cross-sectional nature of the study does not support causal inference. Second, the CMRFs data used in this study are from people already utilising health care service in the study area, so care should be taken in generalising the results to the overall population. Third, the study used a radial buffer distance of 30 km for access calculations rather than travel time/distance, because proprietary road network data were unavailable for this study. Usage of actual travel time/distance might have provided a better estimate of access to primary care service locations within this study region. Within the access index calculations, even though the 30 km buffer distance helped to include a maximum coverage of the population in relation with the geographic location of the primary care providers, this distance might have influenced the discriminatory accuracy of the SA1s in the multilevel analyses.

The main strength of this study is the use of a large population-derived database comprising a wide range of CMRFs. The research adds to the very few studies which consider multiple CMRF variables from the same region.[14, 18, 20, 78-80] and is unusual for its hierarchical analysis of the associations between a range of CMRFs and primary care access in a widely dispersed population.

Future research is required to investigate other area-level attributes contributing to the geographic variation of CMRFs in the study region. Our previous research has reported that area-level disadvantage contributes 14.7–57.8% of the geographic variation in CMRFs. The current study extended the previous findings by identifying

the specific contribution of area-level primary care access, ranging between 0.0–10.5%. Further area-level analyses are required to identify other factors contributing to the geographic inequality of the CMRFs in the study region.

6.7 Conclusion

The findings of the study suggest that adults residing in areas that have a poor primary care access are more likely to be identified with low HDL and obesity. However, the specific contribution of area-level primary care access was small when compared with the contribution of area-level disadvantage. The finding supports the importance of overall socioeconomic development of areas to reduce CMRF risk, while supporting universal approaches for the prevention and control of CMRFs which are proportional to the need and disadvantage level of the individuals. Future research including other aspects primary care access such as physical/road network access, financial affordability and acceptance of the services might help to provide an overall picture of the contributing role of primary care in the study region.

References

1. Alkerwi, A., et al., *Geographic variations in cardiometabolic risk factors in Luxembourg.* International journal of environmental research and public health, 2017. **14**(6): p. 648.
2. Paquet, C., et al., *Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia.* International journal of environmental research and public health, 2016. **13**(5): p. 519.
3. Zhou, M., et al., *Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults.* Diabetes care, 2015. **38**(1): p. 72-81.
4. Barker, L.E., et al., *Geographic distribution of diagnosed diabetes in the US: a diabetes belt.* American journal of preventive medicine, 2011. **40**(4): p. 434-439.
5. Lawlor, D., et al., *Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study.* Journal of Epidemiology & Community Health, 2003. **57**(2): p. 134-140.
6. Valdés, S., et al., *Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@ bet. es study.* Revista Española de Cardiología (English Edition), 2014. **67**(6): p. 442-448.
7. Astell-Burt, T., et al., *Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia.* Diabetes research and clinical practice, 2014. **106**(3): p. e68-e73.
8. Congdon, P., *Estimating diabetes prevalence by small area in England.* Journal of Public Health, 2006. **28**(1): p. 71-81.
9. Toms, R., et al., *Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature.* International journal of health geographics, 2019. **18**(1): p. 1.
10. Toms, R., et al., *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia.* PloS one, 2019. **14**(10).
11. Bonney, A., et al., *Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users–a cross-sectional study.* PloS one, 2015. **10**(8).
12. Maier, W., et al., *Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update'GEDA 2009 and 2010.* PloS one, 2014. **9**(2).
13. Mujahid, M.S., et al., *Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics.* Obesity research, 2005. **13**(8): p. 1412-1421.
14. Unger, E., et al., *Association of neighborhood characteristics with cardiovascular health in the multi-ethnic study of atherosclerosis.* Circulation: Cardiovascular Quality and Outcomes, 2014. **7**(4): p. 524-531.
15. Cubbin, C., et al., *Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects.* Scandinavian journal of public health, 2006: p. 228-237.
16. Dragano, N., et al., *Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany.* BMC Public Health, 2007. **7**(1): p. 255.
17. Lawlor, D.A., et al., *Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study.* American journal of public health, 2005. **95**(1): p. 91-97.

18. Roux, A.V.D., D.R. Jacobs, and C.I. Kiefe, *Neighborhood characteristics and components of the insulin resistance syndrome in young adults: the coronary artery risk development in young adults (CARDIA) study.* Diabetes care, 2002. **25**(11): p. 1976-1982.
19. Andersen, A., et al., *Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study.* Diabetic medicine, 2008. **25**(12): p. 1462-1468.
20. Naimi, A.I., et al., *Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area.* International journal of environmental research and public health, 2009. **6**(12): p. 3082-3096.
21. Hiscock, R., et al., *Is neighborhood access to health care provision associated with individual-level utilization and satisfaction?* Health services research, 2008. **43**(6): p. 2183-2200.
22. Schmidt, M.I., et al., *Chronic non-communicable diseases in Brazil: burden and current challenges.* The Lancet, 2011. **377**(9781): p. 1949-1961.
23. Weinberger, M., E.Z. Oddone, and W.G. Henderson, *Does increased access to primary care reduce hospital readmissions?* New England Journal of Medicine, 1996. **334**(22): p. 1441-1447.
24. Kirby, J.B. and T. Kaneda, *Neighborhood socioeconomic disadvantage and access to health care.* Journal of health and social behavior, 2005. **46**(1): p. 15-31.
25. Basu, J., B. Friedman, and H. Burstin, *Primary care, HMO enrollment, and hospitalization for ambulatory care sensitive conditions: a new approach.* Medical care, 2002: p. 1260-1269.
26. Chang, C.-H., et al., *Primary care physician workforce and Medicare beneficiaries' health outcomes.* Jama, 2011. **305**(20): p. 2096-2104.
27. Laditka, J.N., S.B. Laditka, and J.C. Probst, *More may be better: evidence of a negative relationship between physician supply and hospitalization for ambulatory care sensitive conditions.* Health services research, 2005. **40**(4): p. 1148-1166.
28. Parchman, M.L. and S.D. Culler, *Preventable hospitalizations in primary care shortage areas: an analysis of vulnerable Medicare beneficiaries.* Archives of family medicine, 1999. **8**(6): p. 487.
29. Larkins, S., et al., *Addressing inequities in access to primary health care: lessons for the training of health care professionals from a regional medical school.* Australian Journal of Primary Health, 2011. **17**(4): p. 362-368.
30. Angier, H., et al., *Using geographic information systems (GIS) to identify communities in need of health insurance outreach: an OCHIN practice-based research network (PBRN) report.* The Journal of the American Board of Family Medicine, 2014. **27**(6): p. 804-810.
31. Atallah, A., et al., *Reducing the burden of arterial hypertension: what can be expected from an improved access to health care? Results from a study in 2420 unemployed subjects in the Caribbean.* Journal of human hypertension, 2007. **21**(4): p. 316-322.
32. Kotchen, J.M., et al., *Hypertension control and access to medical care in the inner city.* American Journal of Public Health, 1998. **88**(11): p. 1696-1699.
33. Occelli, F., et al., *Mapping end-stage renal disease (ESRD): spatial variations on small area level in northern France, and association with deprivation.* PloS one, 2014. **9**(11).
34. Harris, M.I., *Racial and ethnic differences in health care access and health outcomes for adults with type 2 diabetes.* Diabetes care, 2001. **24**(3): p. 454-459.

35. Pattenden, S., et al., *Geographical variation in infant mortality, stillbirth and low birth weight in Northern Ireland, 1992–2002.* J Epidemiol Community Health, 2011. **65**(12): p. 1159-1165.
36. Australian Bureau of Statistics. *2011 Census data* [cited 2020 Sept 5]; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280.
37. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Census Dictionary, 2011*. 2011; Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2901.0Chapter23102011.
38. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Volume 1 - Main Structure :STATISTICAL AREA LEVEL 1 (SA1).* 2016; Available from: https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+(ASGS).
39. Ghosh, A., et al., *Using data from patient interactions in primary care for population level chronic disease surveillance: The Sentinel Practices Data Sourcing (SPDS) project.* BMC Public Health, 2014. **14**(1): p. 557.
40. Ghosh A, M.K., Marshall K. . *Illawarra-Shoalhaven Medicare Local - Population Health Profile: 2013.* 2013; Available from: https://www.gph.org.au/assets/Main-Site/Uploads/Resources/Improving-population-health/ISML-Population-Health-Profile-2013-FINAL.pdf.
41. Toms, R., et al., *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia.* PloS one, 2019. **14**(10): p. e0223179.
42. Australian Bureau of Statistics. *Census data* 2011; Available from: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280.
43. The Royal Australian College of General Practitioners & Diabetes Australia. *General Practice Management of Type 2 Diabetes 2016-2018.* . The Royal Australian College of General Practitioners 2016; Available from: Doi: 10.1007/s00125-010-2011-6.
44. National heart foundation of Australia. *Lipid management profile for health professionals.* . Available from: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management.
45. National Kidney foundation(USA). *Albumin creatinine Ratio (ACR).* 2018; Available from: https://www.kidney.org/kidneydisease/siemens_hcp_acr.
46. World Health Organization. *Obesity : Preventing and managing the global epidemic: Technical Report Series. WHO Technical Report Series, no. 894.* 2000; Available from: doi:ISBN 92 4 120894 5.
47. Luo, W. and F. Wang, *Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region.* Environment and Planning B: Planning and Design, 2003. **30**(6): p. 865-884.
48. Vo, A., M. Plachkinova, and R. Bhaskar, *Assessing healthcare accessibility algorithms: A comprehensive investigation of two-step floating catchment methodologies family.* 2015.
49. McGrail, M.R., *Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements.* International journal of health geographics, 2012. **11**(1): p. 50.
50. Australian Bureau of Statistics. *Main Features - IRSD.*; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011.
51. Goldstein, H., W. Browne, and J. Rasbash, *Partitioning variation in multilevel models.* Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences, 2002. **1**(4): p. 223-231.
52. Szmaragd, C. and G. Leckie, *Module 6: Regression Models for Binary Responses R Practical.* 2011.
53. Environmental Systems Research Institute (ESRI). *ArcGIS 10.4.1, Redlands, CA, USA, ESRI Inc,*

. Available from: https://www.esri.com/.

54. R Core Team. *R: A language and environment for statistical computing.R Foundation for Statistical Computing,Vienna, Austria*. 2018; Available from: https://www.R-project.org/.
55. Bates, D., et al., *Fitting linear mixed-effects models using lme4.* arXiv preprint arXiv:1406.5823, 2014.
56. Zeileis, A. and T. Hothorn, *Diagnostic checking in regression relationships.* 2002.
57. AUSTRALIAN BUREAU OF STATISTICS. *Technical Paper: Socio-Economic Indexes for Areas (SEIFA).* 2011; Available from: https://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/22CEDA8038AF7A0DCA257B3B00116E34/$File/2033.0.55.001%20seifa%202011%20technical%20paper.pdf.
58. Ansari, Z., J.N. Laditka, and S.B. Laditka, *Access to health care and hospitalization for ambulatory care sensitive conditions.* Medical care research and review, 2006. **63**(6): p. 719-741.
59. Falster, M.O., et al., *Sociodemographic and health characteristics, rather than primary care supply, are major drivers of geographic variation in preventable hospitalizations in Australia.* Medical care, 2015. **53**(5): p. 436.
60. Katterl, R., et al., *Potentially avoidable hospitalisations in Australia: Causes for hospitalisations and primary health care interventions.* 2012.
61. Giuffrida, A., H. Gravelle, and M. Roland, *Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes.* Bmj, 1999. **319**(7202): p. 94-98.
62. Thomson, S., et al., *International profiles of health care systems 2012: Australia, canada, denmark, england, france, germany, iceland, italy, japan, the netherlands, new zealand, norway, sweden, switzerland, and the united states.* 2012.
63. Toms, R., et al., *Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia.* Scientific Reports, 2020. **10**(1): p. 1-18.
64. Walsh, M.G., et al., *The socioeconomic correlates of global complication prevalence in type 1 diabetes (T1D): a multinational comparison.* Diabetes research and clinical practice, 2005. **70**(2): p. 143-150.
65. Murillo-Castillo, K.D., et al., *Food insecurity was associated with low quality diet and low HDL level in mothers of Northwest Mexico relying on fisheries for livelihood.* Nutr Hosp, 2018. **35**(6): p. 1379-1386.
66. Shin, J.-I., et al., *Food insecurity and dyslipidemia in a representative population-based sample in the US.* Preventive medicine, 2015. **77**: p. 186-190.
67. Shariff, Z.M., et al., *Food insecurity and the metabolic syndrome among women from low income communities in Malaysia.* Asia Pacific journal of clinical nutrition, 2014. **23**(1): p. 138-147.
68. Wedick, N.M., et al., *Access to healthy food stores modifies effect of a dietary intervention.* American journal of preventive medicine, 2015. **48**(3): p. 309-317.
69. Gordon-Larsen, P., *Food availability/convenience and obesity.* Advances in nutrition, 2014. **5**(6): p. 809-817.
70. Lee, H., *The role of local food availability in explaining obesity risk among young school-aged children.* Social science & medicine, 2012. **74**(8): p. 1193-1203.
71. Odoms-Young, A.M., S. Zenk, and M. Mason, *Measuring food availability and access in African-American communities: implications for intervention and policy.* American journal of preventive medicine, 2009. **36**(4): p. S145-S150.
72. Mau, M.K., et al., *Environmental factors of obesity in communities with native Hawaiians.* Hawaii medical journal, 2008. **67**(9): p. 233.

73. Gruen, R.L., et al., *Specialist outreach clinics in primary care and rural hospital settings.* Cochrane Database of Systematic Reviews, 2003(4).
74. Jamtsho, S., R. Corner, and A. Dewan, *Spatio-temporal analysis of spatial accessibility to primary health care in Bhutan.* ISPRS International Journal of Geo-Information, 2015. **4**(3): p. 1584-1604.
75. Bauer, J., et al., *Spatial accessibility of primary care in England: A cross-sectional study using a floating catchment area method.* Health services research, 2018. **53**(3): p. 1957-1978.
76. Evans, D.B., J. Hsu, and T. Boerma, *Universal health coverage and universal access*. 2013, SciELO Public Health.
77. Rose, G., *Sick individuals and sick populations.* Int. J. Epidemiol, , 1985.
78. Gabert, R., et al., *Identifying high-risk neighborhoods using electronic medical records: a population-based approach for targeting diabetes prevention and treatment interventions.* PLoS One, 2016. **11**(7).
79. Clark, C.R., et al., *Neighborhood disadvantage, neighborhood safety and cardiometabolic risk factors in African Americans: biosocial associations in the Jackson Heart study.* PloS one, 2013. **8**(5).
80. Keita, A.D., et al., *Associations of neighborhood area level deprivation with the metabolic syndrome and inflammation among middle-and older-age adults.* BMC Public Health, 2014. **14**(1): p. 1319.

Chapter 7

General Discussion and Conclusion

Chapter 7: General Discussion and Conclusion

7.1 Introduction

The thesis presents an epidemiological and geospatial analysis of the distribution of cardiometabolic risk factors (CMRFs) in the Illawarra-Shoalhaven region of NSW, Australia. It has presented the geographic variation in the distribution of individual CMRFs in the study region and has demonstrated the associations of area-level disadvantage and access to primary care with this variation. Additionally, it demonstrates the utility of routine clinical data for research concerning location specific population health approaches and for informing regional health care service commissioning.

The publications arising from this study have contributed to the knowledge regarding the geographic variation of CMRFs and the contextual factors associated with CMRF distribution. The findings are important in the context of the global paradigm shift from infectious disease to non-communicable diseases as key drivers of illness burden, especially cardiovascular diseases (CVD) as the prime cause of death and health care expenditure worldwide.

The final chapter of the thesis provides: an overall outline of achievement of the objectives of the study; a critical in-depth comparison of the findings of the study with international literature; and a discussion of the significance of the findings and future research directions. The chapter also discusses theoretical considerations along with the study's methodological strengths and limitations, prior to presenting the conclusions.

7.2 Major findings

The study sought to investigate the geographic distribution of CMRFs and their associations with area-level disadvantage and access to primary care in the study region. This was achieved by answering the following research questions sequentially:

1. What is the existing level of evidence on the geographic and socioeconomic variation in the distribution of CMRFs internationally?
2. What is the small-area level geographic distribution pattern of cardiometabolic risk factors, within the Illawarra-Shoalhaven region of NSW Australia?
3. What proportion of any geographic variability in cardiometabolic risk factor prevalence is due to small-area level socioeconomic status, within the Illawarra-Shoalhaven region of NSW Australia?
4. What proportion of any geographic variability in cardiometabolic risk factor prevalence is due to differences in small-area level primary care access, within the Illawarra-Shoalhaven region of NSW Australia?

The study answered the above research questions in a hierarchical manner, with the following outcomes:

1. Systematic literature review of 24 eligible studies from multiple nations across the world revealed variation in the distribution of CMRFs at varying geographic scales. Among these, 16 studies demonstrated consistent associations between area-level socioeconomic disadvantage and higher prevalence of various CMRFs. These reports of associations were mostly independent of individual-level factors and reported mainly from industrialised nations.
2. There was significant geographic variation in the distribution of individual CMRFs in the study region in Australia. The variation included clustering of higher risk categories of: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI); and diabetes mellitus (DM) status. The High-High (HH) clusters of CMRFs were found mainly along the highly populated eastern seaboard of the study region, where as the Low-Low (LL) clusters were predominantly in the less populated northern, central and southern areas of the study region.
3. Area-level socioeconomic disadvantage was associated with all the analysed CMRFs after adjusting for individual-level covariates such as age-group and sex. The estimated proportion of the geographic variation in the higher risk CMRFs' distribution explained by area-level socioeconomic disadvantage varied across the CMRFs. Area-level disadvantage explained 57.8% of the geographic variance in low HDL, 57.1% of variance of high FBSL, 53.3% of variance of high HbA1c, 51.2% of variance of high ACR, 41.8% of variance of low eGFR, 41.1% of variance of BMI and 14.7% of variance of high TC test results. These findings demonstrated a consistent burden of multiple higher risk CMRFs concentrated in areas of greater socioeconomic disadvantage.
4. A primary care access map of the study region was plotted; and access was found to be lower along the sparsely populated central and western areas of the study region. The available primary care services were mainly concentrated along the populated eastern seaboard of the study region. Better primary care access was associated with lower levels of higher risk HDL and obesity. No significant area-level associations were found between primary care access and the remaining CMRFs, including high FBSL, high HbA1c, high TC, high ACR and low eGFR. The geographic variation in CMRFs explained by access to primary care was <=10.5%.

The results have the potential to inform local and regional health care service planning and policy developments; and generate hypotheses internationally for area-level research based in similar settings.

7.3 General discussion

This section presents an overall discussion of the individual findings in this thesis. The section is divided into four subsections, to facilitate the presentation of the discussion based on the individual objectives of the study.

7.3.1 Systematic Literature review on the distribution of CMRFs

In study 1, I had systematically reviewed the existing literature reported globally on geographic and area-level socioeconomic variation in the distribution of CMRFs. Geographic variation in the presence of one

or more CMRFs was reported from multiple nations. Recent advances in GIS and analytical approaches were utilised in the studies reporting geographic variation in CMRFs.

Area-level socioeconomic disadvantage was repeatedly demonstrated to be associated with higher cardiometabolic risk.[1] Higher disadvantage was also consistently reported to have an association with cardiovascular risk; whereas lower disadvantage was associated with reduced cardiovascular risk. Such associations were often demonstrated independently of individual level characteristics such as socioeconomic status, education and duration of exposure to area.[1]

The systematic review included only biological proxies of CMRFs. This is mainly due to the overall focus of this thesis on biological proxies of CMRFs, and the review was also intended to provide a background for their analyses. The thesis did not have any data for the examination of behavioral, dietary and activity related risk factors of individuals, hence these were not included in the review.

Related systematic reviews published in this area of research had investigated associations for different geographically distributed factors with CVD. Chaix (2009) reviewed the associations between neighbourhood social environments and CHD and proposed a theoretical model of a mediating mechanism focussing on the social interactional environment.[2] Consistent associations of obesity or hypertension with lower levels of area socioeconomic status, urbanisation, street intersection, accessibility to supermarkets, social cohesion, service availability and residential density and higher levels of noise pollution and density of convenience stores, were reviewed and reported by Leal (2011).[3] Frequent inverse associations of common indices of area-level socioeconomic disadvantage with childhood obesity were reported in the UK.[4] A consistent association between socioeconomic disadvantage and central adiposity was reported by Slopen (2013).[5] All these reviews reported important methodological limitations and the need for further research in this area, which support the findings of the current review.

Overall, the review indicated that disease patterns at smaller areas in a nation may significantly differ from national and regional prevalence reports. Thus small-area analysis is important in order to understand local patterns and requirements.[6] Small-area level analyses also have the potential to reveal area level contexts and dependencies of CMRFs and such analyses can highlight areas for targeted preventive interventions.[1]

7.3.2 Geographic variation in the distribution of CMRFs in the study region

Study 2 explored the geographic distribution of eight CMRFs in 980 conterminous geographic units within the Illawarra-Shoalhaven region of the NSW, Australia.[7] Higher rates and clustering of higher risk CMRFs were mostly observed along the more densely populated eastern seaboard of the study region. However, not all populated areas were involved in this pattern and less populated areas were also involved with higher rates of CMRFs, especially that of high TC. Spatial analyses had revealed significant spatial autocorrelation of all eight CMRFs and the clustering locations observed were different for each of the CMRFs.[7]

The distribution of high TC values were generally reversed to those distributions described for other CMRFs.[7] The reasons for this observation are yet to be explored, but a possible prescribed medication treatment effect was suspected as the lower risk areas were often densely populated areas. It is possible that people residing in these areas had better access to health care services and more frequently prescribed

cholesterol-lowering drugs.[8, 9] However, not all densely populated areas were involved in this TC distribution pattern, so further research is required to identify the area-level factors contributing to this paradoxical distribution pattern of TC.

The findings regarding the geographic variation in the distribution of CMRFs are supported by previous reports internationally.[10-17] In the UK, geographic variation in the prevalence of risk factors such as obesity, smoking, diabetes, hypertension and high cholesterol were reported across four main regions: South England; Midlands and Wales; Scotland; and North England.[17] A higher prevalence of CMRFs was reported in southern Spain (Andalusia), which was found in close association with sedentary lifestyle and markers of socioeconomic disadvantage, after adjusting for individual level covariates.[14]Variation in the distribution of diabetes, high BMI (≥ 25 kg/m^2), abdominal obesity, hypertension, high cholesterol and low glomerular filtration rate were reported at both canton and municipality levels in Luxemburg, Western Europe.[10]

The contribution of the current study is its ability to describe a wide range of CMRFs simultaneously across the same time period and same geographic region. In addition, it is one of few studies reporting the distribution of multiple parameters from regional Australia. Previous studies from Australia had reported geographic variation in being diagnosed with DM among adults living in Sydney.[11] Another study reported geographic variation in glycated haemoglobin (HbA1c) values across 767 Census Collection Districts (CDs) in Adelaide.[13] The current finding builds on previous research by investigating the distribution of a wide range of CMRFs, across population catchments including urban, rural and semirural areas. The evidence of the small-area level variation of multiple CMRFs in this study is significant and consistent across all CMRFs, which included clustering of the pathology test results of high FBSL, high HbA1c, high TC, low HDL, high ACR, low eGFR, obesity and diabetes.

The study focusses on the geographic distribution of eight individual CMRFs rather than the aggregate count of risk factors in one person and their subsequent analyses and mapping. The latter method might be helpful in generating a 'single map' of 'geographic variations in SA1-level per capita counts of "higher risk" risk factors for the region' to indicate areas of higher CVD risk. But the 'single map' will not tell which risk factor is high in an area and what kind of health care resources are required in those areas to address these risks. As one of the main intentions of this program of research was to inform regional health care service commissioning and their resource allocation, it was important that the study outcomes aligned with existing clinical practice to make the results potentially useful for any further planning and implementation strategies.[18] What is practiced in a clinical context is the treatment and management of the individual components of CMRFs. [18] Also, it should be noted that the study sample is a cohort of pathology service users who may not be representative of the SA-1 populations within the study area, which also makes it methodologically more appropriate to use within-cohort denominators.

The findings of the study enabled comparisons of the areas which have demonstrated clustering of various CMRFs. In addition to providing useful results demonstrating areas of significant geospatial clustering of various CMRFs for area-level health care commissioning, the results also facilitated generation of hypothesis for further research on the association of area-level factors which attribute to this clustering.

Regarding the location of the identified high-high clusters of higher risk CMRFs findings, subsequent stage 3 analyses (presented in Chapter 5) demonstrated that the 'higher risk' CMRFs findings were

positively associated with area-level socioeconomic disadvantage, except for TC which demonstrated an inverse trend. This finding can also be retrospectively applied to chapter 4 results to see that deprived areas and their neighbourhoods are more likely to be identified with high-high clusters of 'higher risk' CMRFs, except TC results. For illustrative purposes, an explicit mapping of the areas of high-high clustering and higher area-level disadvantage using the study data set is presented here (Fig 7.1). In the maps, it can be observed that areas of high-high clustering are more likely to be areas of higher area-level disadvantage though not absolutely, which provides a thoughtful link to the next stage findings presented in this chapter.

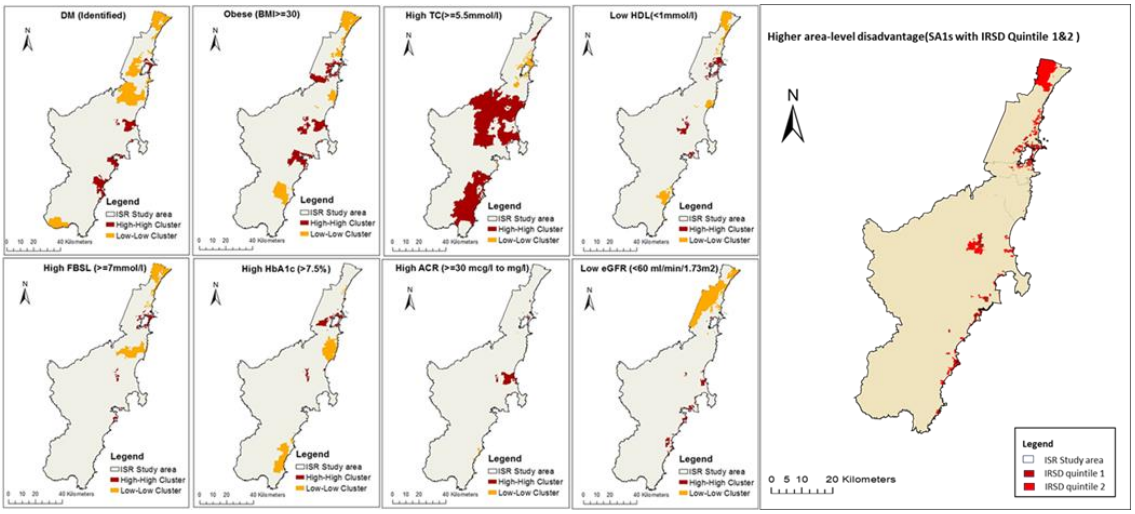

Figure 7.1: Local Moran's I cluster maps showing high-high and low-low spatial associations of CMRFs; and the areas of higher area-level disadvantage, within the Illawarra Shoalhaven region of the NSW Australia.

7.3.3 Association of area-level disadvantage with all CMRFs

Study 3 found consistent evidence for association between area-level disadvantage and CMRFs in the study region.[19] In the individual-level age and sex adjusted models, the odds of a higher risk CMRF finding increased with increasing area-level disadvantage among all CMRFs except TC. Total cholesterol alone demonstrated an inverse pattern of association with increasing area-level disadvantage. The contribution of area-level disadvantage to the observed geographic variance in CMRFs was the highest for low HDL (57.8%), followed by high FBSL (57.1%); high HbA1c (53.3%); high ACR (51.2%); low eGFR (41.8%); and obesity (41.1%) test results. The contribution of area-level disadvantage was comparatively less for the geographic variance of high TC (14.7%).[19]

The TC test results stood apart from the other findings of this study. However, the HDL findings were not consistent with the area-level findings of TC results, even though both are components of the lipid profile. This suggested the possibility of a medication effect on TC in these areas, where the lipid lowering drugs have a less consistent effect in raising HDL than in lowering TC.[20] However the reason for the inverse association demonstrated by TC test results were not within the scope of the current research, but the non-consistent patterns explains possible role of confounders not identified within the study.

The eGFR models demonstrate a very high magnitude of ORs in the higher age groups in Model 2 and Model 4. In table 5.4, where the frequency and proportion of individual CMRF test findings are reported,

it can be found that the proportion of higher risk eGFR findings in the reference age group 18-29 years is very low (0.03%) in comparison with the highest 80+ years age group (5.52%), which is around is 184 times higher and may well be the order of magnitude the ORs suggest. Therefore it is likely that these much-magnified ORs found in the higher age groups of eGFR test findings in Model 2 and Model 4 are due to the effect of very low proportion of higher risk eGFR findings in the reference age group and very high number of findings in the upper age groups. I have presented a sensitivity test for using different reference age groups, in Table 7.1. As shown in the table, changing the reference group only makes a relative difference in their ORs, but not any change in their corresponding interpretation.

Table 7.1: Sensitivity analysis of the Odd's ratios of low eGFR, using different age reference groups.

Ref group:	OR 18-29 years	OR 30-39 years	OR 40 -49 years	OR 50 -59 years	OR 60 – 69 years
Intercept	0.00	0.00	0.01	0.03	0.09
18-29 years	Reference	0.60	0.23	0.08	0.02
30-39 years	2	Reference	0.39	0.14	0.03
40 -49 years	4	2.57	Reference	0.35	0.09
50 -59 years	12	7.40	2.88	Reference	0.27
60 – 69 years	42	25.26	9.82	3.41	Reference
70-79 years	151	91.02	35.38	12.29	3.45
80+ years	509	307.64	119.56	41.55	11.67

In addition, it is also noted that combining 18-29 and 30-39 years as a reference group, only makes a relative difference in their ORs but not any absolute or interpretive difference. This is because the proportion of higher risk eGFR findings in the 30-39 years age group is lower (0.04%) too. The original and modified Table 5.10 is copied below in Table 7.2 for reference purposes, and the changed values are marked in red.

Table 7.2: A comparison of the Multilevel logistic regression model summaries for the area-level analyses of the association between low eGFR (eGFR < 60 mL/min/1.73m^2) and relative socioeconomic disadvantage, with the original and combined age groups.

Variables	Model 1 ***	Model 2 ***	Model 3 ***	Model 4 ***
Low eGFR	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.13(0.12-0.13)***	0.00(0.00-0.00)***	0.08(0.07-0.08)***	0.00(0.00-0.00)***
Sex: Female		Reference		
Male		0.98(0.95-1.01)^{NS}		0.98(0.95-1.01)^{NS}
Age:18—29		Reference		1.66(1.24-2.22)***
30—39		1.66(1.24-2.20)***		4.30(3.36-5.49)***
40—49		4.26(3.34-5.42)***		12.32(9.80-15.47)***
50—59		12.26(9.78-15.35)***		41.86(33.49-52.31)***
60—69		41.81(33.55-52.1)***		149.53(120-186.6)***
70—79		150.66(121-187.6)***		501.47(401.7-626)***
80+		509.18(409-633.9)***		
IRSD : Q-5			Reference	
Q-4			1.23(1.12-1.35)***	1.09(1.03-1.16)***
Q-3			1.59(1.45-1.74)***	1.19(1.13-1.26)***
Q-2			1.65(1.51-1.81)***	1.22(1.15-1.29)***
Most D Q-1			1.97(1.80-2.15)***	1.38(1.31-1.46)***
AIC	167164.8	115257.1	166930.0	115125.7
Variance	0.189	0.024	0.138	0.014
PCV	-	- 87.26%	-26.84%	-92.79%
ICC(%)	5.4	0.7	4.0	0.4
MOR	1.51	1.16	1.43	1.12
Proportional variance explained by IRSD : 41.75%				

Variables	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***
Low eGFR	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.11(0.11-0.12)***	0.00(0.00-0.00)***	0.08(0.07-0.08)***	0.00(0.00-0.00)***
Sex: Female		Reference	--	--
Male		0.99(0.96-1.02)^{NS}		0.99(0.96-1.02)^{NS}
Age:18—39		Reference	--	--
40—59		6.56(5.61-7.67)***		6.60(5.67 – 7.68)***
60—79		64.56(55.64-74.90)***		64.33(55.70-74.29)***
80+		385.85(332.27-448.07)***		379.93(398.71-439.12)***
IRSD : Q-5			Reference	--
Q-4			1.23(1.12-1.35)***	1.10(1.04 -1.17) ***
Q-3			1.59(1.45-1.74)***	1.24(1.17 -1.31) ***
Q-2			1.65(1.51-1.81)***	1.28(1.21 -1.35) ***
Most D Q-1			1.97(1.80-2.15)***	1.46(1.38 -1.54) ***
AIC	167164.8	119484.2	166930.0	119301.2
Variance	0.189	0.030	0.138	0.015
PCV	-	- 84.27%	-26.84%	-92.05%
ICC(%)	5.4	1.0	4.02	0.45
MOR	1.51	1.18	1.43	1.12
Proportional variance explained by IRSD : 61.07%				

**** - p<0.001; eGFR - estimated Glomerular Filtration Rate; IRSD - Index of Relative Socioeconomic Disadvantage; Most D – Most Disadvantaged; LRT – Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1+ Area level: IRSD quintiles of SA1s; Model 4—Model 1+Model 2 + Model 3; NS – Not significant; SA1 — Statistical area-level 1.*

The ICCs of the final models were low and suggest very limited area-level contextual effects and support a homogenic contribution of areas proportional to the level of disadvantage. The ICCs of CMRFs in all the models were comparatively small in all the models. In fully adjusted models, the ICCs were further reduced and ranged between 0.4% and 2.0% in low eGFR and BMI respectively. As per the interpretation framework proposed by Merlo et al (2019), an ICC value less than 10% is indicative of very little geographic difference.[21] However, this has to be interpreted along with the traditional geographic comparisons such as the proportion of the individuals who are affected with higher risk CMRF outcomes. Therefore, a small geographic difference indicates homogeneity of the higher risk CMRF findings within geographic units.[21] Such a situation would call for balanced universal approaches to prevent and control the higher risk CMRFs, with a proportional focus to the need and disadvantage level of affected populations.[22, 23] However, it is also worth noting that when the exposure to an agent is homogenic in a community, traditional epidemiological methods are not very helpful in identifying markers of susceptibility.[24]

The findings of the study add to the existing evidence on the association between area-level disadvantage and CMRFs through the fixed effect estimates of the multilevel regression models.[19, 25, 26] Previous research from Australia demonstrated inverse and positive associations of area-level education and area-level income respectively with cardiometabolic syndrome in a prospective cohort [26], and advocated population interventions based on area-level socioeconomic disadvantage to reduce cardiometabolic risks[25]. The study extends previous reports by providing evidence for a range of CMRFs regarding their association with area-level disadvantage. The findings thus demonstrate the occurrence of multiple higher risk CVD risk factors in disadvantaged areas. Importantly, the specific contributions of area-level disadvantage to the geographic variance of multiple CMRFs were identified through the random effect estimates of the models, which is unique in the literature and highly informative for area-appropriate health care service commissioning.

7.3.4 Association of area-level primary care access with CMRFs

Study 4 was done in two stages. Stage 1 focussed on deriving the area-level primary care access index of the study region, whereas stage 2 aimed to analyse area-level associations between primary care access and the CMRFs.

The access index was higher mainly along the eastern seaboard of the study region, but low and <=1.6 per 1000 people along the western and central areas which were relatively less densely populated. A total of 165 primary care service locations with 611 general practitioners were identified within the study region in 2016. The service locations were mostly along the populated eastern coastal strip of the study region. The primary care access index of the study region was derived using a two-step floating catchment area method (2SFCA), which accounted for both the supply (available providers) and demand (based on the total population) of the primary care services within a specified spatial distance.[27] After sensitivity analyses, a spatial buffer distance of 30 km was chosen to measure the primary care access, from both the geographic centroids of the small-areas within the study region and the service provider locations. This distance was observed to provide a better coverage of the geographic centroids and services providers, considering the sparsely populated areas and the distant locations of the primary care providers. The derived primary care access index of the small-areas within the study region ranged between 0 and 5.4 general partitioners per 1000 people (mean = 2.1, SD = 0.77).

The area-level associations between access to primary care service and distribution of CMRFs in the study region were explored. Area-level primary care access inversely associated with low HDL and obesity after adjusting for individual-level age, sex, and area-level disadvantage. Primary care access was not associated with any of the remaining CMRFs, including high FBSL, high HbA1c, high TC, high ACR and low eGFR. Comparison of the models based on measures of are-level variation demonstrated that area-level primary care access explained 10.5% of the geographic variation in high ACR, followed by 8.5% variation of obesity and 6.6% variation of low HDL. The geographic variation explained by primary care access was close to zero for high FBSL (0.4%), high HbA1c (0.0%), high TC (-0.1%) and eGFR (-0.6%). %). Thus, the observed geographic variation in CMRFs explained by area-level primary care access is found to be minimal, especially in comparison with variation explained by area-level disadvantage which was up to 57.8 % in the previous study. The findings are consistent with the previous related reports from Australia. This does not undermine the crucial role of primary care access in the study region but does reflect both the complexity of defining access and desirability of future research on the performance measures of primary care in the study region. Also, it is important to note that even though no association was found between geographic access to primary care and most of the CMRFs, the unadjusted crude rates of higher risk presented in this study directly indicate the existing burden on the primary care system in the study region and the related requirement of resources and their appropriate delivery to effectively prevent and control CMRFs within the region. Thus, irrespective of the null association findings reported in this study, sufficient resources and their supply would be required to effectively control and manage the observed crude rates of CMRFs in the study region. Effective control and management of existing CMRFs in the study region is expected to reduce the rates of higher risk CMRFs in the study region and thus their CVD related hospitalisation events and their further burden on the health care system – which is the leading cause of death and health care expenditure in Australia.

These results may also result from contextual factors contributing to the geographic variation of CMRFs, over and beyond primary care access. Area-level disadvantage explained more geographic variation in CMRFs than area-level access to primary care. Inclusion of the access index in the final model did not demonstrate any reducing effect on the variance explained by the area-level disadvantage on the geographic variation of CMRFs. This finding supports the importance of overall socioeconomic development of areas to reduce CMRF risk.[22-24] Future research is required to investigate other area-level attributes contributing to the geographic variation of CMRFs in the study region. This research has reported that area-level disadvantage contributes 14.7–57.8% of the geographic variation in CMRFs. My current findings extended the previous findings by identifying the specific contribution of area-level primary care access, ranging between 0.0–10.5%. Further area-level analyses are required to identify other factors contributing to the geographic variation of the CMRFs in the study region.

7.4 Theoretical underpinnings of the findings

This section presents a general discussion on the links of study findings to existing theories. The discussions are based on the concept that the phenomena of the population-level patterns of health are not naturally occurring random events, but are underpinned to how societies are organized.[28] Two theories are discussed in this section in comparison with the major components and outcomes of the individual studies in this thesis.

7.4.1 Socio-ecological theory

The evidence from this thesis is consistent with Socio-ecological theory, proposed by Urie Bronfenbrenner in 1989.[29] The theory was further developed by Daniel Stokols in 1996 to explain the dynamic interrelations among individuals and environmental factors.[30] The thesis findings are most consistent with Stokols's application of Socio-ecological theory.

As per Daniel Stokols, Socio-ecological theory highlights the dynamic relationship between individuals and their surroundings.[31] The contextual factors of an individual can be linked with the aetiology of multiple diseases and risk factors, including CVD. Socio-ecological analyses examine the day to day exposures of individuals to various contextual factors, such as social, demographic and physical environments. Socio-ecological theory suggests that certain environmental contexts within an individual's life situation can exert disproportionate effects on their health and wellbeing. These influential settings can be viewed as high impact 'leverage points' for enhancing or reducing one's health and well-being. Further, Stokols emphasises that rather than allocating large amounts of resources to modify individual-level behaviours, it is sometimes more effective to focus and change the 'health intermediaries', such as policies and decisions to facilitate the desired individual level outcomes.[31]

The study findings are supported by Stokol's Socio-ecological theory applications.[31] Area-level disadvantage contributes to the social environments of individuals living in that area. Area-level disadvantage (or more specifically its components) may constitute common 'leverage' points for the multiple CMRFs analysed in this study. The geographic access to primary care can be considered as the physical environment which has an influencing effect on the ongoing prevention and management of all the CMRFs analysed in this study. The implications of the findings of this study are bi-directional in relation to the past and future policies which determine these environments. Therefore, the related policies in the past might have caused this finding and the current findings have the potential to contribute evidence for the related policy changes in future.

7.4.2 Ecosocial theory

The findings of the study are also supported by the Ecosocial theory proposed by Nancy Krieger of the Harvard School of Public Health in 1994.[32] The theory specifies that distributions of disease and risk factors are determined at multiple levels and its analyses must incorporate all the possible levels. Further, the theory assumes that all factors in the multiple levels must be considered in concert, as they work cohesively in a synergistic way in explanation of the risk distribution.

The key constructs of Ecosocial theory are: 1) embodiment; 2) pathways to embodiment; 3) the cumulative interplay of exposure, susceptibility and resistance; and the 4) agency and accountability. Embodiment denotes the incorporation of the social and biological world into an individual's body. Pathways to embodiment imply various contextual ways which interplay with the embodiment. Pathways to embodiment can affect an individual through various spatio-temporal scales across the life course and can be expressed at multiple levels including individual, community and population levels.[33]

Cumulative Interplay explains how people with different contextual factors have different susceptibility and risk factors to diseases, mainly based on their unique spatio-temporal factors, interaction with the groups and systems and the discrimination and inequality faced throughout their life course.[33] Agency and Accountability argue that the State is a responsible agent in the patterns of disease distribution in a given

society.[33] Thus, the theory suggests that the social system that creates discrimination and inequalities are also responsible for the patterns of disease observed in a society.[33]

Further, Krieger expands the last construct to include that the accountability of epidemiologists and public health researchers in the identification and reporting of the health disparities. This obligation is envisaged as a call to become an activist, rather than researchers, when injustice is observed in inequity.[33]

The current study aligns with the Ecosocial theory of public health in multiple ways. The population distributions of CMRFs were identified in the study through describing the geographic variation in CMRF distribution in the study region. Social determinants or gradients at area-level were used as both predictors and covariates at different stages of the study, through the use of area-level Index of Relative Socioeconomic Disadvantage (IRSD). Differences in area-level access to primary care can also be seen as an area-level gradient, which is included as an explanatory variable in the study. Gender differences were adjusted at all stages of the study as an individual-level covariate. The study also accounts for the levels in the theory and had adopted a multilevel analysis in respect to the nesting of individual data within areas and thus within the study region. Overall, the study has implications regarding the Accountability of the State Agency and recommends changes in the political, economic and ecosystem environments to reduce the observed variation in multiple CMRFs in the study region.

7.5 Significance of the research

The findings of the study contribute evidence for practice and policy developments at regional, national and international levels.

a) Regional level

The findings from this study will contribute to the planning of area-level prevention and control of CMRFs, which is important in the context of CVD being the prime cause of death and health care expenditure in many industrialised nations, including Australia. The study also demonstrates a feasible approach for using population derived regional data for informing the planning and resource allocation of the health care services of the same region. Centralised approaches of prevention may not always suit regional requirements, but the use of local data can provide evidence for regional health care service planning and related policy developments.

b) ***National level***

In Australia, previous reports have acknowledged that the health inequalities experienced by Australians are shaped by their broader socioeconomic circumstances.[34] Australians living in poor socioeconomic areas were reported as subjected to: early death [35]; higher risk for heart disease and diabetes [36]; higher mortality rate [37]; poor mental and physical health [38]; and to have mothers with low birthweight babies[39]. People living in poor socioeconomic areas in Australia were also reported to disproportionately experience: employment restriction due to a disability [34]; unemployment and drug abuse [40]; childhood exposure to tobacco smoke [34]; proportionally low spending on medical and health care [41]; and delay in consultations with dental professionals due to cost [42].

These previous reports had clearly suggested that a ‘social gradient of health’ exists in Australia.[34] The current study contributes to an understanding as to how these nationally recognized social gradients can be addressed by health services locally, i.e. the geographic distributions of health risks and the contribution of

disadvantage can be objectively measured and resource allocation decisions made accordingly. In addition, at least as far as primary care access is concerned, the data suggest new thinking needs to be applied to addressing the inequalities at a health service level. Thus, the study provides indication of pathways from national statistics to evidence based-pathways for action on the ground, especially at regional level.

In addition, Australia's National Health Performance Framework recognises the importance of socioeconomic contexts on health determinants.[34] The framework includes socioeconomic circumstances in the determinants of health and access in the health system performance.[34] Determinants of health are the factors that influence health and illness. Many of the key determinants of health arise from the day to day life of an individual, mainly from the circumstances in which individuals live and work. The biomedical determinants of individual health are intertwined with the behavioural, environmental and social determinants of health.[34]

c) Global significance

Increasing evidence of the relationship between area-level factors and health outcomes leads us to a appreciation of human sensitivity to social environments. Social environments may either strengthen or undermine the health of individuals and communities because of their pervasive effects, which are known as the 'social determinants of health'.[43]

The World Health Organization (WHO) has described social determinants as:

"...the circumstances in which people grow, live, work and age and the systems put in place to deal with illness. The conditions in which people live and die are, in turn, shaped by political, social and economic forces…".[44]

These determinants are often underplayed by economic policies and political systems leading to inequalities in health outcomes.[45] When these determinants are the result of simply unavoidable differences, they are considered as inequalities. However avoidable inequalities are considered as inequities and unjust; and require appropriative policy initiatives to increase equity and social justice. According to WHO, the social determinants can be seen as the 'cause of the causes', thus the fundamental determinant that influences health. The figure illustrates on the inward influence of the general socioeconomic environmental conditions through the social and community networks to the individual health.[34] The thesis provides consistent evidence on the contributing role of socioeconomic contexts on the geographic variations of risk factors and thus future disease.

Implications for health care service planning

It is clear from the study that there exists geographical variation in the distribution of multiple CMRFs in the Illawarra-Shoalhaven region of NSW, Australia.[7] Area-level disadvantage was consistently associated with a range of CMRFs analysed in this study and explained a major proportion of the geographic variation in cardiometabolic risk distribution in the study region.[19] Higher access to primary care was associated with lower risk values of HDL and obesity after adjusting for age, sex and area-level disadvantage but was not associated with the remaining CMRFs.[46] Geographic access to primary care explained only a minimal proportion of the geographic variation of the higher risk CMRFs in the study region.[46] Thus, the study underpins the importance of the overall socioeconomic development of the areas for the area-level prevention and management of CMRFs and desirability of future research concerning the performance measures of primary care in this study region.

With regard to the area-level health care service commissioning, the thesis findings support universal interventions for the prevention and control of CMRFS with a proportional priority to the need and disadvantage level of the populations. As per this, the most disadvantaged areas in the study region could gain the highest proportion of allocation of services and resources based on the identified need. A prioritised service and resource allocations proportional to the disadvantage level of the areas might bring down the rates of higher risk CMRFs, and eventually the rates of future CVD related hospitalisations and health care expenditure in the study region.

Also, it should be noted that the demonstrated non associations and the minimal contributions of the geographic access to primary care on the geographic variation of higher risk CMRFs should not undermine the contributions and existing the burden on primary care to routinely monitor and control the distributed CMRFs in the study region. The findings only indicate the need for future studies on other aspects of primary care access and their performance measures.

7.6 Future directions and recommendations

The section provides recommendations for policy, practice and research, including technical and content area considerations.

7.6.1 Future policy directives: Better health through improving social determinants

The WHO Commission on Social Determinants of Health (CSDH) has made global recommendations to reduce the health inequity between areas through acting on area-level disadvantage.[47] Improving the socioeconomic contexts/determinants of health is identified as the most suitable way to reduce health inequalities, targeting a better health for all across areas.[47]

Evidence from the current study suggests that improvements in the overall socio-economic context of disadvantaged areas may reduce the inequity and thus social injustice, of the observed variation in the distribution of CMRFs in the study region. It is estimated that half a million Australians could avoid chronic illness, $2.3 billion in yearly hospital costs saved and the Pharmaceutical Benefits Scheme prescriptions could be reduced by 5.3 million, if the health inequality between the most and least disadvantaged areas were closed.[48]

The current study provides consistent evidence over multiple parameters for the contributing role of area-level socioeconomic disadvantage in the geographic inequalities in chronic disease parameters. The study also provides some evidence for lower risk factors in areas with better access to primary care. Over and above these findings, the study also demonstrates data and methodological approaches to address the health inequalities by health services and regional area-level planning authorities.

The WHO recommends adopting a ‘health in all policy’ approach to address the socioeconomic determinants of health. The approach suggests policies and interventions from all sectors and levels – to focus and be oriented with the health of individuals and thus of populations and nations.[49] However, there are still many barriers in adopting a socioeconomic determinant approach.

7.6.2 Future Research directions for the extension of current findings

The area-level associations of high TC were often the reverse to other CMRFs analysed in this study. This indicates a need for further studies to explore individual and area-level factors contributing to the geographic distribution of high TC findings in the study region.

Extension of the findings of this study requires inclusion of additional individual and area-level factors to help explain the geographic variation of CMRFs. The current study explains a maximum of 54.7% (on low HDL) of area-level variation. Inclusion of more individual and area-level factors may help to derive the contribution of the remaining factors not included in the current study. Other individual and area-level factors not considered in this study but could be considered in future, include: individual-level SES[50]; type of neighbourhood food outlets[51-54]; physical activity resources [55-57]; residential density and service availability[58]; social cohesion or social capital [59] of the SA1s; environmental pollution[60]; effect of diurnal cycles of light and day, sunlight exposure, seasons, altitude, latitude and greenspaces [61]; other environmental risk factors[62]; and effect of existing area-level policy interventions to reduce CVD risk [63].

7.7 Thesis strength and limitations

A key strength of this study is that it simultaneously analysed a range of eight individual CMRFs identified from the same study region. Consistent evidence on multiple risk factors adds to the diverse nature of the distribution and associations of the risk factors analysed.

The multilevel analytical approaches used in the study accounts for the nesting of individual-level data within different geographic areas. The principal task of multilevel modelling was to decompose the total individual level variance in its components.[21] Area-level generalisations of the findings were possible with the use of random intercepts in a multilevel model. Multilevel models allowed for the estimation of associations between specific area-level characteristics and binary measures of individual-level CMRF outcomes. It also allowed for the analysis of small-area variation and their indices, without disregarding the within-group individual-level variations.[21]

Also, the study used the smallest available geographic units for the area-level analyses. Statistical Area level 1 (SA1) is the smallest geographic unit for the release of census data in 2011 ABS census.[57] Level 1 Statistical Areas generally have a population of 200 to 800 persons (400 average) and the ISR covers a total of 980 conterminous SA1s. The use of small-area in the models has the potential to improve the estimate quality and enhance the area-level precision of understanding the health inequalities.[64]

The study region covers both urban and rural areas and the CMRF data used in the study have a near census coverage of the residential population.[7, 65] However, area-level impacts of rural and urban areas were not analysed in this study as the study region in this thesis is dichotomously divided into rural areas in the south and urban areas in the north. Hence, including a rural-urban analysis would likewise be not informative. While acknowledging possible impacts of rural-urban status on the global distribution of CMRFs, the geographic nature of the study region in this thesis is considered not suitable for such analyses.

The CMRFs data used in the study were population derived and consisted of data extracted from 256, 525 adult residents of ISR, including 144, 418 (56.3%) women. Overall, the Illawarra-Shoalhaven region in Australia had an estimated residential population of 369, 469 in Australian Bureau of Statistic (ABS) Census of Population and Housing in year 2011, of which 285, 385 (77.24%) were adults (>=18 years). This indicates the population coverage of the study data, even though obtained from a privately functioning network of service providers.

The study has to be considered with its limitations: Firstly, the cross-sectional design of the study precludes any causal inferences. In addition, the non-linear and time varying effects of covariates analysed in this study restrict generalisability of their findings.[66] Secondly, the study data were obtained from people already using

health care service facilities from the region. This limits representation of point-estimates to the general population.[66] However, it should also be noted that the study sample had a near census coverage of the adult population residing in this region during the study period. Also it should be noted that further individual-level data additions were not possible with this dataset as the de-identification process precluded the inclusion of any further individual level data. Thus the de-identified dataset does not support any further individual-level explorations into lifestyle related risk factors such as high blood pressure, poor dietary habits, inadequate exercise, sedentary life style patterns and smoking.

Thirdly, the individual and area-level explanatory variables in this thesis are limited. The de-identified dataset used in this thesis does not support any further explorations into individual-level attributes such as individual-level SES, hypertension, poor dietary habits, inadequate exercise, sedentary lifestyle patterns, and smoking. Area-level explanatory variables were limited to area-level disadvantage and primary care access of the study region. Other area-level factors not analysed in this study, but with potential to be associated with the outcome variables, could include the type of neighbourhood food outlets [51-54]; physical activity resources [55-57]; residential density and service availability[58]; social cohesion or social capital [59] of the SA1s; environmental pollution [60]; effect of diurnal cycles of light and day, sunlight exposure, seasons, altitude, latitude and greenspaces [61]; other environmental risk factors[62]; and effect of existing area-level policy interventions to reduce CVD risk [63]. I acknowledge that adjusting for a spectrum of individual and area-level covariates could potentially derive a more precise estimate of the contribution of area-level disadvantage in the study region. However, it should also be noted that unravelling the individual and area-level attributions of the observed geographic variation of CMRFs were not the primary intention in this thesis, but informing regional health care service commissioning on the evidence-based need for targeted area-level approaches for the prevention and control of the observed CMRFs distribution in the study region which is achieved in this thesis.

The standard practice in Australia and globally is to use the raw and/or standardised rates to inform health care service commissioning as they are a more realistic representation of the volume of morbidity distributed in a population.[34] Previous research indicates that unpacking the effects of covariates should be both attributable and responsive to policy changes[67, 68], and research reports from Australia indicate that adjusted rates may mask the actual volume of service requirments in an area[69]. Also it is worth noting that albeit the health service planning principles in Australia is person focused, the actual design/types of planning are done is for:1) geographical catchments; 2) population groups; 3) clinical service streams/areas such as, prevention, primary care, ambulatory care, acute care, sub-acute care and mental health.[70] The study 3 reports age and sex standardised effects of area-level disadvantage on CMRFs, and these estimates are more likely to be informative for area-level health care service commissioning and related health policy developments rather than a spectrum of covariate adjusted estimates. A classic complex systems model of public health conceptualises health inequalities as outcomes of a multitude of interdependent elements within a connected whole.[71] However, it is worth noting that rhetoric complex systems approaches to public health in academic research is only rarely operationalized to generate relevant policies.[71, 72] Complex attribution structuring should definitely be of academic research interest but may actually complicate and possibly nullify the information required for the health care service commissioning. Therefore, while acknowledging the possible direct and subtle attributions of a range of individual and area-level covariates on the estimated effects

of area-level disadvantage on CMRFs, their unravelling is not intended in this thesis given them main purpose of this study that is to inform area-level health care service commissioning.

Fourthly, the IRSD measure used in the study has limitations intrinsic to aggregate measures. The IRSD is one of four Socio-Economic Indexes for Areas (SEIFA) based on socioeconomic status. The IRSD summarises a range of measures of relative socioeconomic disadvantage of people and households within SA1s and includes: level of income; education; employment; family structure; disability; housing; transportation; and internet connection.[65] The IRSD scores of the SA1s were accessed from the published data of 2011 Australian census and a higher score indicated lower levels of disadvantage.[65] Despite the widespread acceptance of these measures in the Australian literature, the aggregate nature of this measure does not account for the heterogeneity of the people residing within a given area.[73] Fifthly, the primary care access measured in this study is limited to a radial buffer distance of 30 km. The 2SFCA method used in the study to calculate access uses this radial buffer distance to define the catchment areas both from the service provider locations and from the geographic centroids of the analysis units. In the preliminary stage of this study, sensitivity analyses were performed using 1 km, 16 km and 30 km spatial buffer distances. The 1 km radial buffer distance from the primary care provider locations covered only 545 (56%) SA1 centroids in the study region, whereas as a 16 km radial buffer distances covered 973 (~99%) and 30 km radial buffer distance covered 978 (~100%) SA1 centroids. Therefore, a radial buffer distance of 30 km was chosen to determine the primary care access of the SA1s. This radial distance was observed to cover the mixed rural, semi-rural and urban distribution of the population in the study region well, more adequately than the distances used in the sensitivity analyses.

Finally, the multilevel logistic regression models and the variance partitioning approach adopted in this thesis has limitations intrinsic to their methodology. The assumptions of the standard multilevel logistic regression modelling methods adopted in this thesis would not be able to account for the autocorrelation of the area-level residuals (if any) of the models. Simulated results comparing the outcomes of standard multilevel models and spatial models indicate that both of these methods produce similar fixed effect estimates, but can vary in their random effect estimates.[74] Previous work demonstrates that both the multilevel and spatial models tend to overestimate the corresponding random effects variances compared with hybrid models.[74, 75] However, we observe that this potential overestimation is not likely to be substantial in the thesis results as the area-level variance shown in all the models was quite small. This is demonstrated by the very low ICC and MOR values of our models, which can be interpreted as demonstrating very small or no area-level difference.[21] In addition, hybrid models are harder to adopt for use with individual level de-identified data such as the data used in this thesis.[74, 75] More accurate geo-referencing of individuals would be necessary to implement such models, especially to account for the correlation of residuals of neighbouring units compared with those far apart.[75] We observe that is beyond the scope of our data type and location specifications. We also observed that artificial/simulated attempts to overcome this data specification requirement through random distribution of samples at constant distances within areas,[75] would not sufficiently reflect the actual nature and distribution of populations in our study areas especially as it consists of a range of highly dense and very sparsely populated areas within the study region. While acknowledging this limitation, I believe the effects of this are not critical in the results. Also, the variance partitioning approach adopted in this thesis may not fully reflect the impact of contextual variables, especially the interactions among area-level and individual-level factors. It is possible that there could be existing relations between the lower-level variable (e.g. higher

age group of the sample), which changes as a function of the upper-level variable (e.g. higher area-level disadvantage) but not captured within the variance partitioning approach of the binary outcome CMRF variables of the study.[76] Analysis of such cross-level interactions and between variable interactions were not a focus in this thesis though they might be of potential epidemiological utility [59], as their outcomes were not believed to add information for health care service planning which is the primary intention of this research.

Notwithstanding these limitations, the thesis is novel in that it analysed a range of CMRFs across a widely dispersed population and included both rural and urban residents. In addition, the thesis used six years (year 2012 – 2017) of CMRF tests data from the region in the hierarchical multilevel analyses. The findings of the thesis indicate that there is a significant geographic variation in the distribution of CMRFs in the study region. Those residing in the most disadvantaged areas are more likely to be identified with higher risk CMRFs than those in lower disadvantage areas. Also, the thesis suggests higher odds of being identified with low HDL and obesity with reduced access to primary care. However, the low ICC and MOR values of the area-level models in the thesis do not support for contextual approaches. Rather, the findings of the study support a proportionate universalism approach in which health resources are made universally available but proportional to the need and disadvantage level of the affected population.[22, 23]

7.8 Conclusion

The thesis has made a significant contribution to understanding the geographic variation of cardiometabolic risk factors in the Illawarra-Shoalhaven region of NSW, Australia. The study reports associations of area-level disadvantage and area-level access to primary care with CMRF distribution. The findings arising from this thesis have demonstrated that area-level disadvantage explains a large proportion of the geographic variation of CMRFs in the study region. The results support future investigations into whether public health activities or interventions in primary care, targeted to localities with greater area-level disadvantage, can ameliorate CMRFs. Based on the findings, the thesis recommends evidence based universal and proportionate interventions to priority populations for the prevention and control of CMRFs. These findings can be used to inform regional health care service commissioning and related policy developments; and are highly relevant in the context of the global paradigm shift from communicable diseases to cardiovascular diseases (CVD) as the leading cause of human death and health care expenditure.

References

1. Toms, R., et al., *Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature.* International journal of health geographics, 2019. **18**(1): p. 1.
2. Chaix, B., *Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda.* Annual review of public health, 2009. **30**: p. 81-105.
3. Leal, C. and B. Chaix, *The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda.* Obesity reviews, 2011. **12**(3): p. 217-230.
4. El-Sayed, A.M., P. Scarborough, and S. Galea, *Socioeconomic inequalities in childhood obesity in the United Kingdom: a systematic review of the literature.* Obesity facts, 2012. **5**(5): p. 671-692.
5. Slopen, N., et al., *Socioeconomic and other social stressors and biomarkers of cardiometabolic risk in youth: a systematic review of less studied risk factors.* PLoS One, 2013. **8**(5): p. e64418.
6. Occelli, F., et al., *Mapping end-stage renal disease (ESRD): spatial variations on small area level in northern France, and association with deprivation.* PloS one, 2014. **9**(11): p. e110132.
7. Toms, R., et al., *Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia.* PloS one, 2019. **14**(10): p. e0223179.
8. Stocks, N., et al., *Gender, socioeconomic status, need or access? Differences in statin prescribing across urban, rural and remote Australia.* Australian Journal of Rural Health, 2009. **17**(2): p. 92-96.
9. Stocks, N.P., et al., *Statin prescribing in Australia: socioeconomic and sex differences.* Medical journal of Australia, 2004. **180**(5): p. 229-231.
10. Alkerwi, A., et al., *Geographic variations in cardiometabolic risk factors in Luxembourg.* International journal of environmental research and public health, 2017. **14**(6): p. 648.
11. Astell-Burt, T., et al., *Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia.* Diabetes research and clinical practice, 2014. **106**(3): p. e68-e73.
12. Oh, W.S., et al., *Geographical variations and influential factors in prevalence of cardiometabolic diseases in South Korea.* PloS one, 2018. **13**(10): p. e0205005.
13. Paquet, C., et al., *Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia.* International journal of environmental research and public health, 2016. **13**(5): p. 519.
14. Valdés, S., et al., *Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@ bet. es study.* Revista Española de Cardiología (English Edition), 2014. **67**(6): p. 442-448.
15. Zhou, M., et al., *Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults.* Diabetes care, 2015. **38**(1): p. 72-81.
16. Barker, L.E., et al., *Geographic distribution of diagnosed diabetes in the US: a diabetes belt.* American journal of preventive medicine, 2011. **40**(4): p. 434-439.
17. Lawlor, D.A., et al., *Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study.* American journal of public health, 2005. **95**(1): p. 91-97.

18. *National Vascular Disease Prevention Alliance. Guidelines for the management of absolute cardiovascular disease risk. 2012. Available at: https://www.heartfoundation.org.au/getmedia/4342a70f-4487-496e-bbb0-dae33a47fcb2/Absolute-CVD-Risk-Full-Guidelines_2.pdf*
19. Toms, R., et al., *Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia.* Scientific Reports, 2020. **10**(1): p. 1-18.
20. Barter, P.J., et al., *Effect of statins on HDL-C: a complex process unrelated to changes in LDL-C: analysis of the VOYAGER Database.* Journal of lipid research, 2010. **51**(6): p. 1546-1553.
21. Merlo, J., P. Wagner, and G. Leckie, *A simple multilevel approach for analysing geographical inequalities in public health reports: The case of municipality differences in obesity.* Health & place, 2019. **58**: p. 102145.
22. Lu, D. and I. Tyler, *Focus on: a proportionate approach to priority populations.* Public Health Ontario. https://www. publichealthontario. ca/en/eRepository/Focus_On_Priority_Populations. pdf. Accessed, 2016. **29**.
23. Marmot, M. and R. Bell, *Fair society, healthy lives.* Public health, 2012. **126**: p. S4-S10.
24. Rose, G., *Sick individuals and sick populations.* International journal of epidemiology, 2001. **30**(3): p. 427-432.
25. Ngo, A.D., et al., *Area-level socioeconomic characteristics, prevalence and trajectories of cardiometabolic risk.* International journal of environmental research and public health, 2014. **11**(1): p. 830-848.
26. Ngo, A.D., et al., *Area-level socioeconomic characteristics and incidence of metabolic syndrome: a prospective cohort study.* BMC public health, 2013. **13**(1): p. 1-11.
27. McGrail, M.R., *Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements.* International journal of health geographics, 2012. **11**(1): p. 50.
28. Frohlich, K.L., et al., *Advancing the population health agenda.* Canadian journal of public health, 2004. **95**(5): p. 392-395.
29. Bronfenbrenner, U., *Ecological systems theory (Vol. 6).* Greenwich, CT: JAI, 1989.
30. Stokols, D., *Translating social ecological theory into guidelines for community health promotion.* American journal of health promotion, 1996. **10**(4): p. 282-298.
31. Stokols, D., *Social ecology and behavioral medicine: implications for training, practice, and policy.* Behavioral medicine, 2000. **26**(3): p. 129-138.
32. Krieger, N., *Epidemiology and the web of causation: has anyone seen the spider?* Social science & medicine, 1994. **39**(7): p. 887-903.
33. Krieger, N., *Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination.* International journal of health services, 1999. **29**(2): p. 295-352.
34. Australian Institute of Health and Welfare. *Australia's health 2016.* Australia's health series no. 15. Cat. no. AUS 199. Canberra: AIHW. 2016 Available from: https://www.aihw.gov.au/getmedia/68294863-481a-4d5e-b9fe-2d94e4bb7aee/ah16-7-1-indicators-australias-health.pdf.aspx.
35. NHPA (National Health Performance Authority) *Healthy communities: avoidable deaths and life expectancies in 2009–2011*. Sydney: NHPA. 2013.
36. Australian Bureau of Statistics. *National Health Survey, first results 2014–15.* . ABS cat. no. 4364.0.55.001. Canberra: ABS. 2015.
37. Australian Institute of Health and Welfare. *Mortality inequalities in Australia 2009–2011.* AIHW bulletin no. 124. Cat. no.AUS 184. 2014 Canberra: AIHW.

38. Mallett, S., *Precarious housing and health inequalities: what are the links?* 2011.
39. Australian Institute of Health and Welfare. *Australia's mothers and babies 2013—in brief.* Perinatal statistics series no. 31.Cat no. PER 72. Canberra: AIHW. 2015.
40. Australian Institute of Health and Welfare. *National Drug Strategy Household Survey detailed report: 2013. Drug statisticsseries no. 28. Cat. no. PHE 183*. Canberra: AIHW; 2014.
41. Australian Bureau of Statistics. *Australian social trends, March quarter 2012.* ABS cat. no. 4102.0. Canberra: ABS. 2012.
42. Australian Bureau of Statistics. *Patient experiences in Australia: summary of findings, 2014–15. .* ABS cat. no. 4839.0.Canberra: ABS. 2015.
43. Wilkinson, R.G. and M. Marmot, *Social determinants of health: the solid facts*. 2003: World Health Organization.
44. World Health Organization. *Social determinants of health: key concepts*. Geneva: World Health Organization 2008; Available from: https://www.who.int/social_determinants/thecommission/finalreport/key_concepts/en/
45. Pan American Health Organisations (PAHO). *SOCIAL DETERMINANTS OF HEALTH IN THE AMERICAS.* Regional Office for the Americas of the World Health Organization; USA. 2013; Available from: https://www.paho.org/salud-en-las-americas-2017/?p=45
46. Toms, R., et al., *Role of area-level access to primary care on the geographic variation of cardiometabolic risk factor distribution: a multilevel analysis of the adult residents in the Illawarra—Shoalhaven Region of NSW, Australia.* International journal of environmental research and public health, 2020. **17**(12): p. 4297.
47. Health, C.o.S.D.o., *Closing the gap in a generation: health equity through action on the social determinants of health: final report of the commission on social determinants of health*. 2008: World Health Organization.
48. Brown, L., L. Thurecht, and B. Nepal, *The cost of inaction on the social determinants of health.* University of Canberra: National Centre for Social and Economic Modelling (NATSEM), 2012.
49. World Health Organization, *Closing the gap: policy into practice on social determinants of health: discussion paper*. Geneva:WHO., 2011.
50. Sodjinou, R., et al., *Obesity and cardio-metabolic risk factors in urban adults of Benin: relationship with socio-economic status, urbanisation, and lifestyle patterns.* BMC public health, 2008. **8**(1): p. 84.
51. Cummins, S. and S. Macintyre, *Food environments and obesity—neighbourhood or nation?* International journal of epidemiology, 2006. **35**(1): p. 100-104.
52. Fraser, L.K., et al., *The geography of fast food outlets: a review.* International journal of environmental research and public health, 2010. **7**(5): p. 2290-2308.
53. Macdonald, L., S. Cummins, and S. Macintyre, *Neighbourhood fast food environment and area deprivation—substitution or concentration?* Appetite, 2007. **49**(1): p. 251-254.
54. Pearce, J., et al., *Neighborhood deprivation and access to fast-food retailing: a national study.* American journal of preventive medicine, 2007. **32**(5): p. 375-382.
55. Buttar, H.S., T. Li, and N. Ravi, *Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation.* Experimental & clinical cardiology, 2005. **10**(4): p. 229.
56. Chomistek, A.K., et al., *Healthy lifestyle in the primordial prevention of cardiovascular disease among young women.* Journal of the American College of Cardiology, 2015. **65**(1): p. 43-51.

57. Fiuza-Luces, C., et al., *Exercise benefits in cardiovascular disease: beyond attenuation of traditional risk factors.* Nature Reviews Cardiology, 2018. **15**(12): p. 731-743.
58. Malambo, P., et al., *Built environment, selected risk factors and major cardiovascular disease outcomes: a systematic review.* PloS one, 2016. **11**(11): p. e0166846.
59. Rhodes, T., *Risk environments and drug harms: a social science for harm reduction approach*. 2009, Elsevier.
60. O'Toole, T.E., D.J. Conklin, and A. Bhatnagar, *Environmental risk factors for heart disease.* Reviews on environmental health, 2008. **23**(3): p. 167-202.
61. Bhatnagar, A., *Environmental determinants of cardiovascular disease.* Circulation research, 2017. **121**(2): p. 162-180.
62. Malambo, P., et al., *The relationship between objectively-measured attributes of the built environment and selected cardiovascular risk factors in a South African urban setting.* BMC public health, 2018. **18**(1): p. 847.
63. Cosselman, K.E., A. Navas-Acien, and J.D. Kaufman, *Environmental factors in cardiovascular disease.* Nature Reviews Cardiology, 2015. **12**(11): p. 627.
64. Moon, G., et al., *The utility of geodemographic indicators in small area estimates of limiting long-term illness.* Social Science & Medicine, 2019. **227**: p. 47-55.
65. Australian Bureau of Statistics. *Main Features - IRSD.*; Available from: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011.
66. Carlson, M.D. and R.S. Morrison, *Study design, precision, and validity in observational studies.* Journal of palliative medicine, 2009. **12**(1): p. 77-82.
67. Adair, C.E., et al., *Performance measurement in healthcare: part II–state of the science findings by stage of the performance measurement process.* Healthcare Policy, 2006. **2**(1): p. 56.
68. Giuffrida, A., H. Gravelle, and M. Roland, *Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes.* Bmj, 1999. **319**(7202): p. 94-98.
69. Falster, M.O., et al., *Sociodemographic and health characteristics, rather than primary care supply, are major drivers of geographic variation in preventable hospitalizations in Australia.* Medical care, 2015. **53**(5): p. 436.
70. Department of Health. *Guide to health service planning.* 2015 [cited 2020; Available from: https://www.health.qld.gov.au/__data/assets/pdf_file/0025/443572/guideline-health-service-planning.pdf.
71. Rutter, H., et al., *The need for a complex systems model of evidence for public health.* The lancet, 2017. **390**(10112): p. 2602-2604.
72. Wutzke, S., et al., *Systems approaches for chronic disease prevention: sound logic and empirical evidence, but is this view shared outside of academia.* Public Health Res Pract, 2016. **26**(3): p. e2631632.
73. Walker, A. and N. Becker, *Health inequalities across socio-economic groups: comparing geographic-area-based and individual-based indicators.* Public health, 2005. **119**(12): p. 1097-1104.
74. Xu, H., *Comparing spatial and multilevel regression models for binary outcomes in neighborhood studies.* Sociological methodology, 2014. **44**(1): p. 229-272.
75. Chaix, B., J. Merlo, and P. Chauvin, *Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France.* Journal of Epidemiology & Community Health, 2005. **59**(6): p. 517-526.
76. Aguinis, H., R.K. Gottfredson, and S.A. Culpepper, *Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling.* Journal of Management, 2013. **39**(6): p. 1490-1528.

Appendix

Appendix I: Search strategy

Appendix I (a): Literature Search Strategy for Article 1)

Database: Ovid MEDLINE(R) <2001 to November 30, 2018>

Search Strategy:

---

1 cardiometabolic.ti, ab, mp. (6616)

2 cardio metabolic.ti, ab, mp. (1069)

3 metabolic syndrome.ti, ab, mp. (42423)

4 metabolic risk.ti, ab, mp. (3856)

5 Geographic.ti, ab, mp. (69264)

6 Geospatial.ti, ab, mp. (1148)

7 Spatial.ti, ab, mp. (211880)

8 regional variation.ti, ab, mp. (3211)

9 area socioeconomic.ti, ab, mp. (147)

10 neighbo?rhood socioeconomic.ti, ab, mp. (658)

11 area poverty.ti, ab, mp. (38)

12 neighbo?rhood deprivation.ti, ab, mp. (346)

13 1 or 2 or 3 or 4 (49526)

14 5 or 6 or 7 or 8 9 or 10 or 11 or 12 (274223)

15 13 and 14 (166)

16 limit 15 to (english language and humans and yr="2001 -Current" and "all adult (19 plus years)") (91)

Appendix I (b): Search URLs for Article 1

I. Database: Scopus <2001 to November 30, 2018>

https://www-scopus-com.ezproxy.uow.edu.au/results/results.uri?sort=plf-f&src=s&sid=ea1c0f3d085ba8ac5196ec70f956b0b0&sot=a&sdt=a&cluster=scolang%2c%22English%22%2ct%2bscoexactkeywords%2c%22Human%22%2ct%2c%22Humans%22%2ct%2c%22Adolescent%22%2cf%2c%22Nonhuman%22%2cf%2c%22Child%22%2cf&sl=364&s=%28+TITLE-ABS-KEY+%28+geographic+OR+geospatial+OR+spatial+OR+regional+AND+variation+%29+OR+TITLE-ABS-KEY+%28+area+AND+socioeconomic+OR+area+AND+socioeconomic+OR+neighbo%3frhood+AND+socioeconomic+OR+area+AND+poverty+OR+neighbo%3frhood+AND+deprivation+%29+AND+TITLE-ABS-KEY+%28+cardiometabolic+OR+%22cardio+metabolic%22+OR+%22metabolic+syndrome%22+OR+%22metabolic+risk%22+%29+%29+AND+PUBYEAR+%3E+2000&origin=searchhistory&txGid=ecc5b4b292f0fe563250cb74eb62adc7

Search history

(TITLE-ABS-KEY (geographic OR geospatial OR spatial OR regional AND variation) OR TITLE-ABS-KEY (area AND socioeconomic OR area AND socioeconomic OR neighbo?rhood AND socioeconomic OR area AND poverty OR neighbo?rhood AND deprivation) AND TITLE-ABS-KEY (cardiometabolic OR "cardio metabolic" OR "metabolic syndrome" OR "metabolic risk")) AND PUBYEAR > 2000 AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (EXACTKEYWORD , "Human") OR LIMIT-TO (EXACTKEYWORD , "Humans") OR EXCLUDE (EXACTKEYWORD , "Adolescent") OR EXCLUDE (EXACTKEYWORD , "Nonhuman") OR EXCLUDE (EXACTKEYWORD , "Child"))

II. Database: PubMed <2001 to November 30, 2018>

https://www-ncbi-nlm-nih-gov.ezproxy.uow.edu.au/pubmed/?term=(((Geographic%5BTitle%2FAbstract%5D+OR+Geospatial%5BTitle%2FAbstract%5D+OR+Spatial%5BTitle%2FAbstract%5D+OR+%E2%80%9Cregional+variation%E2%80%9D%5BTitle%2FAbstract%5D))+OR+(%E2%80%9Carea+socioeconomic%22%5BTitle%2FAbstract%5D+OR+%E2%80%9Cneighbourhood+socioeconomic%E2%80%9D%5BTitle%2FAbstract%5D+OR+%E2%80%9Carea+poverty%E2%80%9D%5BTitle%2FAbstract%5D+OR+%E2%80%9Cneighbourhood+deprivation%E2%80%9D%5BTitle%2FAbstract%5D))+AND+(cardiometabolic%5BTitle%2FAbstract%5D+OR+%E2%80%9Ccardio+metabolic%E2%80%9D%5BTitle%2FAbstract%5D+OR+%E2%80%9Cmetabolic+syndrome%E2%80%9D%5BTitle%2FAbstract%5D+OR+%22metabolic+risk%22%5BTitle%2FAbstract%5D)

Filters activated: Publication date from 2001/01/01 to 2018/11/30, Humans, English, Adult: 19+ years.

III. Database: Web of science <2001 to November 30, 2018>

http://apps.webofknowledge.com.ezproxy.uow.edu.au/summary.do?product=UA&doc=1&qid=31&SID=C3eZJBIA1SNizT9v93Z&search_mode=CombineSearches&update_back2search_link_param=yes

Appendix II: List of Excluded Full Text Studies With Reason in Article 1

	Excluded studies	Reason
1	Inoue, Y., et al. Neighborhood Characteristics and Cardiovascular Risk among Older People in Japan: Findings from the JAGES Project. PLoS ONE .[Electronic Resource] 11, e0164525 (2016).	'Accident prone perception 'is the ASED proxy measurement.
2	Sundquist, K., Eriksson, U., Mezuk, B. & Ohlsson, H. Neighborhood walkability, deprivation and incidence of type 2 diabetes: a population-based study on 512, 061 Swedish adults. Health Place 31, 24-30 (2015).	'Neighbourhood deprivation' is used as a control in analyses.
3	Congdon, P. Estimating diabetes prevalence by small area in England. J Public Health (Oxf) 28, 71-81 (2006).	Methodology oriented paper.
4	Mezuk, B. *et al.* Depression, neighborhood deprivation and risk of type 2 diabetes. *Health Place* **23,** 63–69 (2013).	Depression patients are the study population.
5	Stoddard, P. J. *et al.* Neighborhood deprivation and change in BMI among adults with type 2 diabetes: the Diabetes Study of Northern California (DISTANCE). *Diabetes Care* **36,** 1200–1208 (2013).	Diabetic patients are the study population.
6	Chaikiat, A., Li, X., Bennet, L. & Sundquist, K. Neighborhood deprivation and inequities in coronary heart disease among patients with diabetes mellitus: a multilevel study of 334, 000 patients. Health Place 18, 877–882 (2012).	Diabetic patients are the study population.
7	Yu, Z., et al. Obesity related metabolic abnormalities: distribution and geographic differences among middle-aged and older Chinese populations. Prev Med 48, 272-278 (2009)	Non-continuous geographic units (i.e prevalence in two cities in the north and south of country are compared, and its urban/ rural cross differences were focussed)
8	Chichlowska, K.L., et al. Individual and neighborhood socioeconomic status characteristics and prevalence of metabolic syndrome: the Atherosclerosis Risk in Communities (ARIC) Study. Psychosom Med 70, 986-992 (2008).	No results on discrete CMRFs or its association with ASED.
9	Ardern, C.I. & Katzmarzyk, P.T. Geographic and demographic variation in the prevalence of the metabolic syndrome in Canada. Can 31, 34-46 (2007).	No results on discrete CMRFs or its association with ASED.

10	Traissac, P., et al. Abdominal vs. overall obesity among women in a nutrition transition context: geographic and socio-economic patterns of abdominal-only obesity in Tunisia. Population health metrics 13, 1-1 (2015).	Obesity results are not presented due to small (1.4%) overall prevalence (waist circumference defined as abdominal adiposity is focussed).
11	Jones, M. & Huh, J. Toward a multidimensional understanding of residential eighbourhood: a latent profile analysis of Los Angeles neighborhoods and longitudinal adult excess weight. Health Place 27, 134-141 (2014)	Geographic area based results were not available. Also minimal data on ASED – but three types of neighbourhoods based on 'social context variables' were identified in the study.
13	Kandala N-B, Manda SOM, Tigbe W, Mwambi H, Stranges S. Geographic distribution of cardiovascular comorbidities in South Africa: a national cross-sectional analysis. Journal of Applied Statistics 2014;41(6):1203-1216	Age group of the study sample (aged 15 and over) are under the review defined adult age group (18 years and above).

Appendix III: PRISMA 2009 Checklist for Article 1

Section/topic	#	Checklist item	Reported on page # (in published version of article 1)
Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: A systematic review of the literature			1
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			2
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			3-4
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			4-6
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4-5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	4-5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5

Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	6

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	6
RESULTS			6-11
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	6-7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-10 Table 1&2
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	13
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression .[see Item 16]).	11 Qualitative synthesis
DISCUSSION			11-13
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	13

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12
FUNDING			15
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	15

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097. For more information, visit: **www.prisma-statement**

Appendix IV: SIMLR Data Extraction Document

Appendix IV (a): SIMLR Study Data Extraction and Supply Documentation

This document describes the preparation and extraction of data from the Southern.IML Research (SIMLR) Study database for the project listed below. Access to, and use of, these data are subject to all conditions imposed by a responsible Human Research Ethics Committee, the SIMLR Data Access Agreement and the SIMLR Study Management Committee. You should contact the SIMLR Management Committee and your responsible Human Research Ethics committee immediately if you are unable to comply with any conditions of access, or use or if you become aware that any of these conditions have been breached.

Project:	An epidemiological study of chronic disease parameters at regional level, in view of developing a cardiometabolic risk map to facilitate regional planning activities.
Researcher(s):	Renin Toms, Andrew Bonney, Xiaoqi Feng, Darren Mayne
Ethics reference:	2017/124
Data extraction by:	Darren Mayne
Data extraction date:	Initial: 06 February 2018 Updated: 23 March 2018 Updated: 09 April 2018 (fixed incorrectly assigned SA1 location at time of testing)
Data extraction format:	Coma Separated Values (*.CSV)
File names:	RMBST288_TOMS_RENIN_PHD_DATA_EXTRACT_20180409.CSV
Number of records:	256, 526
Document version:	1.1
Document date:	09 April 2018

Study sample

Data were extracted from the SIMLR database for all non-pregnant persons aged ≥18 years with ≥1 cardiometabolic analyte result between 01 January 2012 and 31 December 2017. For each person, the most recent cardiometabolic test result was extracted to maximise its temporal alignment with residential location at time pathology testing.

Variables

Sex

The sex of individual patients is based on that recorded for their most recent episode of care (test result). We have received previous advice from Southern.IML Pathology Pty Ltd that this value is considered the most accurate record of gender status.

Diabetes status

Fasting blood sugar level (FBSL) tests within ± 24 months of glycated haemoglobin (HbA_{1c}) test dates were matched for each individual. Each testing record was then coded as algorithm-positive for diabetes if either the HbA_{1c} result was ≥ 6.5% or had an associated FBSL result ≥ 7.0 mmol/L within ±24 months. Diabetes case definition status (diab_status) was set to "Yes" for patients the first time the algorithm criterion was met, and then propagated throughout the data set for subsequent testing records. The diabetes diagnosis date (diab_date) corresponds to the date of the HbA_{1c} or FBSL used to assign algorithm-positive status, which may be used as a proxy for time with disease. We are currently preparing a paper on the performance of this algorithm for identifying patients with diabetes in administrative databases that you will be able to reference.

Cardiometabolic analytes

The extract includes the following cardiometabolic analyte variables: fasting blood sugar levels (FBSL), glycated haemoglobin (HbA_{1c}), high density lipoprotein (HDL) and total (TC) cholesterol, albumin creatine ratio (ACR), estimated Glomerular Filtration Rate (eGFR), and body mass index (BMI). For each of these analytes the following variables are provided: collection date (collection_date); age group at time of collection (collection_age05grp); test value in standard units (test_value_num); geocoding match status for residential address at time of collection (geo_match_status); geocoded Statistical Area 1 (SA1) of residential address at

time of collection (sa1_2011_code); 2011 Index of Relative Socioeconomic Advantage/Disadvantage for geocoded SA1 at time of collection (irsad_2011_quintile); 2011 Index of Relative Socioeconomic Disadvantage for geocoded SA1 at time of collection (irsd_2011_quintile); 2011 Index of Economic Resources for geocoded SA1 at time of collection (ier_2011_quintile); 2011 Index of Education and Occupation for geocoded SA1 at time of collection (ieo_2011_quintile). For most records, geocoding-based variables will be the same for all analytes; however, as the collection date between analytes increases for an individual, so does the probability that they will have moved to a residential location in a different quintile. We recommend using the quintile covariates that correspond to the outcome in your analysis.

Data completeness

Two data completeness variables are included in the extract to guide case selection for analysis. The number of missing test results (nmiss_test_results) indicates the number of analytes with missing values either because the test has never been requested or the most recent test predates 01 January 2012. The complete case status (complete_case) variable indicates if the current record is complete (complete_case = Y, i.e. no missing test results) or incomplete (complete_case = N, i.e. ≥ 1 missing analyte).

Table A.1 Variable names, types, values and descriptions for study data

Variable	Type	Values	Description
enc_project_id	Character		Encrypted patient identifier
epi_sex	Character	M = Male F = Female	Sex
diab_status	Character	No Yes (incident) Yes (prevalent)	Diabetes status from SIMLR classification algorithm
diab_date	Date	YYYY-MM-DD	Diabetes date from SIMLR classification algorithm
fbsl_collection_date	Date	YYYY-MM-DD	Fasting BSL collection date
fbsl_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	Fasting BSL age group
fbsl_test_value_num	Numeric	0.7–43.9	Fasting BSL test value (mmol/L)
fbsl_geo_match_status	Character	Exact address Exact street Exact suburb	Fasting BSL geocoding match status
fbsl_sa1_2011_code	Character	11-digit ABS code	Fasting BSL SA1 of residence
fbsl_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	Fasting BSL IRSAD quintile
fbsl_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	Fasting BSL IRSD quintile

Variable	Type	Values	Description
fbsl_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	Fasting BSL IER quintile
fbsl_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	Fasting BSL IEO quintile
hba1c_collection_date	Date	YYYY-MM-DD	HbA1c collection date
hba1c_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	HbA1c collection age group
hba1c_test_value_num	Numeric	2.6–17.8	HbA1c test value (%)
hba1c_geo_match_status	Character	Exact address Exact street Exact suburb	HbA1c geocoding match status
hba1c_sa1_2011_code	Character	11-digit ABS code	HbA1c SA1 of residence
hba1c_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	HbA1c IRSAD quintile
hba1c_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	HbA1c IRSD quintile
hba1c_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	HbA1c IER quintile
hba1c_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	HbA1c IEO quintile
hdl_collection_date	Date	YYYY-MM-DD	HDL collection date

Variable	Type	Values	Description
hdl_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	HDL collection age group
hdl_test_value_num	Numeric	0.06–5.76	HDL test value (mmol/L)
hdl_geo_match_status	Character	Exact address Exact street Exact suburb	HDL geocoding match status
hdl_sa1_2011_code	Character	11-digit ABS code	HDL SA1 of residence
hdl_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	HDL IRSAD quintile
hdl_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	HDL IRSD quintile
hdl_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	HDL IER quintile
hdl_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	HDL IEO quintile
tc_collection_date	Date	YYYY-MM-DD	TC collection date
tc_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	TC collection age group
tc_test_value_num	Numeric	1.08–39.42	TC test value (mmol/L)
tc_geo_match_status	Character	Exact address Exact street Exact suburb	TC geocoding match status

Variable	Type	Values	Description
tc_sa1_2011_code	Character	11-digit ABS code	TC SA1 of residence
tc_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	TC IRSAD quintile
tc_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	TC IRSD quintile
tc_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	TC IER quintile
tc_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	TC IEO quintile
acr_collection_date	Date	YYYY-MM-DD	ACR collection date
acr_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	ACR collection age group
acr_test_value_num	Numeric	0.1–1291.5	ACR test value (mcg/L)
acr_geo_match_status	Character	Exact address Exact street Exact suburb	ACR geocoding match status
acr_sa1_2011_code	Character	11-digit ABS code	ACR SA1 of residence
acr_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	ACR IRSAD quintile
acr_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	ACR IRSD quintile
acr_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	ACR IER quintile

Variable	Type	Values	Description
acr_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	ACR IEO quintile
egfr_collection_date	Date	YYYY-MM-DD	eGFR collection date
egfr_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	eGFR collection age group
egfr_test_value_num	Numeric	2–91 91 = >60 or >90	eGFR test value (mL/min/1.73m2)
egfr_geo_match_status	Character	Exact address Exact street Exact suburb	eGFR geocoding match status
egfr_sa1_2011_code	Character	11-digit ABS code	eGFR SA1 of residence
egfr_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	eGFR IRSAD quintile
egfr_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	eGFR IRSD quintile
egfr_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	eGFR IER quintile
egfr_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	eGFR IEO quintile
bmi_collection_date	Date	YYYY-MM-DD	BMI collection date

Variable	Type	Values	Description
bmi_collection_age05grp	Character	18–24 25–29 30–34 35–39 40–44 45–49 50–54 55-59 60–64 65–69 70–74 75–79 80–84 85+	BMI collection age group
bmi_test_value_num	Numeric	11.2–181.0	BMI test value (kg/m2)
bmi_geo_match_status	Character	Exact address Exact street Exact suburb	BMI geocoding match status
bmi_sa1_2011_code	Character	11-digit ABS code	BMI SA1 of residence
bmi_irsad_2011_quintile	Character	Q1 = Quartile 1 (Disadv) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Adv)	BMI IRSAD quintile
bmi_irsd_2011_quintile	Character	Q1 = Quartile 1 (Most) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (Least)	BMI IRSD quintile
bmi_ier_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	BMI IER quintile
bmi_ieo_2011_quintile	Character	Q1 = Quartile 1 (Low) Q1 = Quartile 2 Q1 = Quartile 3 Q1 = Quartile 4 Q1 = Quartile 5 (High)	BMI IEO quintile
nmiss_test_results	Numeric	0–6	Number of missing test results
complete_case	Character	Y = Yes N = No	Complete case (no missing test results)

Appendix IV (b): Data Extraction Program

```
*************************************************************************;
* Program:    SIMLR_PREVALENT_DIABETES_2010_TO_2014
* Version:    01
* Author:     Darren Mayne
* Date:       28 November 2016
* Contact:    darren.mayne@health.nsw.gov.au
* Purpose:    Extracts data from SIMLR Study database for diabetes
incidence
* Sub macros: METADATA
* Notes:      Requires access to the SIMLR_GEO_MERGED.accdb database on the
*             \\shares3.its.uow.edu.au\1000206_simlr_raw$ share
*
*             Share location
*
\\shares3.its.uow.edu.au\1000206_SIMLR_Diabetes$\STUDIES\Prevalent Diabetes
Extract
*
*             hba1c_mean_updated_year = financial year NOT calendar year
*=======================================================================
* PARAMETERS:
* -Name---------------- -Description--------------------------------------
*
*=======================================================================
* AMENDMENT HISTORY:
* -Ini- -Date-- -Id---- -Description----------------------------------------
*
*=======================================================================
* This is public domain software. No guarantee as to suitability or
accuracy is
* given or implied. Users use this code entirely at their own risk.
*************************************************************************;

**** Session settings ****;

* Session options;

options nofmterr;

* Create SIMLR Sstandarised Data Extracts;

%macro makedata(load);

    %if %upcase(&load) eq YES %then %do;
        %include "S:\1000206_simlr_raw\SIMLRGEO\_merged_extracts\SIMLR
Standardised Data Extracts Code.sas";
    %end;

%mend makedata;
%makedata(NO);

************************************************************************
**** Create coding formats
***
************************************************************************;

proc sql;
    create table seifafmt as
        select
```

```
            "$sa1adv" as fmtname,
            ses_sa1_2011_code as start,
            ses_irsad_2011_quintile as label,
            "C" as type
        from ses;
    insert into seifafmt
        select
            "$sa1dis" as fmtname,
            ses_sa1_2011_code as start,
            ses_irsd_2011_quintile as label,
            "C" as type
        from ses;
    insert into seifafmt
        select
            "$sa1ier" as fmtname,
            ses_sa1_2011_code as start,
            ses_ier_2011_quintile as label,
            "C" as type
        from ses;
    insert into seifafmt
        select
            "$sa1ieo" as fmtname,
            ses_sa1_2011_code as start,
            ses_ieo_2011_quintile as label,
            "C" as type
        from ses;
quit;

proc format library=work cntlin=seifafmt;
    value age05grp
        18-24   = "18-24"
        25-29   = "25-29"
        30-34   = "30-34"
        35-39   = "35-39"
        40-44   = "40-44"
        45-49   = "45-49"
        50-54   = "50-54"
        55-59   = "55-59"
        60-64   = "60-64"
        65-69   = "65-69"
        70-74   = "70-74"
        75-79   = "75-79"
        80-84   = "80-84"
        85-high = "85+"
        other   = " ";
run;

*************************************************************************
**** Create analytic data set
****
*************************************************************************;

**** Reference list of test codes and names ****;

* Create Test Name display format;

proc sql;
    create table TestName as
        select distinct
            res_test_code,
            res_test_mnemonic,
```

```
            res_test_name
    from WORK.RES
    order by res_test_code;
quit;

**** Create Episode-Results table with one result per day ****;

* NOTE: Some patients have multiple results for the same test per day. As
we
        have no reason to prefer one over the other(s) we simply average
over
        the multiple results to ensure one result per test per patient per
day;

proc sql;
    create table WORK.EPIRES as
        select
            a.epi_patient_id,
            a.epi_sex,
            a.epi_collection_date,
            a.epi_collection_age,
            a.epi_study_id,
            b.res_test_code,
            b.res_test_mnemonic,
            mean(b.res_test_value_num) as res_test_value_num length=8
informat=best. format=best.,
            count(*) as res_test_count length=3 informat=best. format=best.
        from WORK.EPI as a inner join WORK.RES as b on
a.epi_episode_id=b.res_episode_id
        where a.epi_sex ~= "P" /*and year(a.epi_collection_date) >= 2012*/
        group by
            a.epi_patient_id,
            a.epi_sex,
            a.epi_collection_date,
            a.epi_collection_age,
            a.epi_study_id,
            b.res_test_code,
            b.res_test_mnemonic
        order by
            a.epi_patient_id,
            b.res_test_code,
            a.epi_collection_date;
    create index epi_patient_id
        on work.epires (epi_patient_id);
quit;

**** Calculate diabetes status using SIMLR algorith ****;

/*

Incident cases

1.  No HbA1c before 01-JAN-2012 AND

    i.  1+ HbA1c on or after 01-JAN-2012 AND

        a.    HbA1c >= 6.5% (WHO, 2011; ABS, 2013) OR
        b.    Fasting BSL >= 7.0 mmol/L (ABS, 2013) within 12 months of
HbA1c
```

```
Prevalent cases

1.  1+ HbA1c on or after 01-JAN-2012 AND

    i.  1+ HbA1c before 01-JAN-2012 AND

        a. HbA1c >= 6.5%  (WHO, 2011; ABS, 2013) OR
        b. Fasting  BSL >= 7.0 mmol/L  (ABS, 2013) within 12 months of
HbA1c

*/

* Extract HbA1c results and match to Fasting BSL results within 12 months;

proc sql;
    create table _diabetes as
        select
            epi_patient_id,
            epi_collection_date,
            res_test_code,
            res_test_mnemonic,
            res_test_value_num
        from epires
        where res_test_code = "E0190" and year(epi_collection_date) <= 2017
        order by epi_patient_id, epi_collection_date;
    create table diabetes as
        select
            a.epi_patient_id,
            a.epi_collection_date,
            a.res_test_code,
            a.res_test_mnemonic,
            a.res_test_value_num,
            b.epi_collection_date as fbsl_collection_date,
            b.res_test_value_num as fbsl_test_value_num,
            case
                when (b.epi_collection_date = .) then (a.res_test_value_num
>= 6.5)
                else (b.res_test_value_num >= 7.0) * (-365 <=
(a.epi_collection_date - b.epi_collection_date) <= 365)
            end as diabetes length=3 informat=1. format=1.,
            min(case
                when (calculated diabetes = 1) then a.epi_collection_date
                else "31DEC2999"d
            end) as diab_diag_date length=8 informat=best.
format=yymmddd10.
        from _diabetes as a
            left join (select * from epires where res_test_code = "C0210")
as b
        on a.epi_patient_id = b.epi_patient_id
        group by a.epi_patient_id
        having diabetes = 1 and (2012 <= year(a.epi_collection_date) <=
2017)
        order by a.epi_patient_id, a.epi_collection_date,
b.epi_collection_date;
    drop table _diabetes;
quit;

* Extract first (diagnostic) HbA1c result:

  If 01-JAN-2003 to 31-DEC-2011 then prevalent case
  If 01-JAN-2012 to 31-DEC-2017 then incident case;
```

```
data diabetes;
    set diabetes;
    by epi_patient_id;

        diabetes = diabetes + (diab_diag_date <= "31DEC2011"d);

        if first.epi_patient_id then output;
run;

**** Select most recent test result for each patient ****;

* Select in-scope tests;

data _epires;
    set epires(drop = res_test_count);
    by epi_patient_id res_test_code epi_collection_date;
    where res_test_code in ("C0080", "C0090", "C0169", "C0210", "C1855",
"C7920", "E0190");

        if last.res_test_code and year(epi_collection_date) >= 2012 then
output;
run;

* Transpose into long format;

proc transpose data=_epires out=_epires;
    var epi_collection_date
        epi_collection_age
        epi_study_id
        res_test_value_num;
    by  epi_patient_id
        epi_sex
        notsorted res_test_mnemonic;
run;
* Get value ranges;

proc sql;
    create table test_value_ranges as
        select
            res_test_mnemonic,
            min(col1) as minimum,
            max(col1) as maximum
        from _epires
        where index(_name_, "num") >= 1
        group by res_test_mnemonic;
quit;

* Create unique names for variable properties for each test;

data _epires;
    attrib epi_patient_id                                          label = "
"
           epi_sex                                                 label = "
"
           res_test_variable length=$25 informat=$25. format=$25. label = "
"
           value                                                   label = "
";
    set _epires(rename=(col1=value));
```

```
        res_test_variable = compress(res_test_mnemonic || substr(_name_,
4));

    drop res_test_mnemonic _name_;
run;

* Transpose into wide format;

proc transpose data=_epires out=_epires;
    var value;
    by epi_patient_id
       epi_sex ;
    id res_test_variable;
    idlabel res_test_variable;
run;
```

```
**** Create final data set ****;

* Combine test results, diabetes status, and geographic variables;

proc sql noprint;
    select max(length(put(epi_patient_id, best.))) * 2 into :idlen from
_epires;
    create table rmbst288 as
        select
            a.epi_patient_id,
            trim(left(put(a.epi_patient_id, best.))) as enc_project_id
length=&idlen label="Encrypted patient identifier",
            . as raw_project_id length=8,
            a.epi_sex label="Sex",

            case
                when (b.diabetes = .) then "No"
                when (b.diabetes = 1) then "Yes (incident)"
                when (b.diabetes = 2) then "Yes (prevalent)"
                else ""
            end as diab_status length=15 informat=$15. format=$15.
label="Diabetes status from SIMLR classification algorithm",
            b.diab_diag_date as diab_date label="Diabetes status=yes date
from SIMLR classification algorithm",

            /*a.fbsl_study_id, */
            a.fbsl_collection_date format=yymmddd10. label="Fasting BSL
collection date",
            put(a.fbsl_collection_age, age05grp5.) as
fbsl_collection_age05grp length=5 informat=$5. format=$5. label="Fasting
BSL age group",
            a.fbsl_test_value_num label="Fasting BSL test value
(mmol/L)",
            c.geo_match_status as fbsl_geo_match_status label="Fasting
BSL geocoding match status",
            c.geo_sa1_2011_code as fbsl_sa1_2011_code label="Fasting BSL
SA1 of residence",
            put(c.geo_sa1_2011_code, $sa1adv2.) as
fbsl_irsad_2011_quintile length=2 informat=$2. format=$2. label="Fasting
BSL IRSAD quintile",
            put(c.geo_sa1_2011_code, $sa1dis2.) as
fbsl_irsd_2011_quintile length=2 informat=$2. format=$2. label="Fasting
BSL IRSD quintile",
            put(c.geo_sa1_2011_code, $sa1ier2.) as fbsl_ier_2011_quintile
length=2 informat=$2. format=$2. label="Fasting BSL IER quintile",
            put(c.geo_sa1_2011_code, $sa1ieo2.) as fbsl_ieo_2011_quintile
length=2 informat=$2. format=$2. label="Fasting BSL IEO quintile",

            /*a.hba1c_study_id, */
            a.hba1c_collection_date format=yymmddd10. label="HbA1c
collection date",
            put(a.hba1c_collection_age, age05grp5.) as
hba1c_collection_age05grp length=5 informat=$5. format=$5. label="HbA1c
collection age group",
            a.hba1c_test_value_num label="HbA1c test value (%)",
            d.geo_match_status as hba1c_geo_match_status label="HbA1c
geocoding match status",
            d.geo_sa1_2011_code as hba1c_sa1_2011_code label="HbA1c SA1
of residence",
```

```
            put(d.geo_sa1_2011_code, $sa1adv2.) as
hba1c_irsad_2011_quintile length=2 informat=$2. format=$2. label="HbA1c
IRSAD quintile",
            put(d.geo_sa1_2011_code, $sa1dis2.) as
hba1c_irsd_2011_quintile length=2 informat=$2. format=$2. label="HbA1c
IRSD quintile",
            put(d.geo_sa1_2011_code, $sa1ier2.) as
hba1c_ier_2011_quintile length=2 informat=$2. format=$2. label="HbA1c IER
quintile",
            put(d.geo_sa1_2011_code, $sa1ieo2.) as
hba1c_ieo_2011_quintile length=2 informat=$2. format=$2. label="HbA1c IEO
quintile",

            /*a.hdl_study_id, */
            a.hdl_collection_date format=yymmddd10. label="HDL collection
date",
            put(a.hdl_collection_age, age05grp5.) as
hdl_collection_age05grp length=5 informat=$5. format=$5. label="HDL
collection age group",
            a.hdl_test_value_num label="HDL test value (mmol/L)",
            e.geo_match_status as hdl_geo_match_status label="HDL
geocoding match status",
            e.geo_sa1_2011_code as hdl_sa1_2011_code label="HDL SA1 of
residence",
            put(e.geo_sa1_2011_code, $sa1adv2.) as
hdl_irsad_2011_quintile length=2 informat=$2. format=$2. label="HDL IRSAD
quintile",
            put(e.geo_sa1_2011_code, $sa1dis2.) as hdl_irsd_2011_quintile
length=2 informat=$2. format=$2. label="HDL IRSD quintile",
            put(e.geo_sa1_2011_code, $sa1ier2.) as hdl_ier_2011_quintile
length=2 informat=$2. format=$2. label="HDL IER quintile",
            put(e.geo_sa1_2011_code, $sa1ieo2.) as hdl_ieo_2011_quintile
length=2 informat=$2. format=$2. label="HDL IEO quintile",

            /*a.tc_study_id, */
            a.tc_collection_date format=yymmddd10. label="TC collection
date",
            put(a.tc_collection_age, age05grp5.) as
tc_collection_age05grp length=5 informat=$5. format=$5. label="TC
collection age group",
            a.tc_test_value_num label="TC test value (mmol/L)",
            f.geo_match_status as tc_geo_match_status label="TC geocoding
match status",
            f.geo_sa1_2011_code as tc_sa1_2011_code label="TC SA1 of
residence",
            put(f.geo_sa1_2011_code, $sa1adv2.) as tc_irsad_2011_quintile
length=2 informat=$2. format=$2. label="TC IRSAD quintile",
            put(f.geo_sa1_2011_code, $sa1dis2.) as tc_irsd_2011_quintile
length=2 informat=$2. format=$2. label="TC IRSD quintile",
            put(f.geo_sa1_2011_code, $sa1ier2.) as tc_ier_2011_quintile
length=2 informat=$2. format=$2. label="TC IER quintile",
            put(f.geo_sa1_2011_code, $sa1ieo2.) as tc_ieo_2011_quintile
length=2 informat=$2. format=$2. label="TC IEO quintile",

            /*a.acr_study_id, */
            a.acr_collection_date format=yymmddd10. label="ACR collection
date",
            put(a.acr_collection_age, age05grp5.) as
acr_collection_age05grp length=5 informat=$5. format=$5. label="ACR
collection age group",
            a.acr_test_value_num label="ACR test value (mcg/L)",
```

```
            g.geo_match_status as acr_geo_match_status label="ACR
geocoding match status",
            g.geo_sa1_2011_code as acr_sa1_2011_code label="ACR SA1 of
residence",
            put(g.geo_sa1_2011_code, $sa1adv2.) as
acr_irsad_2011_quintile length=2 informat=$2. format=$2. label="ACR IRSAD
quintile",
            put(g.geo_sa1_2011_code, $sa1dis2.) as acr_irsd_2011_quintile
length=2 informat=$2. format=$2. label="ACR IRSD quintile",
            put(g.geo_sa1_2011_code, $sa1ier2.) as acr_ier_2011_quintile
length=2 informat=$2. format=$2. label="ACR IER quintile",
            put(g.geo_sa1_2011_code, $sa1ieo2.) as acr_ieo_2011_quintile
length=2 informat=$2. format=$2. label="ACR IEO quintile",

            /*a.egfr_study_id, */
            a.egfr_collection_date format=yymmddd10. label="eGFR
collection date",
            put(a.egfr_collection_age, age05grp5.) as
egfr_collection_age05grp length=5 informat=$5. format=$5. label="eGFR
collection age group",
            a.egfr_test_value_num label="eGFR test value
(mL/min/1.73m2)",
            h.geo_match_status as egfr_geo_match_status label="eGFR
geocoding match status",
            h.geo_sa1_2011_code as egfr_sa1_2011_code label="eGFR SA1 of
residence",
            put(h.geo_sa1_2011_code, $sa1adv2.) as
egfr_irsad_2011_quintile length=2 informat=$2. format=$2. label="eGFR
IRSAD quintile",
            put(h.geo_sa1_2011_code, $sa1dis2.) as
egfr_irsd_2011_quintile length=2 informat=$2. format=$2. label="eGFR IRSD
quintile",
            put(h.geo_sa1_2011_code, $sa1ier2.) as egfr_ier_2011_quintile
length=2 informat=$2. format=$2. label="eGFR IER quintile",
            put(h.geo_sa1_2011_code, $sa1ieo2.) as egfr_ieo_2011_quintile
length=2 informat=$2. format=$2. label="eGFR IEO quintile",

            /*a.bmi_study_id, */
            a.bmi_collection_date format=yymmddd10. label="BMI collection
date",
            put(a.bmi_collection_age, age05grp5.) as
bmi_collection_age05grp length=5 informat=$5. format=$5. label="BMI
collection age group",
            a.bmi_test_value_num label="BMI test value (kg/m2)",
            i.geo_match_status as bmi_geo_match_status label="BMI
geocoding match status",
            i.geo_sa1_2011_code as bmi_sa1_2011_code label="BMI SA1 of
residence",
            put(i.geo_sa1_2011_code, $sa1adv2.) as
bmi_irsad_2011_quintile length=2 informat=$2. format=$2. label="BMI IRSAD
quintile",
            put(i.geo_sa1_2011_code, $sa1dis2.) as bmi_irsd_2011_quintile
length=2 informat=$2. format=$2. label="BMI IRSD quintile",
            put(i.geo_sa1_2011_code, $sa1ier2.) as bmi_ier_2011_quintile
length=2 informat=$2. format=$2. label="BMI IER quintile",
            put(i.geo_sa1_2011_code, $sa1ieo2.) as bmi_ieo_2011_quintile
length=2 informat=$2. format=$2. label="BMI IEO quintile",

            nmiss(a.fbsl_test_value_num,
                  a.hba1c_test_value_num,
                  a.hdl_test_value_num,
```

```
                    a.tc_test_value_num,
                    a.acr_test_value_num,
                    a.egfr_test_value_num,
                    a.bmi_test_value_num) as nmiss_test_results length=3
informat=1. format=1. label="Number of missing test results",
            case
                when (calculated nmiss_test_results = 0) then "Y"
                else "N"
            end as complete_case length=1 informat=$1. format=$1. label
"Complete case (no missing test results)"
        from _epires as a
            left join diabetes as b on a.epi_patient_id=b.epi_patient_id
            left join geo as c on a.epi_patient_id=c.geo_patient_id and
a.fbsl_study_id = c.geo_study_id
            left join geo as d on a.epi_patient_id=d.geo_patient_id and
a.hba1c_study_id = d.geo_study_id
            left join geo as e on a.epi_patient_id=e.geo_patient_id and
a.hdl_study_id = e.geo_study_id
            left join geo as f on a.epi_patient_id=f.geo_patient_id and
a.tc_study_id = f.geo_study_id
            left join geo as g on a.epi_patient_id=g.geo_patient_id and
a.acr_study_id = g.geo_study_id
            left join geo as h on a.epi_patient_id=h.geo_patient_id and
a.egfr_study_id = h.geo_study_id
            left join geo as i on a.epi_patient_id=i.geo_patient_id and
a.bmi_study_id = i.geo_study_id
        order by a.epi_patient_id;
quit;

**** Document data set ****;

proc contents data=rmbst288 OUT=rmbst288_vars(keep=name label type
varnum) noprint nodetails;
run;

    proc sort data=rmbst288_vars;
        by varnum;
    run;

proc export data=rmbst288_vars

outfile="\\ad.uow.edu.au\Shares\1000389_toms\Documentation\RMBST288_TOMS_
RENIN_PHD_DATA_EXTRACT_VARS_20180204.CSV"
    dbms=csv
    replace;
run;

**** Output extract to UOW file share ****;

filename extract
"%sysfunc(getoption(work))\RMBST288_TOMS_RENIN_PHD_DATA_EXTRACT_20180204.
CSV";

* Output in CSV format;

proc export data=rmbst288 outfile=extract dbms=csv replace;
run;

* Open ZIP package;

ods package open;
```

```
* Add CSV data file to the ZIP package;

ods package add file=extract;

* Finalise package compression;

ods package publish archive
    properties(archive_name =
"RMBST288_TOMS_RENIN_PHD_DATA_EXTRACT_20180204.zip"
               archive_path =
"\\ad.uow.edu.au\Shares\1000389_toms\Data");

* Close the package;

ods package close;
```

Appendix V: Human Research Ethics Committee Approval Letter

HREC Approval of Application 2017/124

irma-support@uow.edu.au <irma-support@uow.edu.au>
Mon 03/04/2017 15:30

To: abonney@uow.edu.au <abonney@uow.edu.au>
Cc: xfeng@uow.edu.au <xfeng@uow.edu.au>; dmayne@uow.edu.au <dmayne@uow.edu.au>; Renin Melkias Baby Selvi Toms <rmbst288@uowmail.edu.au>; rso-ethics@uow.edu.au <rso-ethics@uow.edu.au>

Dear Dr Bonney,

I am pleased to advise that the application detailed below has been **approved**.

Ethics Number:	2017/124
Approval Date:	03/04/2017
Expiry Date:	02/04/2018
Project Title:	An epidemiological study of chronic disease parameters at regional level, in view of developing a cardiometabolic risk map to facilitate regional planning activities.
Researcher/s:	Feng Xiaoqi; Mayne Darren; Melkias Baby Selvi Toms Renin; Bonney Andrew
Documents Approved:	UOW Application for HREC Approval V1 17032017 All Investigator Detail forms are noted

Sites:

Site	Principal Investigator for Site
University of Wollongong	Professor Andrew Bonney

The HREC has reviewed the research proposal for compliance with the *National Statement on Ethical Conduct in Human Research* and approval of this project is conditional upon your continuing compliance with this document. Compliance is monitored through progress reports; the HREC may also undertake physical monitoring of research.

Approval is granted for a twelve month period; extension of this approval will be considered on receipt of a progress report prior to the expiry date. Extension of approval requires:

- The submission of an annual progress report and a final report on completion of your project.
- Approval by the HREC of any proposed changes to the protocol or investigators.
- Immediate report of serious or unexpected adverse effects on participants.
- Immediate report of unforeseen events that might affect the continued acceptability of the project.

If you have any queries regarding the HREC review process or your ongoing approval please contact the Ethics Unit on 4221 3386 or email rso-ethics@uow.edu.au.

Yours sincerely,

Dr Susan Thomas,
Chair, UOW & ISLHD Health and Medical Human Research Ethics Committee

The University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical HREC is constituted and functions in accordance with the NHMRC National Statement on Ethical Conduct in Human Research. The processes used by this HREC to review multi-centre research proposals have been certified by the National Health and Medical Research Council.

Supplementary Materials

Supplementary Material: 1 (Published version of research article 1)

Toms et al. Int J Health Geogr (2019) 18:1
https://doi.org/10.1186/s12942-018-0165-5

International Journal of Health Geographics

REVIEW Open Access

Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature

Renin Toms[1*], Andrew Bonney[1,2], Darren J. Mayne[1,2,3], Xiaoqi Feng[2,4] and Ramya Walsan[1]

Abstract

Introduction: A growing number of publications report variation in the distribution of cardiometabolic risk factors (CMRFs) at different geographic scales. A review of these variations may help inform policy and health service organisation.

Aim: To review studies reporting variation in the geographic distribution of CMRFs and its association with various proxy measures of area-level socioeconomic disadvantage (ASED) among the adult (≥ 18 years) population across the world.

Methods: A systematic search for published articles was conducted in four databases (MEDLINE (Ovid), PubMed, Scopus and Web of Science) considering the interdisciplinary nature of the review question. Population-based cross-sectional and cohort studies on geographic variations of one or more biological proxies of CMRFs with/without an analysed contextual association with ASED were included. Two independent reviewers screened the studies and PRISMA guidelines were followed in the study selection and reporting.

Result: A total of 265 studies were retrieved and screened, resulting in 24 eligible studies. The review revealed reports of variation in the distribution of CMRFs, at varying geographic scales, in multiple countries. In addition, consistent associations between ASED and higher prevalence of CMRFs were demonstrated. The reports were mainly from industrialised nations and small area geographic units were frequently used.

Conclusion: Geographic variation in cardiometabolic risk exists across multiple spatial scales and is positively associated with ASED. This association is independent of individual-level factors and provides an imperative for area-based approaches to informing policy and health service organisation. The study protocol is registered in *International prospective register of systematic reviews* (Register No: CRD42018115294) PROSPERO 2018.

Keywords: Cardiometabolic risk factors, Area-level socioeconomic disadvantage, Geographic variation

Introduction

Cardiovascular disease (CVD) associated metabolic risk factors represent major global public health concerns. CVD is the leading cause of human death, accounting for 17.7 million (31%) of the 56.4 million total deaths reported worldwide in 2015 [1]. Coronary heart disease (7.4 million) and stroke (6.7 million) were responsible for the greatest mortality within CVD and have remained the leading cause for mortality for the last 15 years [2]. CVD and its associated metabolic risk factors are listed in the top 15 causes of disability adjusted life years (DALY) globally [3]. In keeping with historical trends, deaths due to CVD are projected to increase steeply and reach more than 23.6 million annually by 2030 [4].

An important way to control CVD is by focussing on reducing associated metabolic risk factors. In low resource settings, vulnerable and disadvantaged groups are more likely to be exposed to unhealthy products and practices and develop metabolic risk factors for the development of CVD [5]. Cardiometabolic risk factors

*Correspondence: rmbst288@uowmail.edu.au
[1] School of Medicine, University of Wollongong, Wollongong, Australia
Full list of author information is available at the end of the article

(CMRFs) such as diabetes mellitus (DM), hyperlipidaemia, high body mass index (BMI), and chronic kidney disease (CKD) can predispose and worsen CVD. Individual level approaches to prevent and control these risk factors have demonstrated limited success as evidenced by its increasing rates [6–8]. Thus it is important, in addition, to discern the contextual associations of development of these risk factors to assist in mitigating this global epidemic.

Geographic inequalities in the distribution of CMRFs at varying scales are reported in multiple studies from different countries in association with area-level socioeconomic disadvantage (ASED). Reviewing the area level distribution patterns and associated area level disadvantages reported in these studies may deepen our understanding of the higher prevalence of CMRFs in some geographic areas. Most recent relevant reviews in this area have broadly covered the influence of physical, social and service environment characteristics on CVD risk [9–12]. However, the potentially important influence of ASED is critically under-examined. Systematic synthesis of evidence regarding this globally reported variation and association can inform policy development and healthcare service planning to detail area level approaches, in addition to individual level measures, to prevent and control CMRFs effectively.

Therefore, the questions attempted to answer in this review are: *Is there any geographic variation in the distribution of CMRFs among the adult population (aged 18 years and above) across the world, and is this variation associated with ASED?* The studies expected to include were epidemiological or population based cross sectional and/or cohort studies.

Methods

A review protocol was developed and registered in International prospective register of systematic reviews, PROSPERO 2018 (Register No: CRD42018115294) Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018115294.

Four databases; MEDLINE (Ovid), PubMed, Scopus and Web of science databases were chosen for the search, considering the breadth of fields they cover and the interdisciplinary nature of the review question. Also, handsearches of related articles served as 'other sources' of studies. The database search strategy commenced with two general search domains [1]: studies on CMRFs in singular and composite forms; and [2] geographic and spatial health studies. An intersectional retrieval of studies from both these domains yielded a narrower list of studies on geographic variation in CMRFs. A third domain [3] studies addressing area-level measures of socioeconomic disadvantages were further intersected with the retrieved studies to create a focal list of studies addressing geographic association of CMRFs with ASED. This approach maximised the number of potentially eligible studies identified compared to using single domain searches. Figure 1 conceptualizes the major search domains and their intersections used in the review.

The review included epidemiological or population-based cross-sectional and cohort studies on: geographic variation of one or more biological proxies of CMRFs, with/without an analysed contextual association with ASED. Obesity, diabetes mellitus (DM), hyperlipidaemia, and indices of low kidney function were the included biological proxies of CMRFs. Hypertension is included only when reported with other biological proxies of CMRFs, but not independently considering its limited summation into an overall cardiometabolic risk in an individual. Studies involving type 1 DM and gestational DM were excluded as they were out of scope for the current review pertaining to the geographic or area based contexts of the CMRFs. Studies measuring area-level characteristics other than ASED were also excluded.

All search outcomes were limited to: human studies; adult population (≥18 years); and availability in English language. The initial search included studies from year 1995; and latter it was modified to 01/01/2001 due to minimal publications on the review topic between the years 1995–2000. The search was last updated on 30/11/2018. Adopted search strategy in Ovid MEDLINE, and search result URLs of remaining databases are available in Additional file 1.

All retrieved studies were screened by two independent reviewers (RT and RW) in three stages to reduce the risk of bias. In stage 1, articles from all databases were combined and screened to remove duplicates. Titles and abstracts of remaining articles were screened for eligibility in stage 2. The final stage of study selection was done after full text reading of the remaining studies. Qualities of the individual studies were assessed using the STROBE checklist for cohort, case–control and cross-sectional studies (www.strobe-statement.org). The second coder repeated all three stages in parallel, and selected studies were matched at the conclusion of each stage and any differences were resolved by consensus and arbitration. Other review team members (AD, DJM and XF) served as additional reviewers when required.

Data extraction and coding of the chosen studies were carried out using two pilot-tested templates for consistency. Template 1 focused on the geographic variation in CMRFs and was used to extract information on author, year, nation, study design, sample size and characteristics, geographic unit of reporting, studied CMRFs, and the study outcome. Data on behavioural risk factors were not extracted as these were not included in the current

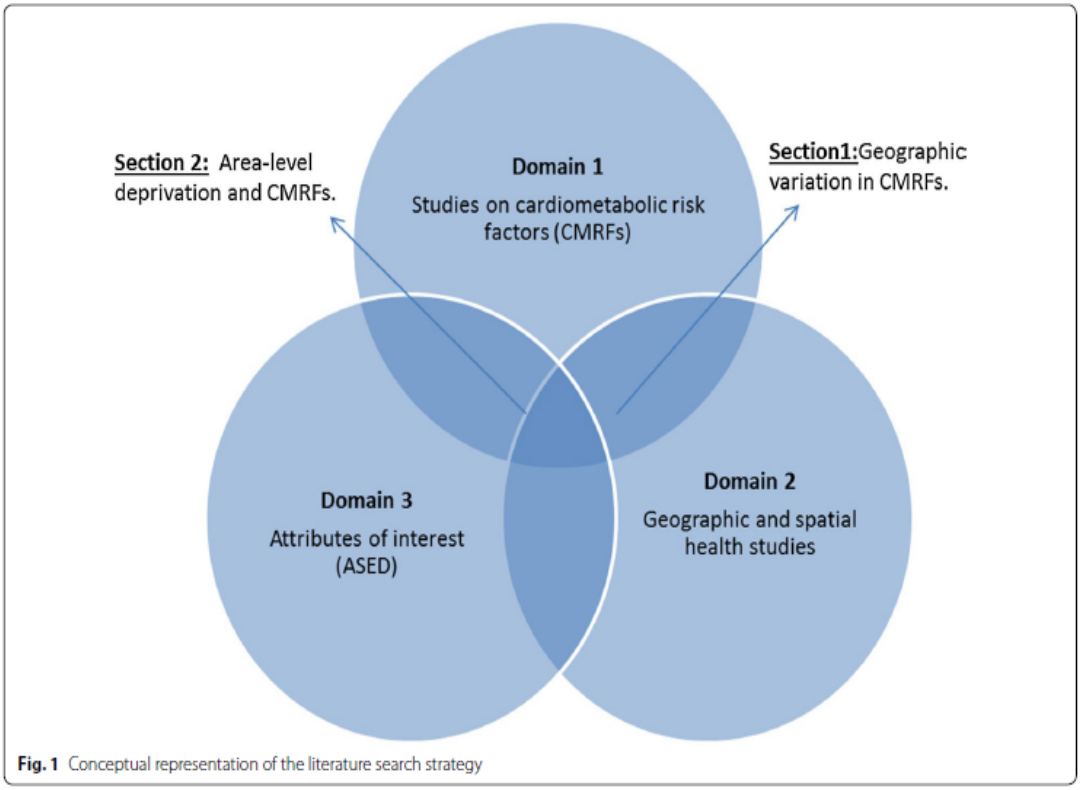

Fig. 1 Conceptual representation of the literature search strategy

review. Template 2 addressed the association of ASED and cardiometabolic risk prevalence, and extracted additional data on the reported proxies of ASED and its association status. An additional template was used for thematic mapping of the data in included studies for further qualitative syntheses. Study origin, representation, nature of problem, ecological context, and evidence strength were the mapped themes.

The two independent review authors extracted and coded the data, and any discrepancies were resolved through discussions between the authors. Summary measures used in this review are descriptive and based on the frequency of relevant studies to its denominator. Endnote software was used to keep track of the bibliographic details of the studies throughout the selection and data extraction process.

Results

A total of 265 individual studies were retrieved from four electronic databases (n = 251) and hand searches of reference lists (n = 14). Studies from electronic data bases included 91 Ovid Medline, 80 PubMed, 58 Scopus, and 22 Web of science publications.

Figure 2 shows the screening process as per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses PRISMA guidelines (www.prisma-statement.org) [13]. Stage 1 screening combined studies from all sources and removed the duplicates (n = 99). Duplicates in removed order: Ovid Medline (n = 0); PubMed (n = 80); Scopus (n = 10); Web of science (n = 3); and hand-searches (n = 6). After removing duplicates, 166 studies were forwarded for stage 2 screening. Stage 2 screening excluded 130 studies based on title and abstract screens, forwarding 36 studies for the full text screen. Studies excluded in stage 2 mainly addressed genetic, cellular, instrumental or pharmacological research regarding CMRFs. Studies on type 1 DM, paediatric or juvenile DM and gestational DM were also excluded at this stage as per the exclusions stated. Stage 3 screening carefully considered the whole full text of articles and 12 records were excluded with reason (list available in Additional file 2) leaving 24 studies for the systematic synthesis. PRISMA 2009 guidelines are followed in reporting the review and the checklist available in Additional file 3.

The review is structured into three sections. Screened research articles retrieved through 'AND' intersections of search domain 1 and 2 (n = 8) are reviewed in

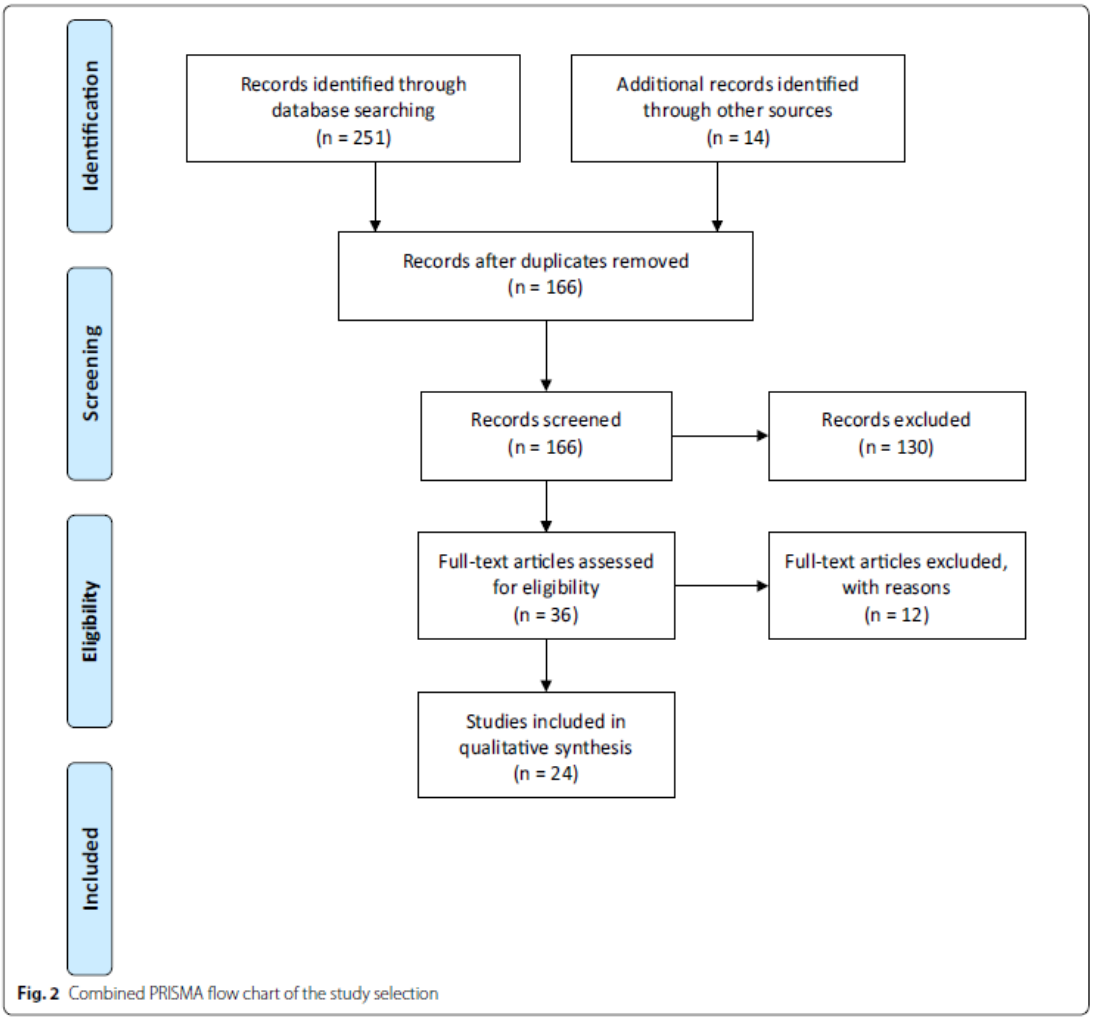

Fig. 2 Combined PRISMA flow chart of the study selection

"Introduction" section: Geographic variation in the prevalence of cardiometabolic risk factors. Screened articles retrieved by intersecting domains 1 and 3 (n = 16) are reviewed in "Methods" section: Area level deprivation and cardiometabolic risk prevalence. Overall synthesis based on the total reviewed studies (n = 24) is presented in "Results" section: Overall synthesis of the studies.

Geographic variation in the prevalence of cardiometabolic risk factors

Table 1 summarizes the eight studies reviewed under this section [13–20]. Geographic variation in the prevalence of one or more CMRFs is reported in each of these studies. Most of the studies (7/8) reported hyperglycaemia as an important biomarker displaying geographic variation in cardiometabolic risk [13–18, 20,] followed by dyslipidaemia (4/8), body mass index (4/8), blood pressure (BP) (3/8) and reduced glomerular filtration rate (GFR) (1/8).

All studies reported geographic variation in the prevalence of CMRFs, regardless of the geographic unit of analysis used [13–20]. Most of these studies were from Europe (4/8), predominantly from Western Europe (3/8) [15, 16, 18, 19]. These reports were from UK, [19] Spain, [18] France, [16] and Luxembourg [15]. In the UK, geographic variation in the prevalence of risk factors such as obesity, smoking, diabetes, hypertension and high cholesterol were reported across four main

Table 1 List of studies reviewed on geographic variation in CMRFs

	First author Country	Sample Age group	Design Geographic units	CMRFs[a] (data source)	Outcome
1	Lawlor et al. UK [19]	4286 (women) 60–79 years	Cross-sectional 4 regions within country	HT, BMI, LDL, TC (data collected)	Geographic variation
2	Barker et al. USA [20]	813,498 DM ≥18 years	Cross-sectional 644 counties in proximity	DM prevalence ≥ 11.0%, (behavioral risk factor surveillance)	Geographic diabetic belt
3	Valdes et al. Spain [18]	5103 adults ≥ 18 years	Cross-sectional 2 region within country	BP, BMI, FPG,TC, WC (Di@bet.es study)	Geographic coherence
4	Astell-Burt et al. Australia [37]	114,755 adults ≥ 45 years	Cross-sectional ~ 40 local government areas (2011 significant urban boundary)	DM (the 45 and up study)	Geographic variation
5	Zhou et al. China [17]	98,058 adults > 18 years	Cross-sectional multilevel 31 provinces in country	DM (National Health Survey)	Geographic variation
6	Paquet et al. AU-France [16]	Au: 3893 (≥ 18 years) Fr: 6430 (30–79 years)	Cross-sectional multilevel Au: 767 CDs (SS, POA, SLA LGA). Fr: 1866 IRIS (TRIRIS, Municipalities)	BP,BMI,WC,FG,HbA1c,HR, TC,HDL, TG, (Au: NWAHS study, Fr: RECORD cohort study)	Geographic clustering
7	Alkerwi et al. Luxemberg [15]	1432 subjects 18–69 years.	Cross-sectional multilevel 106 municipalities (12 cantons)	BMI, FPG,TC, GFR (ORISCAV-LUX national survey)	Geographic variation
8	Oh et al. South Korea [13]	228,921 people ≥ 19 years	Cross-sectional 230 administrative districts	HT, DM (Korean Community Health Surveys)	Geographic clustering

AU Australia, *CD* census collection district, *POA* postal area, *SLA* statistical local area, *LGA* local government area, *IRIS* Ilôts regroupés pour l'information statistique, *TRIRIS* groups of around three IRIS areas, *BP* blood pressure, *BMI* body mass index, *DM* diabetes mellitus, *FBG* fasting plasma glucose, *FPG* fasting glucose, *HbA1c* glycated haemoglobin, *HR* heart rate, *HT* hypertension, *TC* total cholesterol; *TG* triglycerides, *LDL* low density lipoprotein, *GFR* glomerular filtration rate, *WC* waist circumference

[a] Behavioural risk factors excluded

regions: South England: Midlands: and Wales: Scotland: and North England [19]. A higher prevalence of CMRFs was reported in southern Spain (Andalusia), which was found in close association with sedentary lifestyle and markers of socioeconomic disadvantage [18]. Variation in the distribution of diabetes, high BMI (≥ 25 kg/m^2), abdominal obesity, hypertension, high cholesterol and low glomerular filtration rate were reported at both canton and municipality levels in Luxemburg, Western Europe [15].

BMI and resting heart rate were reported to have greater geographic variation among matched cohorts in France and Australia [16].

Other reports in this section were from Oceania (2/8), East Asia (2/8) and North America 1/8)—sourced from Australia, China, South Korea and US [13, 14, 16, 17, 20]. A geographic variation of 42% was reported in the odds of being diagnosed with DM among adults in Sydney, Australia [14]. In another Australian metropolitan based cohort, glycated haemoglobin (HbA1c) was reported to have geographic variation among matched cohorts in Australia and France [16]. In China, significant variation in the regional prevalence of diabetes was reported after adjusting for age, sex and urban/rural socioeconomic circumstances [17]. Geographic clustering of cardiometabolic risk factors were reported at administrative district level in South Korea [13]. The presence of a 'diabetic belt' with higher prevalence of diagnosed diabetes (> 11.0%) was reported in the United States, consisting of 644 counties in its 15 mostly southern states [20]. Though the risk profiles and parameters varied, all these studies consistently reported geographic variation in its CMRFs.

The geographic scales of area-based units reported in all these studies ranged from large regions [17–19] within countries to smaller jurisdictional administration units [13–16, 20] and trended towards smaller geographic areas over time. Easily accessible pre-existing geographic units and boundaries were used in these studies but most weren't explicit on the spatial extention and average population within their geographic units. Three studies had relied only on self-reports on anthropometric, behavioural, biochemical, physiological and diagnostic categories of data, risking for recall bias and misclassifications [13, 14, 20].

Area level deprivation and cardiometabolic risk prevalence

Table 2 summarises the 16 studies reviewed under this section [21–36]. Reported studies were mainly from Europe (7/16) and North America (7/16), followed by Oceania (1/16) and South America (1/16). Studies from Europe were predominantly reported from the western

Table 2 List of studies reviewed on the association of area-level deprivation and cardiometabolic risk prevalence

	First author Country	Sample Age group	Design Spatial unit	CMRFs* (data source)	Proxies of ASED (data source)	Association
1	Bonney et al. Australia [26]	91,776 adults 55.2 ± 15.66	Cross-sectional higherarchical 631 census collection districts	BMI (the SIMLR study)	Index of Relative Socioeconomic Disadvantage (Australian Census 2006)	+ve (women)
2	Unger et al. USA [27]	5805 adults 45–84 years	Prospective cohort higherarchical Census tract level	BMI, BP, BS, TC- CVH score (The MESA study)	Neighbourhood SES (constructed summary score)	+ve
3	Maier et al. Germany [29]	33,690 adults < 30 years	Cross-sectional design 412 districts	T2DM, obesity (GEDA national health interview survey)'	German Index of Multiple Deprivation score (assessed by GIMD)	+ve (women)
4	Silhol et al. France [33]	19,808 adults 35–50 years	Cross-sectional cohort Municipality level	Incidence of CHD (French GAZEL cohort Data)	Area socio - economic position (French Census 1990)	−ve
5	Naimi et al. Canada [36]	342 adults 18–55 years	Cross-sectional 7 census tracts	BMI, HbA1c, TG, TC, HDL—TCR (Montreal Neighbourhood Survey of Lifestyle and Health)	Area-level unemployment (Canada Census 2001)	+ve
6	Cox et al. Scotland [24]	3917 adults < 35 years	Cross-sectional 3382 census output areas (OA)	T2DM (DARTS Diabetes Audit and Research Tayside Scotland dataset)	Area deprivation (The Carstairs score based on 2001 Scotland census data)	+ve
7	Andersen et al. UK [35]	4286 women 60–79 years	Cross-sectional 457 British electoral wards	T2DM, FBG, IR (British Women's Heart and Health Study)	Area deprivation (The Carstairs score based on 2001 census data)	+ve
8	Gabert et al. USA [25]	63,053 DM 18–74 years	Retrospective observational 120 zip code areas	BP, HbA1c, LDL (Minnesota Community Measurement electronic health records)	Area-level indicators of SES (based on American Community Survey 2013)	+ve
9	Dragano et al. GR-Czech [32]	GR: 4814 adults CZ: 8856 adults 57.7 ± 6.6 years	2 longitudinal cohort studies 326 pre-existing administrative units	Obesity, HT (GR: 'Heinz Nixdorf Recall (HNR) Study', Czech: 'Health, Alcohol and Psychosocial Factors in Eastern Europe (HAPIEE) Study')	Area-level socioeconomic status (based on census data)	+ve
10	Cubbin et al. Sweden [31]	18,081 adults 25–64 years	Pooled cross-sectional data 8624 SAMS neighbourhoods	Obesity, DM, HT (Swedish Annual Level of Living Survey (SALLS), 1988–89)	Neighbourhood deprivation (assessed by Care Need Index (CNI) 1997 data)	+ve
11	Mujahid et al. USA [28]	13,167 adults 45–64 years	Crosssectional and longitudinal (3–9 years) Census block	BMI (The Atherosclerosis Risk in Communities ARIC Study)	Neighbourhood SES score (1990 U.S. Census 1990)	−ve
12	Lawlor et al. UK [34]	4286 women 60–79 years	Cross-sectional 457 electoral wards	Coronary heart disease (British Women's Heart and Health Study)	Residential area deprivation (The Carstairs score based on 1991 UK census data)	+ve
13	Roux et al. USA [30]	3093 adults 28–40 years	Cross-sectional 10 years follow up 2260 census block (in 45 states).	BMI, HDL, TG, BP, FI and FG -IRS (Coronary Artery Risk Development in Young Adults CARDIA Study)	Neighbourhood SES score (1990 U.S. Census)	−ve
14	Keita et al. USA [22]	19,079 black/white age > 45 years	Cross-sectional cohort Census block group	Obesity, WC, BP, FBG, TG, low-HDL (REGARDS study).	Neighborhood socioeconomic deprivation (US Census 2000)	+ve (black/white)
15	Clark, et al. USA [23]	3909 Afro-Americans 35–84 years	Cross-sectional cohort 102 census tracts	TG, FBG, BP, WC, low-HDL (Jackson Heart Study).	Neighborhood socioeconomic disadvantage (US Census 2000)	+ve (women)
16	Barber et al. Brazil [21]	10617 adults 35–75 years	Cross sectional cohort Study defined clusters of contiguous census tracts	DM and HT (Brazilian Longitudinal Study of Adult Health)	Area level economic residential segregation (IBGE census 2010)	+ve

BMI body mass index, *BP* blood pressure, *BS* blood sugar, *CHD* coronary heart disease, *CVD* cardiovascular disease, *CVH* cardiovascular health, *DM* diabetes mellitus, *eGFR* estimated Glomerular filtration rate, *FBG* fasting blood glucose, *FG* fasting glucose, *FI* fasting insulin, *GR* Germany, *HbA1c* glycated haemoglobin, *HDL* high density lipoprotein, *HT* hypertension, *IR* insulin resistance, *IRS* insulin resistance syndrome, *LDL* low density lipoprotein, *SES* socioeconomic status, *TC* total cholesterol, *TCR* total cardiometabolic risk, *T2DM* type 2 diabetes mellitus, *TG* triglycerides, *SAMS* small area market statistics

region and sourced from UK, Germany, Czech Republic and France. Reports from North America were mainly from USA (6/7) and Canada (1/7). There was only one study from Oceania, sourced from Australia [26]. Most of these studies were sourced from industrialised nations, except one study from Brazil, [21] a developing nation in South America.

All studies reported associations of higher prevalence of CMRFs with greater ASED [21–36]. Various measures of the biological proxies of CMRFs reported includ biochemical, anthropometric, physiologic, behavioural and diagnostic categories of data. Census sourced data on ASED were used in most of these studies (12/16), whereas other survey sourced data were used in the remaining studies (4/16) to construct summary scores or indices on ASED. The categories of measures used to calculate ASED in these studies were area-level proportions of: median income; education; occupation; housing; transport; dependent population; social class; social capital; environment; security; family structure; disability; internet access; and insurance coverage. A minimum of one category of these measures were used in all the studies [21–36].

The samples characteristics and variables considered were notably heterogeneous across studies. The sampling frame of most (7/16) of these studies were population based lists, however service provider (4/16) and employee (3/16) lists were also used. Two studies had used a combination of both population lists and service provider given lists [27, 30]. Though subjects in all studies used adult age limits ($\geq$18 years), divergent age groups were sampled across all of the studies. Also gender [34, 35] and race [22, 23] specific sampling were used in two studies each. Heterogeneity of these sample characteristics makes a comparison and further quantitative synthesis difficult.

The samples were mostly accessed from existing study cohorts, laboratory databases, national surveys and audit lists. The sample size of studies ranged from 342 adults to a maximum of 91,776 adults, mostly larger in size. Census administration units were the most commonly used neighbourhood proxy, followed by other administrative units and electoral wards. Pre-existing geographic boundaries were mostly adopted to define the spatial unit, but their spatial extents of the unit of analyses were not stated in most of the studies.

Overall synthesis of the studies

Significant features of the included studies were identified to aid synthesis of the findings. These features were the origin of the study, its representativeness, nature of the CMRFs studied, the ecological context and the strength of evidence presented. These features were then formulated into five themes, mapping the related data for further analyses (Table 3).

We had plotted all the included studies to identify their global region of origin and the economic nature of the source country. Most of the studies published were from Europe (11/24), closely followed by America (9/12), (two studies were cross national, hence counted under both the nations and corresponding regions). Fewer publications were found from Oceania (3/24), and Asia (2/24). However no identified studies were from Africa. Studies from developing nations were fewer (3/24) compared with studies from industrialised nations (21/24). This emphasises a gap in related publications from Asia–pacific and African regions, especially from nations of developing and underdeveloped economies. The global representativeness of this review is hence limited, and the review findings may be more generalizable to industrialised nations.

The target populations for included studies are shown in Table 3. The sample frame of most of the studies were population based lists (13/24 studies), however service providers' lists (5/24) and employees lists (3/24) were also used. Both population and service providers' lists were used in three studies (3/24). All the population based studies used a random sampling technique to ensure the population representativeness. However, the response rates varied (15–90.5%) in these studies. Two studies had a response rate $<$ 50%, suggesting a risk of responder bias despite a probability sampling method being employed [29, 36].

Ecological contexts of the included studies were analysed by extracting area level characteristics (Table 3). Area level units used in these studies extended from small areas (10/24), to medium areas (9/24) and large areas (5/24). Small area units were mostly based on census, administrative or zip code area with an average $\sim$ 1000 residing population. Medium area units had an average $\sim$ 5000 population and the large area units were mostly regions, provinces and districts. ASED gradients were based on area level measures of ranged from 1 to 7 measures, however single measures of income or overcrowding as an indirect proxy of ASED raised concerns regarding their comprehensiveness in comparison to aggregate measures of ASED.

The nature of CMRFs and the strength of evidence in relation to associations with outcomes were mapped by extracting data on the categories of CMRFs measured, the source of data and the mode of analyses (Table 3). Biological proxy categories of CMRFs were mostly biochemical (18/24), followed by anthropometric (18/24), physiologic (15/24), and diagnostic (4/24) in nature. Self-reported data on these categories of CMRFs had

Table 3 Thematic mapping of data categories from all included studies

Theme Data map	Study origin Reference Nation (status) Region	Representation Sample frame Sampling Response or retention %	Ecological context Geographic unit and/or ASED	Nature of problem Cardiometabolic risk nature	Evidence Strength Data source Analyses
1.	Roux et al. [30] USA (Industrialised) North America	Population and service providers' lists 79% retention	Small area ASED: income, education, occupation	Biochemical, anthropometric Physiological	Self-report, PE, Specimen tests Statistical
2.	Lawlor et al. [19] UK (Industrialised) Western Europe	Service provider's list Random 60% response	Large area	Biochemical, anthropometric, physiological	Self-report, PE, Specimen tests, MR Statistical
3.	Mujahid et al. [28] USA (Industrialised) North America	Population list1 Random 81% retention	Small area ASED: income, education, occupation	Anthropometric	Self-report, PE Statistical
4.	Lawlor et al. [34] UK (Industrialised) Western Europe	Service provider's list Random 60% response	Median area ASED: employment, housing, transport, social class	Biochemical, anthropometric, physiological	Self-report, PE, Specimen tests, MR Statistical
5.	Cubbin et al. [31] Sweden(Industrialised) Northern Europe	Population list2 Random4 ~80% response	Small area ASED: population structure, education, unemployment etc.	Anthropometric, physiological, diagnostic: DM	Self-report Statistical
6.	Cox et al. [24] Scotland (Industrialised) Western Europe	Service provider's list ~Purposive	Small area ASED: employment, housing, transport, social class	Diagnostic—T2DM	Medical record Spatial and Statistical
7.	Dragano et al. [32] GR-Czech(Industrialised) Western—Central Europe	Population list Random 56 and 55% responses	Small area ASED: unemployment overcrowding	Anthropometric, physiological	Self-report, PE Statistical
8.	Andersen et al. [35] UK (Industrialised) Western Europe	Service provider's lists Random 60% response	Small area ASED: employment, housing, transport, social class	Biochemical	Self-report, PE, MR Statistical
9.	Naimi et al. [36] Canada (Industrialised) North America	Population list Stratified cluster sampling 15% response	Medium area ASED: Education employment	Anthropometric, biochemical	Self-report, PE, Specimen tests Statistical
10.	Barker et al. [20] USA (Industrialised) North America	Population list Random 50.6% response	Medium area	Anthropometric, biochemical	Self-report Statistical
11.	Silhol et al. [33] France (Industrialised) Western Europe	Employees lists ~purposive	Medium area ASED: higher job, education	Anthropometric, biochemical physiological	Self-report, Employers data, Insurance data Spatial and Statistical
12.	Keita et al. [22] USA (Industrialised) North America	Population list	Small area ASED: income, housing education and occupation	Biochemical, anthropometric physiological	Self-report, PE, Specimen tests Statistical
13.	Clark et al. [23] USA (Industrialised) North America 6	Population list Random	Medium area ASED :10 components	Biochemical, anthropometric physiological	PE, Specimen tests Statistical
14.	Valdes et al. [18] Spain (Industrialised) Southern Europe	Population list Cluster –random 54.6% response	Large area	Anthropometric, hysiological, biochemical	Self-report, PE, Specimen tests Statistical
15.	Astell-Burt et al. [14] Australia (Industrialised) Oceania	Population (insurance) lists Random	Medium area	Biochemical, physiological	Self-report Spatial and Statistical
16.	Unger et al. [27] USA (Industrialised) North America	Population and service provider's list (~purposive)	Medium area ASED: income, housing, education, occupation.	Anthropometric biochemical physiological	Self-report, PE, Specimen tests Statistical

Table 3 (continued)

Theme Data map	Study origin Reference Nation (status) Region	Representation Sample frame Sampling Response or retention %	Ecological context Geographic unit and/or ASED	Nature of problem Cardiometabolic risk nature	Evidence Strength Data source Analyses
17.	Maier et al. [29] Germany(Industrialised) West - Central Europe	Population list Random 29.1% response	Large area. ASED: income, employment, education, revenue, social capital, environment, security	Anthropometric Diagnostic—T2DM	Self-report Statistical
18.	Zhou et al. [17] China (Developing) East Asia	Population (survey) list Random 90.5% response	Large area	Anthropometric biochemical	Self-report, Specimen tests Statistical
19.	Bonney et al. [26] Australia (Industrialised) Oceania	Service provider's list (~purposive)	Small area. ASED: income, education, employment, family structure, disability, housing, transport and internet connection	Anthropometric	Medical record Statistical
20.	Gabert et al. [25] USA (Industrialised) North America 8	Employees list 83.6% response	Small area ASED: income, education, insurance	Biochemical	Medical record Spatial and Statistical
21.	Paquet et al. [16] AU-France (Industrialised) Oceania - West Europe	Australia: Population list (Random 13/49.4% response 3) France: Employees list (Purposive/83.6% response)	Small area	Anthropometric, biochemical, hysiological	PE, Specimen tests Spatial and Statistical
22.	Alkerwi et al. [15] Luxemberg (Industrialised) Western Europe	Population (survey) list stratified random 32.2% response	Medium area	Physiological, biochemical	Self-report, PE, Specimen tests Spatial and Statistical
23.	Oh et al. [13] South Korea (Developing) East Asia	Population (ministry) lists ~purposive	Medium area	Biochemical, physiological, diagnostic	Self-report Spatial and Statistical
24.	Barber et al. [21] Brazil (Developing) South America	Employees lists ~purposive	Large area ASED: income	Biochemical, anthropometric, physiological	Self-report, PE, Specimen tests Spatial and Statistical

the highest risk for misclassification due to reporting bias or errors. Studies which adopted a combined mode of both statistical and spatial analyses provided a better ecological context of CMRFs than with statistical analyses alone.

Discussion

ASED was repeatedly demonstrated to be associated with higher cardiometabolic risk. Higher ASED was consistently reported to have an association with cardiovascular risk; whereas lower ASED was associated with reduced cardiovascular risk. Such associations were often demonstrated independently of individual level characteristics such as socioeconomic status, education and duration of exposure to area. Type 2 diabetes and high body mass index (BMI) were reported to be more prevalent in disadvantaged areas. Related studies report that the type of neighbourhood food outlets [37–39], poor physical activity resources [39], individual perception of area level features [40] residential density and service availability [41] were all explanatory variables associated with cardiometabolic risk prevalence among people living in disadvantaged neighbourhoods.

Related systematic reviews published in this area of research investigate associations for different geographically distributed factors with CVD. Chaix (2009) reviewed the associations between neighbourhood social environments and CHD, and proposed a theoretical model of a mediating mechanism focussing on the social interactional environment [10]. Consistent associations of obesity or hypertension with lower levels of area socioeconomic status, urbanization, street intersection, accessibility to supermarkets, social cohesion, service availability and residential density; and higher levels of

noise pollution and density of convenience stores, were reviewed and reported by Leal [11]. Frequent inverse associations of the common indices of ASED with childhood obesity were reported in the UK [9]. Consistent associations between socio economic disadvantage and central adiposity was reported by Slopen [12]. All these reviews report important methodological inadequacies and the need for further research in this area, which support the findings of the current review.

Recent advances in geographic information systems (GIS) and analytical approaches were utilised in the studies reporting geographic variation in CMRFs. These studies have demonstrated advances in various analytical tools and the potential for plotting area level risk parameters. Geocoding and mapping of existing large population based datasets has become feasible with newer computational tools through linking location data; such as map co-ordinates, addresses or postcodes [42]. These tools have the capacity to visually display area based factors, in contrast with traditional table and graph methods, and this has the potential to enhance impact on subsequent area level health care policy development and resource allocation [43–45]. In addition, systematic quantitative analyses are possible with these spatial tools which create opportunities to investigate the role of environmental factors in explaining any geographic aggregations beyond random effects [46].

National estimates of CVD have limited utility in informing prevention and management of CVD within discrete communities. The disease patterns at smaller areas may significantly differ from national and regional prevalence reports, thus small area analysis is important in order to understand local patterns and requirements [47]. Small-area level analyses also have the potential to reveal area level contexts and dependencies of CMRFs and such analyses can highlight areas for targeted preventive interventions.

CVD and its associated CMRFs continue to evolve as a major global health threat. It is the highest cause of mortality and the highest absorber of health care expenditure in many developed nations [6, 48, 49]. Once diagnosed, the ongoing costs of care and productivity loss due to consequent disability and premature death creates a large economic burden not only to the individual and family, but to the nation—especially when half the people dying are found to be in their prime productive years [50]. Thus, CVD and its associated metabolic risk factors emerge as a threat not only to human health and life, but to the sustainable development and economies of nations. Hence, improving public health program effectiveness in reducing CVD must be a research priority.

Limitations

Firstly, the cross sectional nature of the reviewed studies precluded causative interpretations. Second, the global representativeness of the review is limited mainly due to publication gaps from Asia–pacific and African regions of the World. Third, the scope of our review excluded examination of behavioural, dietary and activity related risk factors and also other area level characteristics to focus only on the biological proxies of CMRFs. Fourth, methodological heterogeneity within the retrieved studies prohibited a meta-analytical synthesis of the findings. The sample characteristics, geographical scales and the CMRFs' risk profiles varied substantially across the studies impeding any further quantitative synthesis.

Recommendations and future directions

Finding geographic variation in CMRFs (if any) and its association with ASED may assist in understanding the contexts of risk. Such studies have the potential to inform contextual planning of interventions for prevention and management of cardiometabolic risk. However, most of the studies in this review do not report the spatial extents of their units of analysis. This is important as associations are likely to be different at different levels of aggregation, and limits the ability to assess the likelihood of spatial scale effects in these studies [22, 23] known as the Modifiable Areal Unit Problem [51, 52]. When data are aggregated to larger geographic units, small-area anomalies may be diluted or smoothed over [22]. Using smaller rather than larger area scales can help to reduce the likelihood of missing important small area anomalies [53]. Similarly, supplementing individual level data along with area level data could minimise group effects due to area level aggregation of data [53]. Leveraging both individual- and area-level data provides a more complete picture to inform planning, policy and practice [46, 53]. Future research directions should include hierarchical multilevel analyses to yield comprehensive picture of the contextual aspects of risk factors, to help aid both individual and area-level better preventive initiatives.

Conclusion

Cardiometabolic risk distribution varied significantly across different geographic scales reported in multiple studies. In addition, there is strong evidence that area-level disadvantage is significantly associated with CMRFs, irrespective of individual-level characteristics. This review highlights the need for area-based preventive approaches in addition to individual-level approaches to prevent and control CMRFs and their consequent CVD outcomes.

Additional files

Additional file 1. The search strategy and results URLs.

Additional file 2. List of excluded full text studies with reason.

Additional file 3. PRISMA 2009 systematic review content checklist.

Abbreviations

AU: Australia; ASED: area-level socioeconomic disadvantage; BP: blood pressure; BMI: body mass index; BS: blood sugar; CD: census collection district; CHD: coronary heart disease; CMRFs: cardiometabolic risk factors; CKD: chronic kidney disease; CVD: cardiovascular disease; CVH: cardiovascular health; DM: diabetes mellitus; eGFR: estimated Glomerular filtration rate; FBG: fasting blood glucose; FG: fasting glucose; FI: fasting insulin; FPG: fasting plasma glucose; GFR: glomerular filtration rate; GR: Germany; HbA1c: glycated haemoglobin; HDL: high density lipoprotein; HR: heart rate; HT: hypertension; IR: insulin resistance; IRS: insulin resistance syndrome; IRIS: ilôts regroupés pour l'information statistique; LDL: low density lipoprotein; LGA: local government area; POA: postal area; PRISMA: preferred reporting items for systematic reviews and meta-analyses; SES: socioeconomic status; SLA: statistical local area; TC: total cholesterol; TCR: total cardiometabolic risk; T2DM: type 2 diabetes mellitus; TG: triglycerides; TRIRIS: groups of around three IRIS areas; WC: waist circumference.

Authors' contributions

Review conception and design: RT, AB, DM, and XF. Search strategy and literature search: RT. Study coding, selection and data extraction: RT, RW. Review and interpretation: RT, AB, DM. Drafting of manuscript: RT, AB, DM. Critical revision: RT, AB, DM. Final revision: RT, AB, DM, and XF

Author details

[1] School of Medicine, University of Wollongong, Wollongong, Australia. [2] Illawarra Health and Medical Research Institute, Wollongong, Australia. [3] Public Health Unit, Illawarra Shoalhaven Local Health District, Warrawong, Australia. [4] Population Wellbeing and Environment Research Lab (PowerLab), School of Health and Society, Faculty of Social Sciences, University of Wollongong, Wollongong, Australia.

Acknowledgements

This review has been conducted with the support of the Australian Government Research Training Program Scholarship.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Not applicable.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

Not applicable.

Publisher's Note

Received: 4 September 2018 Accepted: 31 December 2018
Published online: 08 January 2019

References

1. WHO. Cardiovascular diseases (CVDs): Key facts [Internet]. 2016. http://www.who.int/mediacentre/factsheets/fs317/en/.
2. Schutzer SE, Fraser-Liggett CM, Casjens SR, Qiu WG, Dunn JJ, Mongodin EF, et al. Whole-genome sequences of thirteen isolates of *Borrelia burgdorferi*. J Bacteriol. 2011;193(4):1018–20.
3. Murray CJL, Lopez AD. Measuring the global burden of disease. N Engl J Med. 2013;369(5):448–57.
4. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, et al. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. Circulation. 2017;135(10):e146–603.
5. World Health Organisation. WHO | The top 10 causes of death [Internet]. World Health Organization; 2017 http://www.who.int/mediacentre/factsheets/fs310/en/. Accessed 2018 Mar 10.
6. Cannon CP. Cardiovascular disease and modifiable cardiometabolic risk factors. Clin Cornerstone. 2008;9(2):24–41.
7. Brunzell JD, Davidson M, Furberg CD, Goldberg RB, Howard BV, Stein JH, et al. Lipoprotein management in patients with cardiometabolic risk: consensus conference report from the American Diabetes Association and the American College of Cardiology Foundation. J Am Coll Cardiol. 2008;51(15):1512–24.
8. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet. 2014;383(9921):999–1008.
9. El-Sayed AM, Scarborough P, Galea S. Socioeconomic inequalities in childhood obesity in the United Kingdom: a systematic review of the literature. Obes Facts. 2012;5:671–92.
10. Chaix B. Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda. Annu Rev Public Health. 2009;30(1):81–105. https://doi.org/10.1146/annurev.publhealth.031308.100158.
11. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. Obes Rev. 2011;12(3):217–30.
12. Slopen N, Goodman E, Koenen KC, Kubzansky LD. Socioeconomic and other social stressors and biomarkers of cardiometabolic risk in youth: a systematic review of less studied risk factors. PLoS ONE. 2013;8(5):e64418.
13. Oh WS, Yoon S, Noh J, Sohn J, Kim C, Heo J. Geographical variations and influential factors in prevalence of cardiometabolic diseases in South Korea. PLoS ONE. 2018;13(10):e0205005.
14. Astell-Burt T, Feng X, Kolt GS, McLean M, Maberly G. Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia. Diabetes Res Clin Pract. 2014;106(3):e68–73.
15. Alkerwi A, Bahi IE, Stranges S, Beissel J, Delagardelle C, Noppe S, et al. Geographic variations in cardiometabolic risk factors in luxembourg. Int J Environ Res Public Health. 2017;14(6):16.
16. Paquet C, Chaix B, Howard NJ, Coffee NT, Adams RJ, Taylor AW, et al. Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia. Int J Environ Res Public Health. 2016;13(5):21.
17. Zhou M, Astell-Burt T, Bi Y, Feng X, Jiang Y, Li Y, et al. Geographical variation in diabetes prevalence and detection in china: multilevel spatial analysis of 98,058 adults. Diabetes Care. 2015;38(1):72–81.
18. Valdes S, Garcia-Torres F, Maldonado-Araque C, Goday A, Calle-Pascual A, Soriguer F, et al. Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (Southern Spain). Comparison with national prevalence data. The Diabetes study. Rev Esp Cardiol. 2014;67(6):442–8.
19. Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. J Epidemiol Commun Health. 2003;57(2):134–40.
20. Barker LE, Kirtland KA, Gregg EW, Geiss LS, Thompson TJ. Geographic distribution of diagnosed diabetes in the U.S.: a diabetes belt. Am J Prev Med. 2011;40(4):434–9.
21. Barber S, Diez Roux AV, Cardoso L, Santos S, Toste V, James S, et al. At the intersection of place, race, and health in Brazil: residential segregation and cardio-metabolic risk factors in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Soc Sci Med. 2018;199:67–76.
22. Keita AD, Judd SE, Howard VJ, Carson AP, Ard JD, Fernandez JR. Associations of neighborhood area level deprivation with the metabolic syndrome and inflammation among middle- and older-age adults. BMC Public Health. 2014;14:1319.

23. Clark CR, Ommerborn MJ, Hickson DMA, Grooms KN, Sims M, Taylor HA, et al. Neighborhood disadvantage, neighborhood safety and cardiometabolic risk factors in African Americans: biosocial associations in the Jackson Heart Study. PLoS ONE. 2013;8(5):e63254.
24. Cox M, Boyle PJ, Davey PG, Feng Z, Morris AD. Locality deprivation and Type 2 diabetes incidence: a local test of relative inequalities. Soc Sci Med. 2007;65(9):1953–64.
25. Gabert R, Thomson B, Gakidou E, Roth G. Identifying high-risk neighborhoods using electronic medical records: a population-based approach for targeting diabetes prevention and treatment interventions. PLoS ONE. 2016;11(7):e0159227.
26. Bonney A, Mayne DJ, Jones BD, Bott L, Andersen SE, Caputi P, et al. Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users—a cross-sectional study. PLoS ONE. 2015;10(8):e0137261.
27. Unger E, Diez-Roux AV, Lloyd-Jones DM, Mujahid MS, Nettleton JA, Bertoni A, et al. Association of neighborhood characteristics with cardiovascular health in the multi-ethnic study of atherosclerosis. Circ Cardiovasc Qual Outcomes. 2014;7(4):524–31.
28. Mujahid MS, Diez Roux AV, Borrell LN, Nieto FJ. Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics. Obes Res. 2005;13(8):1412–21.
29. Maier W, Scheidt-Nave C, Holle R, Kroll LE, Lampert T, Du Y, et al. Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys "German Health Update" GEDA 2009 and 2010. PLoS ONE. 2014;9(2):e89661.
30. Roux AVD, Jacobs DR, Kiefe CI. Neighborhood characteristics and components of the insulin resistance syndrome in young adults the coronary artery risk development in young adults (CARDIA) study. Diabetes Care. 2002;25(11):1976–82.
31. Cubbin C, Sundquist K, Ahlén H, Johansson S-E, Winkleby MA, Sundquist J. Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects. Scand J Public Health. 2006;34(3):228–37.
32. Dragano N, Bobak M, Wege N, Peasey A, Verde PE, Kubinova R, et al. Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany. BMC Public Health. 2007;7(1):1.
33. Silhol R, Zins M, Chauvin P, Chaix B. Investigating the spatial variability in incidence of coronary heart disease in the Gazel cohort: the impact of area socioeconomic position and mediating role of risk factors. J Epidemiol Commun Health. 2011;65(2):137–43.
34. Lawlor DA, Davey Smith G, Patel R, Ebrahim S. Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study. Am J Public Health. 2005;95(1):91–7.
35. Andersen AF, Carson C, Watt HC, Lawlor DA, Avlund K, Ebrahim S. Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study. Diabet Med. 2008;25(12):1462–8.
36. Naimi AI, Paquet C, Gauvin L, Daniel M. Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area. Int J Environ Res Public Health. 2009;6(12):3082–96.
37. Astell-Burt T, Feng X. Geographic inequity in healthy food environment and type 2 diabetes: Can we please turn off the tap? Med J Aust. 2015;203:246–8.
38. Millstein RA, Yeh H-C, Brancati FL, Batts-Turner M, Gary TL. Food availability, neighborhood socioeconomic status, and dietary patterns among blacks with type 2 diabetes mellitus. Medscape J Med. 2009;11(1):15.
39. Christine PJ, Auchincloss AH, Bertoni AG, et al. Longitudinal associations between neighborhood physical and social environments and incident type 2 diabetes mellitus: the multi-ethnic study of atherosclerosis (mesa). JAMA Intern Med. 2015;175(8):1311–20. https://doi.org/10.1001/jamainternmed.2015.2691.
40. Baldock K, Paquet C, Howard N, Coffee N, Hugo G, Taylor A, et al. Associations between resident perceptions of the local residential environment and metabolic syndrome. J Environ Public Health. 2012;2012:589409.
41. Chaix B. Geographic life environments and coronary heart disease: a literature review, theoretical contributions, methodological updates, and a research agenda. Annu Rev Public Health. 2009;30:81–105.
42. Stevens CD, Schriger DL, Raffetto B, Davis AC, Zingmond D, Roby DH. Geographic clustering of diabetic lower-extremity amputations in low-income regions of California. Health Aff. 2014;33(8):1383–90.
43. Angier H, Likumahuwa S, Finnegan S, Vakarcs T, Nelson C, Bazemore A, et al. Using geographic information systems (GIS) to identify communities in need of health insurance outreach: an OCHIN practice-based research network (PBRN) report. J Am Board Fam Med. 2014;27(6):804–10. https://doi.org/10.3122/jabfm.2014.06.140029.
44. Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. A review of spatial methods in epidemiology. Ann Rev Public Health. 2012;33:107–22.
45. Bazemore A, Phillips RL, Miyoshi T. Harnessing geographic information systems (GIS) to enable community-oriented primary care. J Am Board Fam Med. 2010;23(1):22–31. https://doi.org/10.3122/jabfm.2010.01.090097.
46. Elliott P, Wartenberg D. Spatial epidemiology: current approaches and future challenges. Environ Health Perspect. 2004;112:998–1006.
47. Occelli F, Deram A, Genin M, Noel C, Cuny D, Glowacki F, et al. Mapping end-stage renal disease (ESRD): spatial variations on small area level in northern France, and association with deprivation. PLoS ONE. 2014;9(11):e110132.
48. World Health Organization; World Heart Federation and World Stroke Organization. Global Atlas on cardiovascular disease prevention and control. Glob Atlas Cardiovasc Dis Prev Control. 2011;155. https://www.cabdirect.org/cabdirect/abstract/20123402600%0Afile:///C:/Users/USER/Downloads/9789241564373_eng(2).pdf.
49. World Health Organization. WHO | Noncommunicable diseases. WHO. WHO; 2017. http://www.who.int/mediacentre/factsheets/fs355/en/. Accessed 2018 Mar 10.
50. Bloom DE, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Reddy Bloom L, Fathima S, et al. The global economic burden of noncommunicable diseases. World Econ Forum. 2011;1–46. http://ideas.repec.org/p/gdm/wpaper/8712.html.
51. Openshaw S. Modifiable areal unit problem. Concepts Tech Mod Geogr. 1989;38:169–74.
52. Openshaw S, Taylor PJ. A million or so correlation coefficients, three experiments on the modifiable areal unit problem. In: Wrigley N, editor. Statistical applications in the spatial science. London: Pion; 1979. p. 127–44. https://trove.nla.gov.au/work/10094088
53. Wakefield, Jonathan HL. Spatial aggregation and the ecological fallacy. Handbook of spatial statistics. 2010; pp. 541–558. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4209486/.

Supplementary Material: 2 (Published version of research article 2)

PLOS ONE

RESEARCH ARTICLE

Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia

Renin Toms[1,2]*, **Darren J. Mayne**[1,2,3,4], **Xiaoqi Feng**[2,5,6], **Andrew Bonney**[1,2]

1 School of Medicine, University of Wollongong, Wollongong, NSW, Australia, **2** Illawarra Health and Medical Research Institute, Wollongong, NSW, Australia, **3** Public Health Unit, Illawarra Shoalhaven Local Health District, Warrawong, NSW, Australia, **4** School of Public Health, The University of Sydney, Sydney, NSW, Australia, **5** School of Public Health and Community Medicine, University of New South Wales, Sydney, NSW, Australia, **6** Population Wellbeing and Environment Research Lab (PowerLab), School of Health and Society, University of Wollongong, Wollongong, NSW, Australia

* rmbst288@uowmail.edu.au

OPEN ACCESS

Citation: Toms R, Mayne DJ, Feng X, Bonney A (2019) Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia. PLoS ONE 14(10): e0223179. https://doi.org/10.1371/journal.pone.0223179

Editor: Wenhao Yu, China University of Geosciences, CHINA

Received: May 26, 2019

Accepted: September 16, 2019

Published: October 1, 2019

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0223179

Data Availability Statement: Access to, and use of, Southern IML Research (SIMLR) Study data are subject to a License Agreement — Provision of

Abstract

Introduction

Metabolic risk factors for cardiovascular disease (CVD) warrant significant public health concern globally. This study aims to utilise the regional database of a major laboratory network to describe the geographic distribution pattern of eight different cardiometabolic risk factors (CMRFs), which in turn can potentially generate hypotheses for future research into locality specific preventive approaches.

Method

A cross-sectional design utilising de-identified laboratory data on eight CMRFs including fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI); and diabetes mellitus (DM) status was used to undertake descriptive and spatial analyses. CMRF test results were dichotomised into 'higher risk' and 'lower risk' values based on existing risk definitions. Australian Census Statistical Area Level 1 (SA1) were used as the geographic units of analysis, and an Empirical Bayes (EB) approach was used to smooth rates at SA1 level. Choropleth maps demonstrating the distribution of CMRFs rates at SA1 level were produced. Spatial clustering of CMRFs was assessed using Global Moran's I test and Local Indicators of Spatial Autocorrelation (LISA).

Results

A total of 1,132,016 test data derived from 256,525 individuals revealed significant geographic variation in the distribution of 'higher risk' CMRF findings. The populated eastern seaboard of the study region demonstrated the highest rates of CMRFs. Global Moran's I

Data (LA) between Southern IML Pathology Pty Ltd (Data Owner) and The University of Wollongong (License Holder), and a Data Access Agreement (DAA) between the License Holder and researchers (Data Users). This process is facilitated by the Illawarra Health and Medical Research Institute (IHMRI) (Data Custodian) through the Southern IML Research Study — Cohort Management Committee (SIMLR—CMC). The Data License does not allow for "public access" to data; however, researcher may access to SIMLR Study data subject to approval by the SIMLR—CMC and an appropriately constituted Australian Human Research Ethics Committee (HREC) as defined in the National Health and Medical Research Council's National Statement on Ethical Conduct in Human Research (2007) (available from https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018). The Data License requires at least one of the research team be affiliated with IHMRI. SIMLR—CMC contact details are: C/o- Associate Professor Kathryn Weston; Southern IML Research Study — Cohort Management Committee; Illawarra Health and Medical Research Institute; Building 32, University of Wollongong, Northfields Avenue, Wollongong NSW 2522, Australia; Phone +61 2 4221 4333; Email: info@ihmri.org.au; Web Link: https://www.ihmri.org.au/research-projects/simlr-cohort-study/.

Funding: The Australian Government Research Training Program Scholarship covers the university tuition fees of the PhD candidate who is the principal author in this study (RT).

Competing interests: The authors have declared that no competing interests exist.

values were significant and positive at SA1 level for all CMRFs. The highest spatial autocorrelation strength was found among obesity rates (0.328), and the lowest for albuminuria (0.028). LISA tests identified significant High-High (HH) and Low-Low (LL) spatial clusters of CMRFs, with LL predominantly in the less populated northern, central and southern regions of the study area.

Conclusion

The study describes a range of CMRFs with different distributions in the study region. The results allow generation of hypotheses to test in future research concerning location specific population health approaches.

Introduction

Uncontrolled cardiometabolic risk factors (CMRFs) such as hyperglycaemia, dyslipidaemia, albuminuria, inadequate glomerular filtration, overweight and/or obesity and diabetes can predispose and heighten the risk for cardiovascular disease (CVD).[1–6] Cardiovascular diseases are the leading cause of death worldwide, and the highest absorber of health care expenditure in many developed nations, including Australia.[7–9]

In Australia, CVD remain the single leading cause of death; the largest health problem; and a major economic burden.[10,11] Nine in 10 adult Australians have at least one CVD risk factor, and one in four have three or more risk factors.[11] CVD kills one Australian every 12 minutes and one in six Australians (3.7 million people) are thought to be at risk.[12] In addition, the prevalence of CVD is projected to steeply increase in the coming decades.[11] A deceleration in the rapid growth of this major health care issue is possible only through the prevention and control of CMRFs. The role of CMRFs in the population, over and above individual level factors such as age, are being questioned in regard to discriminatory accuracy for development of CVD.[13] However identification of one or more CMRFs in a person at any age can initiate preventive lifestyle changes which may have significant benefits.[14–18] Similarly, identification of areas with higher rates of CMRFs can potentially trigger further area-level analyses investigating the potential for targeted health service commissioning.[19–21]

Advances in Geographic Information System (GIS) over the last quarter of a century have provided various tools to integrate epidemiological and geographical data.[22–24] Geocoding of risk parameters became feasible with such tools for its area-level analyses, which has facilitated area-level mapping of risk parameters, which has the potential generate hypothesis for regional health care research.[22] Thus integrating risk parameters through GIS has the potential to facilitate area-level health research, [25–28]; however, not without potential pitfalls [29–31]. A limitation of GIS-based mapping is that its outputs may be misleading, especially if maps are not smoothed using appropriate spatial or multilevel analyses.[32–34] However, it is well recognised in the literature that area level community interventions based on GIS approaches have been successful in a number of countries. [19–21,35,36]

There has been a significant increase in the number of epidemiological studies using spatial analytical methods in the last decade, including international studies reporting significant geographic variation in CMRFs at different spatial scales of measurements.[37–45] Hyperglycaemia was the most commonly reported CMRF displaying variation, followed by dyslipidaemia, overweight and/or obesity, and inadequate glomerular filtration.[37] Multiple risk factors were rarely analysed in these studies, though most CMRFs are interrelated and often coexist.[46] In

this study, we aim to demonstrate the feasibility of utilising laboratory based routine test data to generate basic distribution maps of eight different CMRFs in regional New South Wales (NSW), Australia. The research questions we address are: (1) what is the geographic distribution pattern of CMRFs in the study area; and (2) is there any significant spatial clustering of CMRFs rates? The research sought to identify area-level patterns in the distribution of CMRFs that could be used to generate hypotheses for future research with the goal of improving health service commissioning in the study region.

Methods

The study adopted a cross-sectional design and was approved by the University of Wollongong (UOW) and Illawarra and Shoalhaven Local Health District (ISLHD) Human Research Ethics Committee (HREC 2017/124).

Setting

The study was undertaken in the Illawarra-Shoalhaven region (ISR) of the NSW, Australia. The ISR region stretches from the immediate south of the metropolitan boarders of Sydney, and extends along the south-eastern coastal belt of NSW—bordered by the Pacific Ocean in the east and the coastal escarpment of the Southern Tablelands in the West. This region encompasses multiple cities, towns and rural areas and includes the four local government areas of Wollongong, Shellharbour, Kiama and Shoalhaven. Overall, the ISR covers a land area of 5615 square kilometres and had an estimated residential population of 369,469 persons at the 2011 Australian Census of Population and Housing, of which 285, 385 (77.24%) were adults (> = 18 years).[47] De-identified data for this study were obtained from the Southern IML Research (SIMLR) Study, a large-scale community-derived cohort of internally-linked and geographically referenced pathology data collected in routine practice by the largest pathology provider servicing the study area. More details on this data source, its access and maintenance are published elsewhere.[48]

Statistical Area level 1 (SA1) was used as the geographic unit of analysis in this study, which was the smallest geographic unit for the release of Census data in 2011.[49] SA1s generally have a population of 200 to 800 persons (400 average), and the ISR includes a total of 980 conterminous SA1s. Fig 1 shows the study area with SA1 units and the major landmarks of the region. Very small and crowded SA1s similar to the areas shown the inset map tend to be more densely populated.

Participants and variables

The CMRF test data of the adult residents of ISR between 1 Jan 2012–31 Dec 2017 (6 years) were extracted for analyses from the SIMLR database. Test data were extracted for eight CMRFs: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); body mass index (BMI) and diabetes mellitus (DM) status. The SIMLR database uses an algorithm to identify DM status based on diagnosis guidelines published by the Royal Australian College of General Practitioners (RACGP) and Diabetes Australia, and methods from the National Health Survey of the Australian Bureau of Statistics (ABS).[50,51] The algorithm identifies DM for HbA1c $\geq$ 6.5% or FBSL $\geq$ 7.0 mmol/l within +/- 24 months of HbA1c < 6.5%. The study data included both prevalent and incident DM cases. Study data included only the most recent CMRF test result for each individual. We excluded extreme BMI values <12 and >80 based on cut-off points reported by Cheng (2016), Li (2009) and Littman (2012).[52–54] Table 1 lists the CMRFs value definitions adopted in this study and their source references.

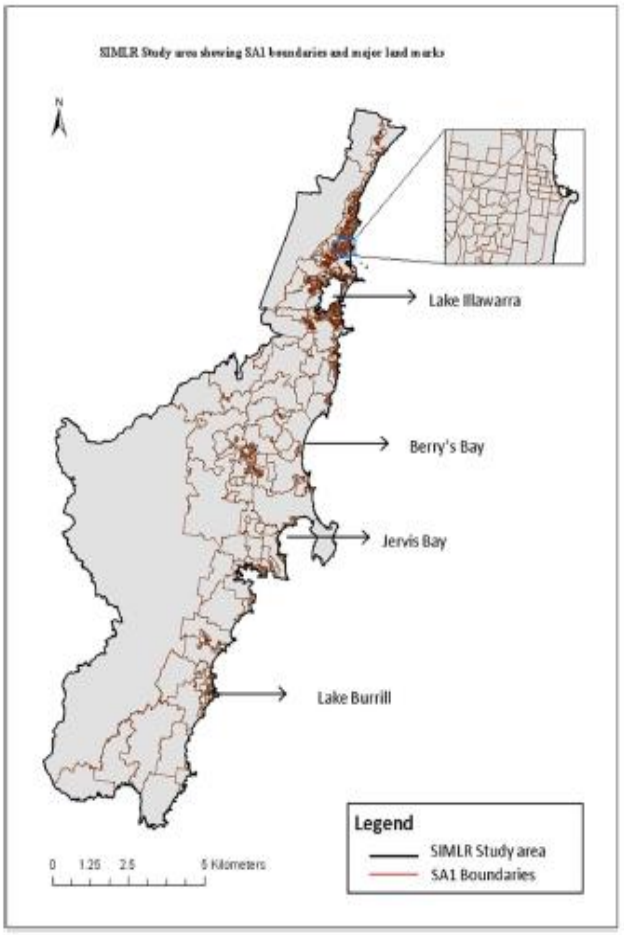

Fig 1. Map of the Illawarra-Shoalhaven region of NSW Australia showing SA1 areas and major landmarks.

https://doi.org/10.1371/journal.pone.0223179.g001

Statistical and spatial analyses

First, individual-level descriptive analyses of CMRFs were performed. The total number of each CMRF tests and summary statistics of each tests' results are reported. The summary

Table 1. Cardiometabolic risk classification.

'Higher risk' CMRFS	Value definition	Adopted from
High FBSL	FBSL ≥7.0 mmol/l	RACGP guidelines[50]
High HbA1c	HbA1c > 7.5%	RACGP guidelines[50]
High TC	TC ≥ 5.5 mmol/l	Australian Health Survey[55]
Low HDL	HDL < 1 mmol/l[56]	National heart foundation of Australia[57]
High ACR	ACR ≥ 30 mcg/L to mg/l	Kidney Health Australia[58]
Low eGFR	eGFR < 60 mL/min/1.73m2	Kidney Health Australia[58]
High BMI	BMI ≥ 30 (Obese)	World Health Organization (WHO)[59]
DM Status	+ve DM test algorithm	RACGP guidelines[50] and Australian Health Survey[55]

https://doi.org/10.1371/journal.pone.0223179.t001

values for eGFR test results are calculated using the approach for grouped data as eGFR test result values are truncated at >90 in the SIMLR Study data. Test results were dichotomised into 'higher risk' and 'lower risk' categories based on the CMRF definitions in Table 1.

Second, area-level analyses of CMRFs were undertaken. Within-cohort prevalence of 'higher risk' CMRF findings are calculated using the total number of tests within each SA1 as the denominator. The exception were DM cases, which are likely to include most prevalent cases in the study area, so SA1 adult populations aged 18 years and over were used as the denominators (accessed from ABS census 2011 data). Thereafter, an Empirical Bayes (EB) approach was used to smooth all the CMRFs' raw rates to minimise extreme values arising from small sample sizes. The EB smoothed rates were then imported into GIS software for mapping and spatial statistical analyses.

As individuals with CMRFs are assumed randomly distributed within the study area, the geographic distribution of CMRFs is assumed spatially independent in this study. Global Moran's I test was used to identify spatial autocorrelation of CMRFs at a 0.05 level of significance. Global Moran's I tests if the geographic distribution of rates is clustered, dispersed or random based.[60] The global Moran's I also indicates the general strength of spatial autocorrelation in the study area, which theoretically ranges between -1 to +1. Values of I significantly above -1/(N-1) indicate positive spatial autocorrelation, where N is the number of spatial units indexed.[61] When significant spatial autocorrelation was detected, Local Indicator of Spatial Autocorrelation (LISA) spatial statistics were used to identify any clustering of CMRFs.[62] LISA was used to indicate spatial clustering of High-High (HH) or Low-Low (LL) CMRFs rates at SA1-level within the study region. False Discovery Rate (FDR) corrections were applied to LISA tests to correct p-values for multiple testing.

All descriptive statistics and EB smoothing were performed using R version 3.4.4.(R Foundation for Statistical Computing, Vienna, Austria).[63] Mapping and spatial analyses were performed using ArcGIS version 10.4.1(ESRI Inc. Redlands, CA, USA).[64]

Results

The study sample comprised 1,132,016 test results contributed by 256,525 adult individuals residing in the study region. Of the 256,525 individuals, 193,679 (75.5%) had FBSL, 73,885 (28.8%) had HbA1, 194,816 (75.9%) had TC, 182,237 had HDL (71.0%), 50,790 had ACR (19.8%), 244,166 had eGFR (95.2%), and 192,443 had BMI (75.0%) test results. It was estimated 23,704 (9.2%) of persons met the clinical criteria for diabetes. Table 2 provides the summary statistics of CMRF test results.

The CMRF test result values were dichotomised into 'higher risk' and 'lower risk' categories based on the CMRF definitions in Table 1. The proportion of individuals with 'higher risk' CMRFs findings varied considerably between tests. The largest 'higher risk' proportions were

Table 2. Summary statistics of CMRFs test results.

CMRFs	Tests	Mean	SD	Min	1st Qu	Median	3rd Qu	Max
FBSL	193679	5.6	1.6	0.7	4.9	5.3	5.8	43.9
HbA1c	73885	6.0	1.3	2.6	5.3	5.6	6.4	17.8
TC	194816	5.0	1.1	1.1	4.2	4.9	5.7	39.4
HDL	182237	1.5	1.2	0.1	0.5	1.4	1.8	5.8
ACR	50790	7.4	40.3	0.1	0.4	0.8	2.3	1291.5
eGFR	244166	75.8	13.8	2.0	-	83.2	-	>90.0
BMI	192443	28.4	6.1	12.0	24.1	27.5	31.6	78.1

https://doi.org/10.1371/journal.pone.0223179.t002

found for BMI (33.74%) and TC (32.55%), and the lowest for ACR (4.03%). Table 3 provides details on the CMRF test results classification and the identified proportions.

Geographic distribution of cardiometabolic risk factors

Fig 2 shows the geographic distribution of CMRFs at SA1 level in the ISR region with red indicating the highest and blue the lowest rates of risk. SA1s with no test data appear in white. Areas with higher rates of CMRFs were found to be clustering within the study region. The highest rates were found mainly along the populated eastern board of the study region; notably among SA1s around Lake Illawarra, south-east of Berry's bay, and east of Lake Burill. However, the high TC rates showed a reversed pattern, and higher rates were found in the relatively less populated central and westerly aspects of the study area. HDL rates did not follow this reversed pattern.

Spatial autocorrelation of CMRFs

The global Moran's I tests were significant and positive for all CMRFs (Table 4). The highest spatial autocorrelation strength was found among obesity rates (0.328), followed by high FBSL (0.184) and low HDL (0.174). The spatial autocorrelation strength was the lowest for albuminuria (0.028) and low eGFR (0.069).

LISA tests identified significant spatial clustering of CMRFs in the ISR region. The HH clusters were found mainly along the populated areas of the study region, except for TC. Areas around the immediate surroundings of Lake Illawarra had the most HH clusters, followed by the areas to the south-west of Berry's Bay and south of Jervis Bay. A few areas around Lake

Table 3. Frequency and proportion of 'higher risk' results of CMRFs tests.

Cardiometabolic risk	Classification	Tests n (%)*
FBSL		193679 (100)
FBSL ≥7.0 mmol/L	Higher risk	16280(8.4)
FBG < 7.0 mmol/L	Lower risk	177399(91.6)
HbA1c		73885(100)
HbA1c > 7.5%	Higher risk	7927(10.7)
HbA1c ≤ 7.5%	Lower risk	65958(89.3)
TC		194816(100)
TC ≥ 5.5 mmol/L	Higher risk	63422(32.5)
TC < 5.5 mmol/L	Lower risk	131394(67.5)
HDL		182237 (100)
HDL < 1 mmol/l	Higher risk	21261(11.7)
HDL ≥ 1 mmol/l	Lower risk	160976(88.3)
ACR		50790(100)
ACR ≥30 mcg/L to mg/L	Higher risk	2047 (4.1)
ACR <30 mcg/L to mg/L	Lower risk	48743(95.9)
eGFR		244166(100)
eGFR < 60 mL/min/1.73m^2	Higher risk	27241(11.2)
eGFR20 ≥ 60 mL/min/1.73m^2	Lower risk	216925(88.8)
BMI		192455(100)
BMI ≥ 30 (Obesity)	Higher risk	64832(33.7)
BMI < 30	Lower risk	127511 (66.3)

* The denominators for percentages are the total number of each CMRFs tests

https://doi.org/10.1371/journal.pone.0223179.t003

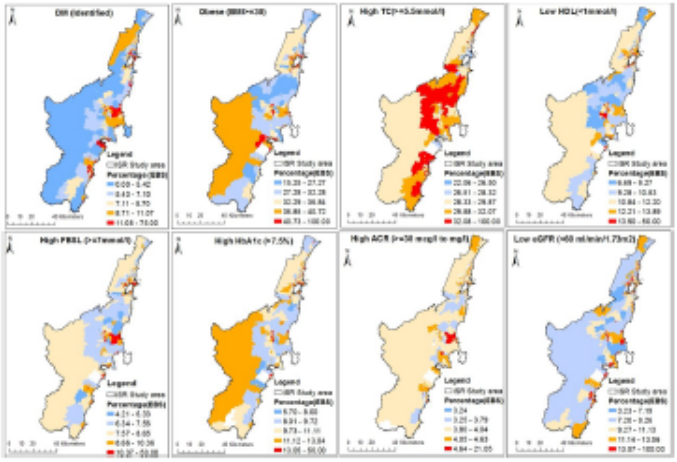

Fig 2. Geographic distribution of the proportion of CMRFs within the Illawarra Shoalhaven region of the NSW Australia.

https://doi.org/10.1371/journal.pone.0223179.g002

Burrill had HH clusters of DM, TC and eGFR. The LL clusters were mainly around the less populated north, central and south ends of the study area, except for TC. The TC clusters demonstrated a reverse pattern in comparison with all other CMRFs, where HH clusters were mainly around the less populated central and southern ends of the ISR and a few instances in the north-eastern end of the study area. LL clusters of TC were found around the immediate surroundings of Lake Illawarra. Fig 3 illustrates the spatial clustering of CMRFs in the study area.

Discussion

Place has always been a key element in human health and epidemiology. In the present study, we explored the geographic distribution of eight CMRFs in 980 SA1s in a regional area of NSW, Australia. The study is a first of its kind known to us in providing a comprehensive small area-level profile of a wide range of cardiometabolic risk factors, and provides an example of using population-derived routine laboratory data for area-level research.

Higher rates and clustering of CMRFs were mostly observed along the more densely populated eastern coast line of the study region. Also, some areas were common for multiple risk factors as their distribution pattern frequently converged in these areas, for example areas

Table 4. Spatial autocorrelation (Moran's I) of CMRFs.

CMRFs	Moran's I	z-score	p-value
DM	0.097	27.952	<0.0001
Obesity	0.328	92.086	<0.0001
High FBSL	0.184	51.539	<0.0001
High HbA1c	0.101	28.030	<0.0001
High TC	0.146	41.154	<0.0001
Low HDL	0.174	48.733	<0.0001
Albuminuria	0.028	8.096	<0.0001
Low eGFR	0.069	19.699	<0.0001

https://doi.org/10.1371/journal.pone.0223179.t004

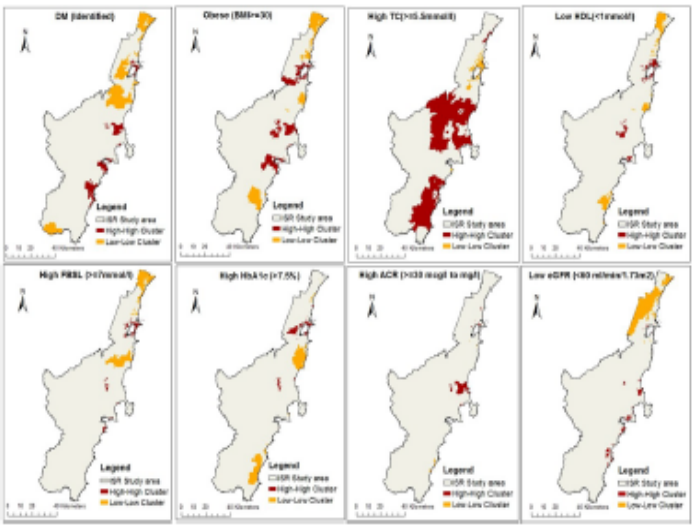

Fig 3. Local Moran's I cluster maps showing high-high and low-low spatial associations of CMRFs within the Illawarra Shoalhaven region of the NSW Australia.

https://doi.org/10.1371/journal.pone.0223179.g003

around Lake Illawarra and south of Jervis Bay. However, not all populated areas were involved in this pattern, and some less populated areas also had higher rates of risk. Spatial analyses revealed significant spatial autocorrelation for all eight CMRFs. Patterns of clustering were different for each CMRFs at the small-area scale used in this study, which provides directions for future research using multilevel analytic methods.[65]

The distribution of high TC values were generally reversed to those distributions of other CMRFs described in this study. The reason for this observation is yet to be explored, but a possible treatment effect is suspected as the lower risk areas were often densely populated areas. It is possible that the people residing in these areas have better access to health care services and more frequently prescribed anti-cholesterol drugs.[66,67] However, not all densely populated areas were involved in this 'higher risk' TC distribution pattern and further research is required.

The current study adds to the limited studies from Oceania reporting on geographic variation of CMRFs, and the first from regional Australia. Previous studies from Australia have reported geographic variation of 42% in the odds of being diagnosed with DM among adults living in Sydney.[38] Another study reported geographic variation in glycated haemoglobin (HbA1c) values across 767 Census Collection Districts (CDs) in Adelaide. [44] The study builds on previous research by investigating the distribution of a wide range of CMRFs, which appears to be unique in the literature.

This study must be considered within its limitations. First, the cross-sectional design of the study precludes causal inference. Second, the descriptive analyses performed in this study indicate only significant variations in the geographic distribution of CMRFs, but does not differentiate the individual and/or area-level attributes which might be contributing to this variation. [13] Third, the maps include areas with no test data. Fourth, the study data were obtained from people attending health care services; therefore its point-estimates may not be representative of the general population. Fifth, we cannot exclude the possibility that a higher proportion of positive tests in an area could be due to greater access to pathology services; however exploring this possibility was beyond the scope of the current study.

Future research is required to understand the reasons for the geographic variation reported in this paper. The findings reported in this study suggest hypotheses that will be further explored using appropriate multilevel/hierarchical analyses to differentiate and quantify the individual and area-level contributions to this variation.[65,68–70] Such hierarchical analyses will have the potential to inform development of appropriate area-level health care service policy initiatives. It is important to differentiate the contributions of individual (e.g. age, sex, etc.) and area (e.g. socioeconomic disadvantage, access or proximity to health care services, etc) level attributes to the different patterns of clustering to inform targeted area-level preventive interventions and future health service commissioning decisions to these areas.

In conclusion, area-level descriptive analyses of CMRFs have the potential to highlight inequalities in the geographic distribution of CMRFs. Regional planning for the prevention and management of CMRFs requires information about its epidemiology within specific communities or areas. Centralised approaches of disease prevention and management may not suit regional requirements as the disease pattern in regional areas may differ to those in metropolitan areas and cities. Area specific evidence through regional health care research is important to inform health care service commissioning for area specific decisions and policy developments. This paper demonstrates an initial step in such regional health care research, and a feasible method using population data derived from routine clinical practice.

Acknowledgments

We would like to thank Southern IML Pathology and staff for their generosity in providing data for the SIMLR Cohort Study and ongoing support. In particular we would like to thank Mr Bryan Jones for providing the technical expertise for the data acquisition and helpful comments on the manuscript. Southern IML Pathology are the owners of the data contained within this publication and the Illawarra Health and Medical Research Institute (IHMRI) is the custodian facilitating access to the data. https://www.ihmri.org.au/research-projects/simlr-cohort-study/

Author Contributions

Conceptualization: Renin Toms, Darren J. Mayne, Xiaoqi Feng, Andrew Bonney.

Data curation: Renin Toms, Darren J. Mayne.

Formal analysis: Renin Toms.

Investigation: Renin Toms.

Methodology: Renin Toms, Darren J. Mayne, Andrew Bonney.

Project administration: Andrew Bonney.

Resources: Andrew Bonney.

Software: Renin Toms, Darren J. Mayne.

Supervision: Darren J. Mayne, Xiaoqi Feng, Andrew Bonney.

Validation: Andrew Bonney.

Writing – original draft: Renin Toms.

Writing – review & editing: Darren J. Mayne, Xiaoqi Feng, Andrew Bonney.

References

1. Danaei G, Lu Y, Singh GM, Carnahan E, Stevens GA, Cowan MJ, et al. Cardiovascular disease, chronic kidney disease, and diabetes mortality burden of cardiometabolic risk factors from 1980 to 2010: A

comparative risk assessment. Lancet Diabetes Endocrinol. 2014; https://doi.org/10.1016/S2213-8587(14)70102-0 PMID: 24842598

2. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, Jafar TH, Heerspink HJL, Mann JF, et al. Chronic kidney disease and cardiovascular risk: Epidemiology, mechanisms, and prevention. The Lancet. 2013. https://doi.org/10.1016/S0140-6736(13)60595-4
3. Wadwa RP, Urbina EM, Daniels SR. Cardiovascular disease risk factors. Epidemiology of Pediatric and Adolescent Diabetes. 2008.
4. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: The Framingham heart study. Circulation. 2008; https://doi.org/10.1161/CIRCULATIONAHA.107.699579 PMID: 18212285
5. Liu M, Li XC, Lu L, Cao Y, Sun RR, Chen S, et al. Cardiovascular disease and its relationship with chronic kidney disease. European review for medical and pharmacological sciences. 2014.
6. Hubert HB, Feinleib M, McNamara PM, Castelli WP. Obesity as an independent risk factor for cardiovascular disease: A 26-year follow-up of participants in the Framingham Heart Study. Circulation. 1983; https://doi.org/10.1161/01.CIR.67.5.968 PMID: 6219830
7. World Health Organisation. WHO | The top 10 causes of death [Internet]. World Health Organization; 2017. Available: http://www.who.int/mediacentre/factsheets/fs310/en/
8. World Health Organization. WHO | Noncommunicable diseases. In: WHO [Internet]. WHO; 2017 [cited 10 Mar 2018]. Available: http://www.who.int/mediacentre/factsheets/fs355/en/
9. World Health Organization; World Heart Federation and World Stroke Organization. Global Atlas on Cardiovascular disease prevention and control. Glob atlas Cardiovasc Dis Prev Control. 2011; 155. NLM classification: WG 120
10. AIHW-Australian Institute of Health and Welfare. Cardiovascular disease, diabetes and chronic kidney disease: Australian facts: morbidity—hospital care [Internet]. 2017. Available: https://www.aihw.gov.au/reports/heart-stroke-vascular-disease/cardiovascular-diabetes-chronic-kidney-morbidity/contents/table-of-contents
11. AIHW-Australian Institute of Health and Welfare. Cardiovascular disease, diabetes and chronic kidney disease—Australian facts: Morbidity–Hospital care. Cardiovascular, diabetes and chronic kidney disease series no. 3. Canberra; 2014.
12. AIHW-Australian Institute of Health and Welfare. Australia's Health 2014 [Internet]. Canberra; 2014. Available: http://www.aihw.gov.au/publication-detail/?id=60129547206
13. Merlo J, Mulinari S, Wemrell M, Subramanian S V., Hedblad B. The tyranny of the averages and the indiscriminate use of risk factors in public health: The case of coronary heart disease. SSM—Popul Heal. 2017; https://doi.org/10.1016/j.ssmph.2017.08.005 PMID: 29349257
14. Ma J, King AC, Wilson SR, Xiao L, Stafford RS. Evaluation of lifestyle interventions to treat elevated cardiometabolic risk in primary care (E-LITE): A randomized controlled trial. BMC Fam Pract. 2009; https://doi.org/10.1186/1471-2296-10-71 PMID: 19909549
15. Tourlouki E, Matalas AL, Panagiotakos DB. Dietary habits and cardiovascular disease risk in middle-aged and elderly populations: a review of evidence. Clinical interventions in aging. 2009.
16. Dehghani A, Kumar Bhasin S, Dwivedi S, Kumar Malhotra R. Influence of Comprehensive Life Style Intervention in Patients of CHD. Glob J Health Sci. 2015; https://doi.org/10.5539/gjhs.v7n7p6 PMID: 26153198
17. Ard JD, Gower B, Hunter G, Ritchie CS, Roth DL, Goss A, et al. Effects of Calorie Restriction in Obese Older Adults: The CROSS ROADS Randomized Controlled Trial. J Gerontol A Biol Sci Med Sci. 2017; https://doi.org/10.1093/gerona/glw237 PMID: 28003374
18. Weiss EP, Fontana L. Caloric restriction: Powerful protection for the aging heart and vasculature. American Journal of Physiology—Heart and Circulatory Physiology. 2011. https://doi.org/10.1152/ajpheart.00685.2011 PMID: 21841020
19. Nissinen A, Berrios X, Puska P. Community-based noncommunicable disease interventions: Lessons from developed countries for developing ones. Bulletin of the World Health Organization. 2001.
20. O'Connor Duffany K, Finegood DT, Matthews D, McKee M, Venkat Narayan KM, Puska P, et al. Community Interventions for Health (CIH): A novel approach to tackling the worldwide epidemic of chronic diseases. CVD Prev Control. 2011; https://doi.org/10.1016/j.cvdpc.2011.02.005
21. Parker DR, Assaf AR. Community interventions for cardiovascular disease. Primary Care—Clinics in Office Practice. 2005. https://doi.org/10.1016/j.pop.2005.09.012 PMID: 16326217
22. Fradelos EC, Papathanasiou I V, Mitsi D, Tsaras K, Kleisiaris CF, Kourkouta L. Health Based Geographic Information Systems (GIS) and their Applications. Acta Inform Med. 2014; 22: 402–5. https://doi.org/10.5455/aim.2014.22.402-405 PMID: 25684850

23. Auchincloss AH, Gebreab SY, Mair C, Diez Roux A V. A review of spatial methods in epidemiology, 2000–2010. Annu Rev Public Health. 2012; 33: 107–22. https://doi.org/10.1146/annurev-publhealth-031811-124655 PMID: 22429160

24. Lawson A (Andrew B. Statistical methods in spatial epidemiology [Internet]. Wiley; 2006. Available: https://www.wiley.com/en-us/Statistical+Methods+in+Spatial+Epidemiology%2C+2nd+Edition-p-9780470014844

25. Craglia M, Maheswaran R, Maheswaran R. GIS in Public Health Practice [Internet]. Craglia M, Maheswaran R, editors. CRC Press; 2016. https://doi.org/10.1201/9780203720349

26. Jarrahi AM, Zare M, Sadeghi A. Geographic Information Systems (GIS), an Informative Start for Challenging Process of Etiologic Investigation of Diseases and Public Health Policy Making. Asian Pacific J Cancer Care. 2017; 2: 1–1. https://doi.org/10.31557/APJCC.2017.2.1.1

27. Nykiforuk CIJ, Flaman LM. Geographic Information Systems (GIS) for Health Promotion and Public Health: A Review. Health Promot Pract. 2011; 12: 63–73. https://doi.org/10.1177/1524839909334624 PMID: 19546198

28. Cromley EK. Using GIS to Address Epidemiologic Research Questions. Curr Epidemiol Reports. 2019; 6: 162–173. https://doi.org/10.1007/s40471-019-00193-6

29. Shafran-Nathan R, Levy I, Levin N, Broday DM. Ecological bias in environmental health studies: the problem of aggregation of multiple data sources. Air Qual Atmos Heal. 2017; 10: 411–420. https://doi.org/10.1007/s11869-016-0436-x

30. Portnov BA, Dubnov J, Barchana M. On ecological fallacy, assessment errors stemming from misguided variable selection, and the effect of aggregation on the outcome of epidemiological study. J Expo Sci Environ Epidemiol. 2007; 17: 106–121. https://doi.org/10.1038/sj.jes.7500533 PMID: 17033679

31. Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health. 2007; 61: 98–102. https://doi.org/10.1136/jech.2005.043117 PMID: 17234866

32. Lawson A (Andrew B. Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology, Third Edition.

33. Lawson A (Andrew B., Browne WJ (William J, Vidal Rodeiro CL. Disease mapping with WinBUGS and MLwiN [Internet]. J. Wiley; 2003. Available: https://www.wiley.com/en-us/Disease+Mapping+with+WinBUGS+and+MLwiN-p-9780470856048

34. Lawson A (Andrew B., Kleinman K. Spatial and syndromic surveillance for public health [Internet]. John Wiley; 2005. Available: https://www.wiley.com/en-us/Spatial+and+Syndromic+Surveillance+for+Public+Health-p-9780470092484

35. Gabert R, Thomson B, Gakidou E, Roth G. Identifying High-Risk Neighborhoods Using Electronic Medical Records: A Population-Based Approach for Targeting Diabetes Prevention and Treatment Interventions. PLoS ONE [Electronic Resour. 2016; 11: e0159227. https://doi.org/10.1371/journal.pone.0159227 PMID: 27463641

36. Barber S, Diez Roux A V, Cardoso L, Santos S, Toste V, James S, et al. At the intersection of place, race, and health in Brazil: Residential segregation and cardio-metabolic risk factors in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Soc Sci Med. 2017/08/20. 2018; 199: 67–76. https://doi.org/10.1016/j.socscimed.2017.05.047 PMID: 28821371

37. Toms R, Bonney A, Mayne DJ, Feng X, Walsan R. Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature. Int J Health Geogr. 2019; 18: 1. https://doi.org/10.1186/s12942-018-0165-5 PMID: 30621786

38. Astell-Burt T, Feng X. Geographic inequity in healthy food environment and type 2 diabetes: can we please turn off the tap? Med J Aust. 2015; 203: 246–248.

39. Zhou M, Astell-Burt T, Bi Y, Feng X, Jiang Y, Li Y, et al. Geographical variation in diabetes prevalence and detection in china: multilevel spatial analysis of 98,058 adults. Diabetes Care. 2015; 38: 72–81. https://doi.org/10.2337/dc14-1100 PMID: 25352654

40. Alkerwi A, Bahi IE, Stranges S, Beissel J, Delagardelle C, Noppe S, et al. Geographic Variations in Cardiometabolic Risk Factors in Luxembourg. Int J Environ Res Public Heal [Electronic Resour. 2017; 14: 16. https://dx.doi.org/10.3390/ijerph14060648

41. Lawlor DA, Bedford C, Taylor M, Ebrahim S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. J Epidemiol Community Heal. 2003; 57: 134–140. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=med4&AN=12540690

42. Barker LE, Kirtland KA, Gregg EW, Geiss LS, Thompson TJ. Geographic distribution of diagnosed diabetes in the U.S.: a diabetes belt. Am J Prev Med. 2011; 40: 434–439. https://doi.org/10.1016/j.amepre.2010.12.019 PMID: 21406277

43. Valdes S, Garcia-Torres F, Maldonado-Araque C, Goday A, Calle-Pascual A, Soriguer F, et al. Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@bet.es study. Rev Esp Cardiol. 2014; 67: 442–448. http://dx.doi.org/10.1016/j.rec.2013.09.029

44. Paquet C, Chaix B, Howard NJ, Coffee NT, Adams RJ, Taylor AW. Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia. Int J Env Res Public Heal. 2016;13. https://doi.org/10.3390/ijerph13050519 PMID: 27213423

45. Ocana-Riola R. Common errors in disease mapping. Geospat Health. 2010; 4: 139–154. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=20503184 https://doi.org/10.4081/gh.2010.196 PMID: 20503184

46. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. BMC Public Health. 2009; https://doi.org/10.1186/1471-2458-9-88 PMID: 19320986

47. Australian Bureau of Statistics. 2011 Census data [Internet]. Commonwealth of Australia; Available: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280

48. Bonney A, Mayne DJ, Jones BD, Bott L, Andersen SE, Caputi P. Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users—a cross-sectional study. PLoS ONE. 2015; 10. https://doi.org/10.1371/journal.pone.0137261 PMID: 26317861

49. Australian Bureau of Statistics. Australian Statistical Geography Standard (ASGS): Volume 1—Main Structure and Greater Capital City Statistical Areas:STATISTICAL AREA LEVEL 1 (SA1) [Internet]. 2016 [cited 21 Oct 2018]. Available: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by Subject/1270.0.55.001~July 2016~Main Features~Statistical Area Level 1 (SA1)~10013

50. The Royal Australian College of General Practitioners, Diabetes Australia. General Practice Management of Type 2 Diabetes 2016–2018 [Internet]. The Royal Australian College of General Practitioners. 2016. https://doi.org/10.1007/s00125-010-2011-6

51. Australian Bureau of Statistics. Australian Health Survey: Biomedical Results for Chronic Diseases, 2011–12 [Internet]. Commonwealth of Australia; ou = Australian Bureau of Statistics; Available: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.005main+features12011-12

52. Cheng FW, Gao X, Mitchell DC, Wood C, Still CD, Rolston D, et al. Body mass index and all-cause mortality among older adults. Obesity. 2016; 24: 2232–2239. https://doi.org/10.1002/oby.21612 PMID: 27570944

53. Littman A, Boyko E, . . . MM-P chronic, 2012 undefined. Evaluation of a weight management program for veterans. ncbi.nlm.nih.gov. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3437789/

54. Li W, Kelsey J, Zhang Z, . . . SL-. . . journal of public, 2009 undefined. Small-area estimation and prioritizing communities for obesity control in Massachusetts. ajph.aphapublications.org. Available: https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2008.137364

55. Australian Bureau of Statistics. Australian Health Survey: Biomedical Results for Chronic Diseases, 2011–12. Commonwealth of Australia. 2013. 4364.0.55.005

56. National heart foundation of Australia. Lipid management profile for health professionals [Internet]. Available: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management

57. National Kidney foundation (USA). Albumin creatinine Ratio (ACR) [Internet]. 2018. Available: https://www.kidney.org/kidneydisease/siemens_hcp_acr

58. Kidney Health Australia. Fact sheet: Estimated Glomerular Filtration Rate (eGFR) [Internet]. Available: www.kidney.org.au

59. WHO. Obesity: Preventing and managing the global epidemic. World Health Organization: Technical Report Series. WHO Technical Report Series, no. 894. 2000. ISBN 92 4 120894 5

60. Li H, Calder CA, Cressie N. Beyond Moran's I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model. Geogr Anal. 2007; 39: 357–375. https://doi.org/10.1111/j.1538-4632.2007.00708.x

61. Moran PAP. Notes on Continuous Stochastic Phenomena. Biometrika. 1950; 37: 17. https://doi.org/10.2307/2332142 PMID: 15420245

62. Anselin L. Local Indicators of Spatial Association-LISA. Geogr Anal. 2010; 27: 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

63. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at https://www.R-project.org/.

64. Environmental Systems Research Institute (ESRI). ArcGIS 10.4.1, ESRI Inc. Redlands, CA, USA. (https://www.esri.com/).

65. Merlo J, Wagner P, Ghith N, Leckie G. An Original Stepwise Multilevel Logistic Regression Analysis of Discriminatory Accuracy: The Case of Neighbourhoods and Health. Moerbeek M, editor. PLoS One. 2016; 11: e0153778. https://doi.org/10.1371/journal.pone.0153778 PMID: 27120054

66. Stocks NP, McElroy H, Ryan P, Allan J. Statin prescribing in Australia: socioeconomic and sex differences. Med J Aust. 2004; 180: 229–231. https://doi.org/10.5694/J.1326-5377.2004.TB05891.X PMID: 14984343

67. Stocks N, Ryan P, Allan J, Williams S, Willson K. Gender, socioeconomic status, need or access? Differences in statin prescribing across urban, rural and remote Australia. Aust J Rural Health. 2009; 17: 92–96. https://doi.org/10.1111/j.1440-1584.2009.01043.x PMID: 19335599

68. Merlo J, Viciana-Fernández FJ, Ramiro-Fariñas D, Research Group of Longitudinal Database of Andalusian Population (LDAP). Bringing the individual back to small-area variation studies: A multilevel analysis of all-cause mortality in Andalusia, Spain. Soc Sci Med. 2012; 75: 1477–1487. https://doi.org/10.1016/j.socscimed.2012.06.004 PMID: 22795359

69. Merlo J, Wagner P, Leckie G. A simple multilevel approach for analysing geographical inequalities in public health reports: The case of municipality differences in obesity. Health Place. 2019; 58: 102145. https://doi.org/10.1016/j.healthplace.2019.102145 PMID: 31195211

70. Merlo J, Asplund K, Lynch J, Rastam L, Dobson A, World Health Organization MONICA Project. Population Effects on Individual Systolic Blood Pressure: A Multilevel Analysis of the World Health Organization MONICA Project. Am J Epidemiol. 2004; 159: 1168–1179. https://doi.org/10.1093/aje/kwh160 PMID: 15191934

Supplementary Material: 3 (Published version of research article 3)

www.nature.com/scientificreports

SCIENTIFIC REPORTS
nature research

Check for updates

OPEN

Geographic variation in cardiometabolic risk factor prevalence explained by area-level disadvantage in the Illawarra-Shoalhaven region of the NSW, Australia

Renin Toms[1,2✉], Darren J. Mayne[1,2,3,4], Xiaoqi Feng[2,5,6] & Andrew Bonney[1,2]

Cardiometabolic risk factors (CMRFs) demonstrate significant geographic variation in their distribution. The study aims to quantify the general contextual effect of the areas on CMRFs; and the geographic variation explained by area-level socioeconomic disadvantage. A cross sectional design and multilevel logistic regression methods were adopted. Data included objectively measured routine pathology test data between years 2012 and 2017 on: fasting blood sugar level; glycated haemoglobin; total cholesterol; high density lipoprotein; urinary albumin creatinine ratio; estimated glomerular filtration rate; and body mass index. The 2011 Australian census based Index of Relative Socioeconomic Disadvantage (IRSD) were the area-level study variables, analysed at its smallest geographic unit of reporting. A total of 1,132,029 CMRF test results from 256,525 individuals were analysed. After adjusting for individual-level covariates, all CMRFs significantly associated with IRSD and the probability of *higher risk* CMRFs increases with greater area-level disadvantage. Though the specific contribution of IRSD in the geographic variation of CMRF ranged between 57.8 and 14.71%, the general contextual effect of areas were found minimal (ICCs 0.6–3.4%). The results support universal interventions proportional to the need and disadvantage level of populations for the prevention and control of CMRFs, rather than any area specific interventions as the contextual effects were found minimal in the study region.

The prevalence of cardiometabolic risk factors (CMRFs) varies geographically[1,2]. Previous research has reported higher prevalence of CMRFs in certain localities: typically in areas of higher socioeconomic disadvantage[3–23]. However, frequently these studies have been based on measures of association or geographical variance rather than reporting clustering or the share of the total variance that is at the area-level[3–23]. Quantifying the clustering and the proportion of geographic variation in CMRFs contributed by area-level socioeconomic disadvantage can aid in designing appropriate area-level approaches to help prevent CMRFs. Chronic and uncontrolled CMRFs predispose individuals to the development of cardiovascular disease (CVD), which continues to be the leading cause of health care expenditure and premature mortality worldwide[24].

In Australia, a social gradient is observed in the prevalence of many chronic conditions including various CMRFs (e.g.diabetes and chronic kidney disease)[25]. Generally, Australians enjoy better health than people in many other countries in the world. However, within Australia this better health is not equally distributed[26]. It is

[1]School of Medicine, University of Wollongong, Wollongong, NSW 2522, Australia. [2]Illawarra Health and Medical Research Institute, Wollongong, NSW 2522, Australia. [3]Public Health Unit, Illawarra Shoalhaven Local Health District, Warrawong, NSW 2502, Australia. [4]School of Public Health, The University of Sydney, Sydney, NSW 2006, Australia. [5]Population Wellbeing and Environment Research Lab (PowerLab), School of Health and Society, Faculty of Social Sciences, University of Wollongong, Wollongong, Australia. [6]School of Public Health and Community Medicine, University of New South Wales, Sydney, NSW, Australia. ✉email: rmbst288@uowmail.edu.au

often-reported that socioeconomically disadvantaged individuals in Australia, on average, experience a greater disease burden than their less disadvantaged counterparts[25–28]. This tendency is also evident at a contextual level when studies have investigated association of CMRFs with area-level socioeconomic disadvantage in Australia [5,7] and globally [4,9,10,12,14,16–18,20,23,29–33].

Consistent with this, men from highly urbanised environments have been reported to have higher incidence of coronary heart disease with increasing residential area socioeconomic disadvantage, after adjusting for individual characteristics[18]. Also, lower area-level disadvantage has been reported as being associated with lower prevalence of some behavioural cardiac risk factors such as smoking, physical inactivity and obesity etc. in some studies[9,10,34]. Most of the reported associations of CMRFs with area-level socioeconomic disadvantage were independent of individual-level characteristics such as age and educational attainment. Even though the area-level associations of CMRFs were significant in these studies, the results were often dependent on the CMRF analysed, the measures of area-level socioeconomic disadvantage and the geographic scale at which associations were examined[35].

Multilevel analyses of CMRFs based on the average measures of association or variation alone are insufficient to report the geographical variance as similar associations were possible with very different scenarios of area variance[36]. Multilevel findings extending on the general contextual effects and reporting the proportion of the total area-level variance along with the measures of clustering and the average measures of association or variation are appropriate and informative in reporting area-level influences, but less common[23,36–38]. To differentiate the relative importance of individual versus area-level interventions for the prevention and control of CMRFs, the geographical component of the total individual risk variance has to be identified in a multilevel approach.

Therefore, the aims of this study are to (1) quantify the general contextual or geographic effect of areas on CMRFs, over and above their individual-level compositions; and to (2) quantify the geographic variation across multiple CMRFs specifically explained by area-level socioeconomic disadvantage, within the Illawarra-Shoalhaven region of NSW Australia. Quantification of the general contextual effect and the variation specifically explained by area-level socioeconomic disadvantage will assist our understanding of the socioeconomic context of CMRFs in the study region and provide guidance for health service commissioning more generally nationally.

Methods

A cross-sectional multilevel design was adopted to account for the hierarchical nature of the data and analyses. No informed consent were obtained for the individual-level data used in this study, as the study used existing data which were already de-identified. The study was approved by the University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical Human Research Ethics Committee (HREC protocol No: 2017/124). All the methods and analyses were performed meeting the relevent ethical guidelines and regulations of the committee.

Study area and data. The study was conducted in the Illawarra-Shoalhaven region of the New South Wales (NSW) state in Australia. The Illawarra-Shoalhaven region is a coastal plain along the south-east border of NSW; situates at the immediate south of the metropolitan boundaries of Sydney; and encompasses multiple regional cities, towns and rural areas. This region covers a land area of 5,615 km^2, and had an estimated residential population of 369,469 at the time of the 2011 Australian Census of Population and Housing conducted by the Australian Bureau of Statistics (ABS)[39]. Statistical Area level 1 (SA1), the smallest geographical unit of the 2011 census data release, was the area-level unit of analysis in this study[39]. SA1s typically have a population size of 200 to 800 persons (average 400), and the Illawarra-Shoalhaven region covers a total of 980 conterminous SA1s[39].

The CMRF test data in this study were extracted from the Southern IML Research (SIMLR) Study database, which is comprised of de-identified and internally linked pathology results from a major network of pathology services in the study region. The individual-level data in SIMLR database are geocoded to their corresponding SA1 areas, but not to their residential address, for privacy and confidentiality concerns. More details on this data source, procurement and access are published elsewhere[7]. The CMRF test data were extracted for non-pregnant individuals aged 18 years or older presenting for testing between 01 January 2012 and December 2017. Only the most recent test result was included if an individual had undergone the same test multiple times in this data period. Test data with missing details on the individual and area-level factors analysed in this study were excluded from the analyses.

Variables. *Outcome variable.* Results of the CMRF tests were the individual-level outcome variables. Data on the seven CMRF tests analysed in this study included: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); urinary albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); and objectively-measured body mass index (BMI). These CMRF test results were dichotomised into *higher risk* and *lower risk* values based on the current national and international guidelines on risk definitions (Table 1).

Study variable. The 2011 ABS census based Index of Relative Socioeconomic Disadvantage (IRSD) of the SA1s was the study variable. IRSD summarises a range of measures of relative socioeconomic disadvantage of people and households within SA1s and includes: level of income; education; employment; family structure; disability; housing; transportation; and internet connection[45]. This study uses IRSD reported as quintiles; the lowest quintile (Q1) indicating the most disadvantaged SA1s and the highest quintile (Q5) the least disadvantaged SA1s[45]. The IRSD quintiles in the study were derived by ABS from the distribution of IRSD scores for the Illawarra-Shoalhaven region based on the 2011 census. The study region has a diverse IRSD profile with representation across IRSD scores in comparison with Australia as a whole, making the region useful for population-level studies[46].

	Higher risk CMRFs	Definition
1	High FBSL	FBSL ≥ 7.0 mmol/L[40]
2	High HbA1c	HbA1c > 7.5%[40]
3	High TC	TC ≥ 5.5 mmol/L[41]
4	Low HDL	HDL < 1 mmol/L[42]
5	High ACR	ACR ≥ 30 mcg/L to mg/L[45]
6	Low eGFR	eGFR < 60 mL/min/1.73 m^2[45]
7	Obesity	BMI ≥ 30 kg/m^2[44]

Table 1. Definitions of *higher risk* CMRFs test results. *CMRFs* Cardiometabolic risk factors, *FBSL* Fasting Blood Sugar Level, *HbA1c* Glycated Haemoglobin, *TC* Total Cholesterol, *HDL* High Density Lipoprotein, *ACR* Albumin Creatinine Ratio, *eGFR* estimated Glomerular Filtration Rate, *BMI* Body Mass Index.

Covariates. Analyses were adjusted for sex (male and female) and age group (18–29 years, 30–39 years, 40–49 years, 50–59 years, 60–69 years, 70–79 years, 80+ years) of each individual at the time of the pathology collection of the CMRFs tests analysed in his study.

Statistical analyses. Initially, descriptive statistics of all individual and area-level variables were performed. Thereafter, single level and multilevel logistic regression models were fitted for the CMRF test data of individuals (Level 1), nested within SA1s (Level 2). For each of the seven CMRFs analysed in this study, a hierarchy of four multilevel models at SA1 level were fit that included fixed effects for age, sex and IRSD and random effect (intercept) for SA1. Model 0 was a single level model adjusted for age and sex; Model 1 (M1) was null model at level 2; Model 2 (M2) adjusted for age and sex at level 2; Model 3 (M3) adjusted for the area-level study variable (IRSD) only at level 2; and the final model Model 4 (M4) included both M2 and M3 covariates (age, sex and IRSD) at level 2. The estimated regression coefficients of the derived models were exponentiated to calculate odds ratios (ORs).The goodness of fit of the models were identified using Likelihood Ratio Tests (LRT) at $p<0.05$ level of significance. The general equation of the fully adjusted model is:

$$y_{ij} \sim \text{Binomial}\,(1, \pi_{ij}) \tag{1}$$

$$\text{logit}(\pi_{ij}) = \beta_0 + \beta_1 \text{Age}_{ij} + \beta_2 \text{Sex}_{ij} + \beta_3 \text{IRSD}_{ij} + u_j \tag{2}$$

$$u_j \sim N(0, \tau_u^2) \tag{3}$$

where y_{ij} denote the binary response of CMRF test outcome (as '*higher risk*' or '*lower risk*', based on the adopted definitions) for individual i in the area (SA1) j; π_{ij} denotes the probability that individual i in area (SA1) j has a '*higher risk*' CMRF test outcome given their individual-level age_{ij} and sex_{ij} and their area-level IRSD index. The β_1, β_2, β_3 are the regression coefficients which measure the associations between the log-odds of the CMRF outcome and each covariate all else equal, and when exponentiated these are translated to ORs[36]. u_j is the random effect for the area (SA1) j and τ^2_u is the area level variance, which has to be estimated.

Model comparison. The Akaike Information Criterion (AIC) was used to evaluate model fit. The derived multilevel models were compared for: area-level variance (τ^2) at SA1 (level 2) level; proportional change in variance (PCV); Intra-cluster Correlation Coefficients (ICC); Median Odds Ratios (MORs); area under the receiver operating characteristic (AUC) curve; and the change in AUC.

The τ^2s of the multilevel models were initially identified from each models. PCVs were calculated for models M2s to M4s relative to M1s. The ICCs of the fitted models were calculated using the latent variable approach[47]. This approach assumes that a latent continuous outcome underlies the observed dichotomous outcomes and it is this latent outcome for which the ICC is calculated and interpreted. The ICC measured the expected correlation in CMRF outcomes between two individuals from the same SA1. The higher the ICC, the more relevant area-level context is for understanding individual latent outcome variation[36]. The MOR is calculated as an alternative way of interpreting the magnitude of area-level variance. The MOR translated the area-level variance which were estimated on the log-odds scale to the commonly used OR scale. The MOR result value is interpreted as the median increased odds of identifying the outcome if an individual move to another SA1 with higher risk. Thus, the higher the MOR the greater the general area-level effect and it will equal to 1 in the absence of area-level variance[36]. The general contextual effect of the geographic areas over and above their individual-level composition of the *higher risk* CMRFs, is obtained through the measure of clustering (ICC) in M2s. The geographic variance and ICC in the null models (M1s) of *higher risk* CMRFs may depend on both the contextual and individual-level variables. Therefore, M2s of the *higher risk* CMRFs which adjusted for individual-level attributes is better to provide information on the 'general contextual effect' of the areas. The unique contribution of the area-level study variable (IRSD) to the area-level variance of *higher risk* CMRFs were assessed through the PCVs between M2s and M4s.

The receiver operating characteristic (ROC) curves are created by plotting the true positive rates (TPR) i.e. sensitivity, against the false positive rates (FPR), i.e. 1 specificity for different binary classification thresholds of the predicted probabilities in all the models[48]. Post-estimation, predicted probabilities (π_{ij}) are calculated for

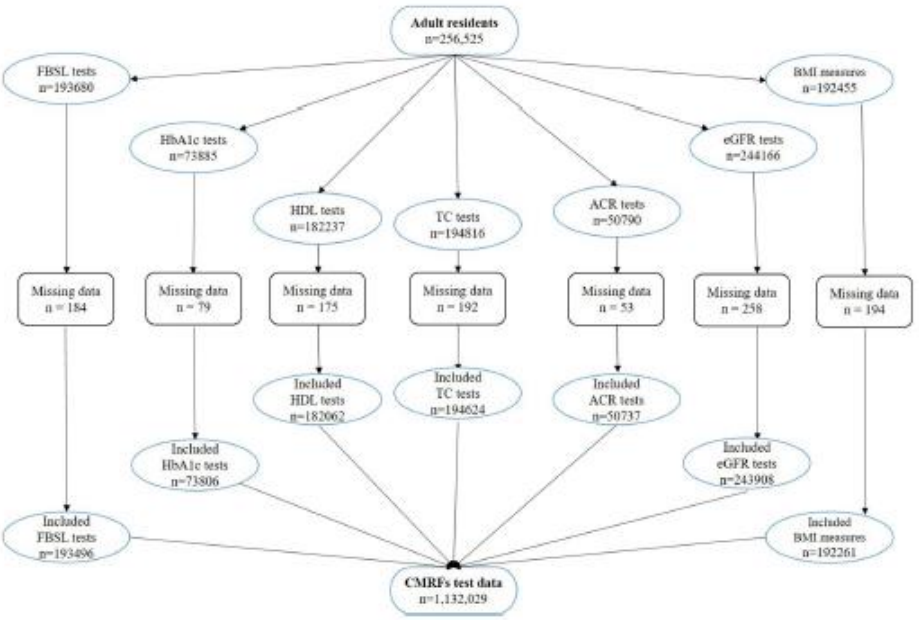

Figure 1. Flow chart of the included/excluded tests in the CMRFs test data.

each individual and are used to calculate the AUC for the model. The AUCs of the models measure the capacity of the models to correctly classify individuals with or without the outcome of a *higher risk* CMRFs analysed in this study, as a function of their predicted probabilities[36]. The AUC values range from 1 and 0.5, where 1 is the perfect predictive discrimination and 0.5 have no predictive power[49]. The AUCs also indicate the general contextual effects and can be compared it to the ICC and the MOR values[36]. The added value of knowing an individual's area of residence besides individual-level information (age and sex) can be obtained through the AUC change in Model 2 in reference to Model 0, where a higher AUC change would indicate higher relevance of areas in relation to CMRFs.

Statistical package. All analyses were performed using R version 3.4.4. (R Foundation for Statistical Computing, Vienna, Austria)[50]. Multi-level models were fit using the *glmer* function in the *lme4* package[51]-likelihood ratio tests were calculated using the *lrtest* function in the *lmtest* package[52]-and ROC curves using the *roc* function in the pROC package[53].

Results

A total of 1,132,029 CMRFs test data which belong to 256,525 individuals were extracted for the analyses. Figure 1 provides a flow chart of the individual tests in CMRF test data. The mean number of tests per person was 4.4. After removing 1,162 (1.0%) test results data with missing details, a total of 1,130,894 tests were included in the analytic data set.

Table 2 provides details of the missing data and test data distribution of each CMRF tests. Most frequently missing data were the IRSD indices from SA1s in the study area for which an IRSD index was not available from ABS 2011 census either due to low populations or poor data quality[54].

Tables 3 and 4 shows the frequencies and relative frequencies of CMRF tests results. Overall, the *higher risk* frequencies of all CMRFs increased with increasing area-level socioeconomic disadvantage, except for TC which demonstrated an inverse trend.

Single and multilevel models for each of the CMRFs analysed in this study are presented in Tables 5, 6, 7, 8, 9, 10 and 11. After adjusting for the covariates, all seven CMRFs were found to be significantly associated with area-level IRSD in the study region. For all but one variable the associations were positive (i.e. increased with area-level disadvantage). TC was the exception; being inversely associated with area-level disadvantage, with the most disadvantaged quintile (Q1) displaying the lowest odds for *higher risk* test results. Among the covariates, there was no significant association between gender and *higher risk* test results of eGFR or BMI. It was also noted that the odds of *higher risk* eGFR tests results accelerated with increasing age group, and the 80+ age group demonstrated a very high odds of being identified with a *higher risk* eGFR tests result in the study region.

	FBSL	HbA1c	TC	HDL	ACR	eGFR	BMI	Total
Extracted	193,680	73,885	194,816	182,237	50,790	244,166	192,455	1,132,029
Missing data								
Test value	1	0	0	0	0	0	0	1
Age	1	1	1	1	0	2	1	7
Sex	0	0	0	0	0	0	0	0
IRSD	182	78	191	174	53	256	193	1,154
Excluded tests	184	79	192	175	53	258	194	1,162
Included tests								
Total n	193,496	73,806	194,624	182,062	50,737	243,908	192,261	1,130,894
(%)	(17.11)	(6.53)	(17.21)	(16.10)	(4.49)	(21.57	(17.00)	(100.00)

Table 2. Table of excluded test data which had missing details. *FBSL* Fasting Blood Sugar Level, *HbA1c* Glycated Haemoglobin, *TC* Total Cholesterol, *HDL* High Density Lipoprotein, *ACR* Albumin Creatinine Ratio, *eGFR* estimated Glomerular Filtration Rate, *BMI* Body Mass Index, *IRSD* Index of the Relative Socioeconomic Disadvantage, *SA1* Statistical Area level 1.

	FBSL		HbA1c		TC		HDL	
CMRFs	Total tests	*Higher risk** results, n (%)	Total tests	*Higher risk** results, n (%)	Total tests	*Higher risk** results, n (%)	Total tests	*Higher risk** results, n (%)
Rates	193,496	16,259 (8.4)	73,806	7,920 (10.73)	194,624	57,506 (29.55)	182,062	21,238 (11.67)
Sex								
Male	83,603	9,279 (4.8)	35,757	4,444 (6.02)	90,950	23,503 (12.0)	85,266	15,872 (8.72)
Female	109,893	6,980 (3.6)	38,049	3,476 (4.71)	103,674	34,003 (17.47)	96,796	5,366 (2.95)
Age (years)								
18–29	19,747	238 (0.1)	3,480	250 (0.34)	14,247	2,127 (1.09)	11,435	1,377 (0.76)
30–39	23,515	459 (0.2)	4,889	293 (0.40)	18,960	4,889 (2.51)	16,787	2,301 (1.26)
40–49	29,424	1,265 (0.65)	8,447	760 (1.03)	31,395	10,719 (5.51)	29,339	3,585 (1.97)
50–59	37,085	2,948 (1.52)	13,510	1,507 (2.04)	39,663	16,316 (8.38)	37,824	4,283 (2.35)
60–69	37,962	4,670 (2.41)	17,665	2,064 (2.80)	40,471	13,620 (7.00)	39,134	4,227 (2.32)
70–79	29,009	4,396 (2.27)	15,715	1,860 (2.52)	31,186	6,748 (3.47)	30,114	3,419 (1.88)
80+	16,754	2,283 (1.18)	10,100	1,186 (1.61)	18,702	3,087 (1.59)	17,429	2046 (1.12)
IRSD								
Most D Q-1	38,885	4,495 (2.32)	17,024	2,429 (3.29)	39,347	10,631 (5.46)	36,625	5,520 (3.03)
Q-2	41,545	3,757 (1.94)	16,680	1,875 (2.54)	41,937	12,015 (6.17)	39,050	4,901 (2.69)
Q-3	39,828	3,386 (1.75)	15,376	1,585 (2.15)	40,401	12,045 (6.19)	37,794	4,201 (2.31)
Q-4	37,137	2,594 (1.34)	13,101	1,138 (1.54)	36,865	11,163 (5.74)	34,566	3,581 (1.97)
Least D Q-5	36,101	2027 (1.05)	11,625	893 (1.21)	36,074	11,652 (5.99)	34,027	3,035 (1.67)

Table 3. Cross-tabulation of individual CMRFs (FBSL, HbA1c, TC and HDL) with the variables in study. *FBSL* Fasting Blood Sugar Level, *HbA1c* Glycated Haemoglobin, *TC* Total Cholesterol, *HDL* High Density Lipoprotein, *Most D* Most Disadvantaged, *Least D* Least Disadvantaged *Refer to Table 1 for *high risk* threshold levels of CMRFs.

The overall comparisons of model random effects are presented in Table 12. Reductions in the AIC values were observed among all CMRFs from the null model (M1) to the final model (M4) indicating a better fit for the final models. In the unadjusted null models, *higher risk* test results of eGFR demonstrated the most area-level variance (0.189) and TC the least (0.026). Adjusting the CMRFs for age and sex initially increased the τ^2 of M2 for FBSL (PCV = + 1.88%), HbA1c (PCV = + 3.02%), HDL (PCV = + 15.25%) and BMI (PCV = + 1.48%). The τ^2 was reduced in the final model among all CMRFs compared with the null models.

The Akaike Information Criterion (AIC) was used to evaluate model fit. The derived multilevel models were compared for: area-level variance (τ^2) at SA1 (level 2) level; proportional change in variance (PCV); Intra-cluster Correlation Coefficients (ICC); Median Odds Ratios (MORs); Area under the receiver operating characteristic (AUC) curve; and the change in AUC.

The ICCs of the unadjusted models ranged between 0.8% in high TC to 5.4% in low eGFR. Inclusion of IRSD after adjusting for age and sex had reduced the ICCs of all CMRFs in the final models, which ranged between 0.4% in low eGFR to 2.0% in obesity test results. The ICCs of the final models were low and suggest very limited area-level contextual effects. The AUC changes in model 2 and MORs of the final model support these findings.

CMRFs	ACR		eGFR		Obesity	
	Total tests	Higher risk* results, n (%)	Total tests	Higher risk* results, n (%)	Total tests	Higher risk* results, n (%)
Rates	50,737	2046 (4.03)	243,908	27,205 (11.15)	192,261	64,875 (33.7)
Sex						
Male	25,043	1,265 (2.49)	108,140	12,441 (5.1)	86,853	29,585 (15.3)
Female	25,694	781 (1.54)	135,768	14,764 (6.05)	105,408	35,290 (18.3)
Age (years)						
18–29	1546	47 (0.09)	32,961	72 (0.03)	23,277	4,582 (2.38)
30–39	2,278	71 (0.14)	29,047	105 (0.04)	22,799	6,535 (3.40)
40–49	4,870	108 (0.21)	35,778	330 (0.14)	30,401	10,595 (5.51)
50–59	9,272	230 (0.45)	42,695	1,112 (0.46)	37,285	13,825 (7.19)
60–69	13,388	412 (0.81)	43,423	3,626 (1.49)	38,370	15,310 (7.96)
70–79	12,337	605 (1.19)	34,406	8,507 (3.49)	30,074	11,324 (5.89)
80+	7,046	573 (1.13)	25,598	13,453 (5.52)	10,055	2,704 (1.41)
IRSD						
Most D Q-1	11,915	638 (1.26)	49,288	7,061 (2.89)	37,476	15,365 (7.99)
Q-2	11,350	485 (0.96)	52,947	6,354 (2.61)	40,172	14,334 (7.46)
Q-3	10,494	391 (0.77)	50,816	5,917 (2.43)	39,133	13,007 (6.77)
Q-4	8,732	308 (0.61)	46,440	4,406 (1.81)	37,370	11,766 (6.12)
Least D Q-5	8,246	224 (0.44)	44,417	3,467 (1.42)	38,110	10,403 (5.41)

Table 4. Cross-tabulation of individual CMRFs (ACR, eGFR, and Obesity) with the variables in study. ACR Albumin Creatinine Ratio, *eGFR* estimated Glomerular Filtration Rate, *BMI* Body Mass Index, *Most D* Most Disadvantaged, *Least D* Least Disadvantaged. *Refer to Table 1 for *higher risk* threshold levels of CMRFs.

Figure 2 provides a comparison of the ROC curves of the fitted models. Model 4 s (age + sex + IRSD adjusted models) and models 3 s (IRSD adjusted models) were chosen for the ROC curve plotting for comparative purpose. The predicted outcomes in the CMRFs plots are for the reference individual, i.e., individuals residing in the least disadvantaged areas (model 3) + female + age group 18–29 years (Model 4). A model curve closer to the top left corner of the subfigures indicate a better predictive accuracy of the model. The single measure summary of the ROC curves, AUCs of the final models ranged 0.62–0.88. The highest AUC value was observed for the final model of low eGFR. The AUC changes of model 2 s in relation to M0s ranged 0.01–0.08, which reconfirm the contextual findings of ICCs that the general contextual effects observed in the models were minimal.

The proportions of the geographic variance in CMRFs contributed by IRSD were estimated through the PCV between M2 and M4. Adjusting the models for IRSD and individual-level variables explained a maximum 92.79% of the variance expressed by the null model of eGFR, reducing the ICC from 5.4 to 0.4%. The changes were least among the adjusted models of TC, with a marginal reduction of ICC from 0.8% to 0.5%. Thus, in the final models, the proportional reduction in variance was the largest for eGFR (PCV = 92.79%) and the least for TC (PCV = 33.27%).

The identified specific contribution of IRSD in the geographic variation of CMRF was the highest among the geographic variance of *higher risk* findings of HDL tests (57.8%), which was closely followed by FBSL (57.14%); HbA1c (53.31%); and ACR (51.17%) test results. The contribution of IRSD was comparatively lower among the geographic variance of the *higher risk* findings of eGFR (41.75%); BMI (41.06%); and TC (14.71%) test results, though not the least. Even though these specific proportions are large, it should be noted that it actually explained a lot of very little (i.e., variance of 0.01–0.07).

Discussion

The study reports on the influence of areas on *higher risk* CMRF distribution and quantifies the specific proportion of geographic variance explained by IRSD. The work adds to the very few studies which consider multiple CMRF variables within the same region, or which are based on population derived data over extended years[14,17,20,29,31,32],and reports on both single and multilevel analyses[38,55]. The results present both the measures of association and area-level variance based on multilevel logistic regression analyses[36]. The findings of the study add to the existing evidence and discussion regarding the relevance of individual versus area-level interventions for the prevention and control of CMRFs.

We found consistent evidence for the association between area-level disadvantage and seven CMRFs among adult health service using residents of the Illawarra-Shoalhaven region in NSW Australia. In adjusted models, the odds of a *higher risk* finding increased with increase in area-level disadvantage among all CMRFs excepting TC, which showed an inverse pattern of association with increase in area-level disadvantage. Thus, in the final models we observed that, over and above individual age and sex, living in a disadvantaged neighbourhood proportionally increased the individual-level probability of being identified with a *higher risk* CMRF. The findings highlight the importance of including of area-level variables into health risk analyses.

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
High FBSL	OR (95% CI)	OR (95% CI)	OR (95% CI)OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.01 (0.01–0.01)	0.09 (0.09–0.09)	0.01 (0.01–0.01)	0.06 (0.06–0.06)	0.01 (0.01–0.01)
Sex					
Female	Reference		–		
Male	1.62 (1.56–1.67)		1.63 (1.56–1.7)		1.63 (1.58–1.7)
Age					
18–29	Reference		–		1.64 (1.41–1.9)
30–39	1.60 (1.36–1.87)		1.63 (1.39–1.9)		3.58 (3.12–4.1)
40–49	3.41 (2.97–3.93)		3.53 (3.07–4.1)		6.81 (5.98–7.8)
50–59	6.48 (5.68–7.42)		6.77 (5.93–7.7)		11.05 (9.71–12.6)
60–69	10.48 (9.21–11.98)		11.07 (9.72–12.6)		13.74 (12.07–15.6)
70–79	13.35 (11.73–15.27)		13.93 (12.22–15.9)		12.02 (10.52–13.7)
80+	12.01 (10.51–13.78)		12.33 (10.79–14.1)		
IRSD					
Q-5				Reference	–
Q-4				1.27 (1.18–1.36)	1.27 (1.18–1.37)
Q-3				1.58 (1.47–1.69)	1.49 (1.39–1.61)
Q-2				1.68 (1.57–1.80)	1.62 (1.50–1.74)
Most D Q-1				2.20 (2.06–2.36)	2.11 (1.96–2.26)
AIC	103,645	111,022.8	103,066.2	110,552.5	102,689.6
Variance		0.101	0.103	0.034	0.044
PCV		–	+ 1.88%	− 66.41%	− 56.33%
ICC (%)		3.00	3.00	1.0	1.3
MOR		1.35	1.36	1.19	1.22
AUC	0.70	0.61	0.73	0.60	0.72
AUC change[†]			+ 0.03		
Proportional variance explained by IRSD: 57.14%					

Table 5. Single and multilevel logistic regression model summaries for high FBSL (FBSL ≥ 7.0 mmol/L). ***$p < 0.001$; [†]Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

The ICCs of CMRFs in all the models were comparatively small (Table 12) in all the models. In the fully adjusted models, the ICCs were further reduced and ranged between 0.4% and 2.0% in low eGFR and BMI respectively. As per the interpretation framework proposed by Merlo et al., an ICC value less than 10% is indicative of very little geographic difference[56]. The AUC change of the model 2 s in relation to the single level models (range 0.01–0.08) reconfirm on these findings. However, this has to be interpreted along with the traditional geographic comparisons such as the proportion of the individuals who are affected with *higher risk* CMRF outcomes. Therefore, a small geographic difference with uniformly higher, medium, or lower proportion of affected individuals indicates homogeneity of the *higher risk* CMRF findings within their geographic units[56]. Such a situation would call for balanced universal approaches to prevent and control the *higher risk* CMRFs, with a proportional focus to the need and disadvantage level of affected populations[57,58]. However, it is also worth noting that when the exposure to an agent is homogenic in a community, the traditional epidemiological methods are not very helpful in identifying their markers of susceptibility[59].

Our results confirm, and are comparable with, associations between area-level disadvantage and CMRFs reported in previous studies[3–23], and extends their findings. The results primarily confirm the geographic variation of CMRFs and associations with area level disadvantage, as reported in previous studies. Further, the study provides means to compare this association which were observed consistently with a range of multiple CMRFs analysed in this study. The study extends on previous reports by differentiating the individual and area-level contributors to the exhibited geographic variance of CMRFs. And most importantly, the general contextual effect and the specific contributions of IRSD on the geographic variance of multiple CMRFs were identified, which is unique in the literature and highly informative for health care service commissioning.

The TC test results often stood apart from the major findings of this study, demonstrating inverse associations with IRSD. However, this was not reflected in the HDL findings, even though both are components of the lipid profile in an individual. This raises the possibility of a medication effect on TC in these areas, where the lipid lowering drugs have a less consistent effect in raising HDL than in lowering TC[60]. Other factors associated with the *higher risk* HDL test results may include uncontrolled diabetes[61], smoking[62], sedentary life style[63,64], obesity[65],

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
High HbA1c	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.07 (0.06–0.08)	0.12 (0.12–0.12)	0.07 (0.06–0.07)	0.08 (0.08–0.09)	0.05 (0.04–0.05)
Sex					
Female	Reference		–		–
Male	1.37 (1.31–1.43)		1.38 (1.32–1.45)		1.39 (1.32–1.45)
Age					
18–29	Reference		–		–
30–39	0.81 (0.68–0.96)		0.81 (0.68–0.96)		0.81 (0.68–0.96)
40–49	1.22 (1.06–1.42)		1.24 (1.07–1.44)		1.26 (1.08–1.46)
50–59	1.53 (1.34–1.77)		1.56 (1.36–1.80)		1.57 (1.36–1.81)
60–69	1.61 (1.40–1.85)		1.64 (1.43–1.88)		1.64 (1.43–1.88)
70–79	1.63 (1.42–1.87)		1.64 (1.42–1.88)		1.62 (1.41–1.86)
80+	1.64 (1.43–1.90)		1.63 (1.41–1.88)		1.60 (1.39–1.85)
IRSD					
Q-5				Reference	–
Q-4				0.08 (0.08–0.09)	1.15 (1.04–1.28)
Q-3				1.14 (1.03–1.27)	1.39 (1.26–1.54)
Q-2				1.40 (1.27–1.54)	1.55 (1.41–1.71)
Most D Q-1				1.54 (1.40–1.69)	2.02 (1.84–2.22)
AIC	49,897	50,114.5	49,690.2	49,875.3	49,453.3
Variance		0.103	0.106	0.047	0.049
PCV		–	+3.02%	−54.82%	−51.91%
ICC (%)		3.0	3.1	1.4	1.5
MOR		1.36	1.36	1.23	1.24
AUC	0.56	0.63	0.64	0.61	0.63
AUC change†			+0.08		
Proportional variance explained by IRSD: 53.31%					

Table 6. Single and multilevel logistic regression model summaries for high HbA1c (HbA1c > 7.5%). ***$p < 0.001$; †Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—Model 1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

and poor diet quality[66,67]. However the reason for the inverse association demonstrated by TC test results are not clearly established within the current study results and requires further research to explore possible individual and area-level contributions.

The study has to be considered within its limitations. Primarily, the cross sectional nature of analyses adopted in this study do not yield support for any causal relationships. In addition, the non-linear and time varying effects of covariates analysed in this study restrict generalisability of their findings though very informative for regional health care service commissioning. Secondly, the IRSD quintiles included as the key explanatory variable represent relative disadvantage in an area and have limitations intrinsic to aggregate measures. Thirdly, it should be noted that the data used in this study are extracted from people already utilising the health care service facilities in the area. Fourthly , the readers should be mindful that the variance reported in this study are attributable to (1) individual level factors (age, sex) analysed at the area-level, (2) area-level contextual influences (IRSD), and (3) other individual and area-level characteristics not considered in this study. However, further individual-level data extractions or collections are not possible with this study's dataset as the de-identification process precludes the inclusion of any further individual level data. Other individual and area-level factors not considered in this study could include: individual-level SES[68], type of neighbourhood food outlets[69–72], poor physical activity resources[73,74], residential density and service availability[75]. Finally, the assumptions of the standard multilevel logistic regression modelling methods adopted in this study would not be able to account for the autocorrelation of the area-level residuals (if any) of the models. Expected shortcomings due to this could be an overestimation of random effects in our models[76]. However, any such effects were observed to be very marginal in our results as the random effect estimates are already at their lower limits. While acknowledging this limitation, we believe the effects of this are not critical in our results. Hybrid models which provide more precise estimates of random effects are becoming increasingly available with advances in computational technologies[77]. However, they

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
High TC	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.20 (0.19–0.21)	0.42 (0.42–0.42)	0.20 (0.19–0.21)	0.08 (0.08–0.09)	0.22 (0.21–0.23)
Sex					
Female	Reference		–		–
Male	0.69 (0.68–0.71)		0.69 (0.68–0.71)		0.69 (0.68–0.71)
Age					
18–29	Reference		–		–
30–39	2.01 (1.90–2.13)		2.02 (1.91–2.14)		2.01 (1.90–2.13)
40–49	3.01 (2.86–3.17)		3.01 (2.86–3.17)		3.00 (2.85–3.16)
50–59	4.09 (3.89–4.30)		4.08 (3.88–4.29)		4.07 (3.87–4.28)
60–69	2.97 (2.83–3.13)		2.95 (2.80–3.10)		2.95 (2.80–3.10)
70–79	1.61 (1.53–1.70)		1.60 (1.52–1.69)		1.61 (1.52–1.70)
80 +	1.14 (1.07–1.21)		1.13 (1.07–1.20)		1.14 (1.07–1.21)
IRSD					
Q-5				Reference	–
Q-4				0.91 (0.87–0.95)	0.94 (0.90–0.98)
Q-3				0.88 (0.85–0.92)	0.94 (0.90–0.98)
Q-2				0.84 (0.80–0.87)	0.90 (0.87–0.94)
Most D Q-1				0.77 (0.74–0.81)	0.84 (0.81–0.88)
AIC	227,464	235,931.6	227,254.6	235,795.4	227,199.2
Variance		0.026	0.020	0.018	0.017
PCV		-	−21.76%	−27.81%	−33.27%
ICC (%)		0.8	0.6	0.6	0.5
MOR		1.16	1.14	1.14	1.13
AUC	0.63	0.56	0.64	0.56	0.64
AUC change[†]			+0.01		
Proportional variance explained by IRSD: 14.71%					

Table 7. Single and multilevel logistic regression model summaries for high TC (TC ≥ 5.5 mmol/L). ***$p < 0.001$; [†]Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

would not be directly applicable to our data sets, mainly due to the non-availability of location specific data at individual-level in our study data.

Notwithstanding these limitations, the study is unique in that it analysed a range of CMRFs across a widely dispersed population and included both rural and urban residents. In addition, the study used six years (year 2012–2017) of CMRF tests data from the region in the hierarchical multilevel analyses. The findings of the study indicate that those residing in the most disadvantaged areas are more likely to be identified with *higher risk* CMRFs than those in lower disadvantage areas. However, the low ICC, AUC change and MOR values of the area-level models do not support for contextual approaches. Rather, the findings of the study support a proportionate universalism approach in which health resources are made universally available but proportional to the need and disadvantage level of the affected population[57,58].

Conclusion

The study demonstrates that in the Illawarra Shoalhaven region of Australia, people residing in socioeconomically disadvantaged areas have a higher probability of being identified with *higher risk* CMRFs across a range of factors. The low general contextual effects of the areas suggest for universal intervention for the prevention and control of CMRFs in this study region, but proportional to the need and disadvantage level. The patterns were consistent across the six CMRFs analysed in this study; and comparable with similar studies reported nationally and globally. Based on our findings, we recommend further area-level research to discern the role of other contextual factors not analysed in this study especially the area-level access to health care services to determine its existing role and adequacy[78], and evidence based universal interventions for the prevention and control of CMRFs but proportionate to the priority level of the populations based on area-level disadvantage.

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
Low HDL	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.06 (0.06–0.07)	0.13 (0.13–0.13)	0.06 (0.06–0.07)	0.10 (0.09–0.10)	0.05 (0.04–0.05)
Sex					
Female	Reference		–		–
Male	3.92 (3.80–4.05)		3.98 (3.85–4.11)		3.98 (3.85–4.11)
Age					
18–29	Reference		–		–
30–39	1.11 (1.03–1.20)		1.11 (1.03–1.20)		1.12 (1.04–1.21)
40–49	0.97 (0.91–1.04)		0.99 (0.92–1.05)		1.00 (0.93–1.07)
50–59	0.87 (0.81–0.93)		0.88 (0.82–0.94)		0.89 (0.83–0.95)
60–69	0.81 (0.76–0.87)		0.82 (0.77–0.88)		0.82 (0.77–0.88)
70–79	0.85 (0.80–0.91)		0.86 (0.80–0.92)		0.85 (0.79–0.91)
80+	0.94 (0.87–1.01)		0.93 (0.86–1.00)		0.91 (0.85–0.98)
IRSD					
Q-5				Reference	–
Q-4				1.18 (1.11–1.26)	1.20 (1.13–1.28)
Q-3				1.29 (1.21–1.37)	1.32 (1.24–1.41)
Q-2				1.48 (1.39–1.57)	1.51 (1.42–1.61)
Most D Q-1				1.81 (1.71–1.92)	1.90 (1.78–2.02)
AIC	123,277	130,649.7	122,700.0	130,294.3	122,328.3
Variance		0.071	0.081	0.030	0.034
PCV		–	+15.25%	−58.05%	−51.37%
ICC (%)		2.1	2.4	0.9	1.0
MOR		1.29	1.31	1.18	1.19
AUC	0.67	0.60	0.71	0.59	0.70
AUC change†			+0.04		
Proportional variance explained by IRSD: 57.8%					

Table 8. Single and multilevel logistic regression model summaries for low HDL (HDL < 1 mmol/L). ***$p < 0.001$; †Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
High ACR	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.02 (0.02–0.03)	0.04 (0.04–0.04)	0.02 (0.02–0.03)	0.03 (0.02–0.03)	0.02 (0.01–0.02)
Sex					
Female	Reference	–	–		–
Male	1.74 (1.59–1.91)		1.75 (1.59–1.92)		1.76 (1.60–1.93)
Age					
18–29	Reference		–		–
30–39	0.99 (0.69–1.45)		1.00 (0.69–1.45)		0.99 (0.68–1.44)
40–49	0.68 (0.49–0.98)		0.69 (0.49–0.97)		0.70 (0.49–0.98)
50–59	0.76 (0.56–1.06)		0.77 (0.56–1.05)		0.77 (0.56–1.05)
60–69	0.95 (0.70–1.30)		0.95 (0.70–1.30)		0.95 (0.70–1.29)
70–79	1.54 (1.15–2.11)		1.55 (1.14–2.09)		1.52 (1.12–2.05)
80+	2.73 (2.04–3.74)		2.74 (2.02–3.71)		2.65 (1.96–3.59)
IRSD					
Q-5				Reference	–
Q-4				1.31 (1.10–1.57)	1.25 (1.04–1.50)
Q-3				1.39 (1.16–1.65)	1.27 (1.06–1.51)
Q-2				1.61 (1.36–1.90)	1.45 (1.23–1.72)
Most D Q-1				2.02 (1.72–2.38)	1.84 (1.56–2.16)
AIC	16,596	17,130.0	16,585.2	17,053.0	16,527.2
Variance		0.092	0.073	0.044	0.036
PCV		–	−20.53%	−52.88%	−61.19%
ICC (%)		2.7	2.2	1.3	1.1
MOR		1.34	1.30	1.22	1.20
AUC	0.65	0.70	0.69	0.62	0.67
AUC change†			+0.04		
Proportional variance explained by IRSD: 51.17%					

Table 9. Single and multilevel logistic regression model summaries for high ACR (ACR ≥ 30 mcg/L to mg/L). ***$p < 0.001$; †Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
Low eGFR	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.00 (0.00–0.00)	0.13 (0.12–0.13)	0.00 (0.00–0.00)	0.08 (0.07–0.08)	0.00 (0.00–0.00)
Sex					
Female	Reference		–		–
Male	0.98 (0.95–1.01)		0.98 (0.95–1.01)		0.98 (0.95–1.01)
Age					
18–29	Reference		–		–
30–39	1.66 (1.23–2.25)		1.66 (1.24–2.20)		1.66 (1.24–2.22)
40–49	4.26 (3.32–5.54)		4.26 (3.34–5.42)		4.30 (3.36–5.49)
50–59	12.24 (9.72–15.68)		12.26 (9.78–15.35)		12.32 (9.80–15.47)
60–69	41.72 (33.30–53.19)		41.81 (33.55–52.1)		41.86 (33.49–52.31)
70–79	150.44 (120.24–191.57)		150.66 (121–187.6)		149.53 (120–186.6)
80+	506.80 (405.05–645.35)		509.18 (409–633.9)		501.47 (401.7–626)
IRSD					
Q-5				Reference	–
Q-4				1.23 (1.12–1.35)	1.09 (1.03–1.16)
Q-3				1.59 (1.45–1.74)	1.19 (1.13–1.26)
Q-2				1.65 (1.51–1.81)	1.22 (1.15–1.29)
Most D Q-1				1.97 (1.80–2.15)	1.38 (1.31–1.46)
AIC	115,340	167,164.8	115,257.1	166,930.0	115,125.7
Variance		0.189	0.024	0.138	0.014
PCV		–	−87.26%	−26.84%	−92.79%
ICC (%)		5.4	0.7	4.0	0.4
MOR		1.51	1.16	1.43	1.12
AUC	0.88	0.64	0.89	0.63	0.88
AUC change†			+0.01		
Proportional variance explained by IRSD: 41.75%					

Table 10. Single and multilevel logistic regression model summaries for low eGFR (eGFR < 60 mL/min/1.73 m^2). ***$p < 0.001$; †Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

	Single level model	Multilevel models			
	Model 0	Model 1	Model 2	Model 3	Model 4
Significance (LRT)	***	***	***	***	***
Obesity	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Intercept	0.25 (0.24–0.25)	0.51 (0.50–0.51)	0.25 (0.24–0.26)	0.37 (0.35–0.39)	0.18 (0.17–0.19)
Sex					
Female	Reference		–		–
Male	0.99 (0.97–1.00)		0.99 (0.97–1.01)		0.99 (0.97–1.01)
Age					
18–29	Reference		–		–
30–39	1.64 (1.57–1.71)		1.63 (1.56–1.71)		1.64 (1.57–1.71)
40–49	2.18 (2.10–2.27)		2.20 (2.11–2.29)		2.21 (2.12–2.30)
50–59	2.41 (2.32–2.50)		2.44 (2.34–2.53)		2.44 (2.35–2.54)
60–69	2.71 (2.61–2.82)		2.73 (2.63–2.84)		2.74 (2.63–2.84)
70–79	2.47 (2.37–2.57)		2.44 (2.34–2.54)		2.43 (2.33–2.53)
80+	1.50 (1.42–1.59)		1.46 (1.38–1.55)		1.45 (1.37–1.53)
IRSD					
Q-5				Reference	-
Q-4				1.25 (1.17–1.33)	1.26 (1.18–1.34)
Q-3				1.37 (1.29–1.46)	1.38 (1.30–1.47)
Q-2				1.51 (1.42–1.61)	1.54 (1.44–1.64)
Most D Q-1				1.90 (1.79–2.03)	1.94 (1.83–2.07)
AIC	242,064	242,793.2	239,122.6	242,443.7	238,748.4
Variance		0.115	0.117	0.071	0.069
PCV		–	+1.48%	−38.76%	−40.30%
ICC (%)		3.4	3.4	2.1	2.0
MOR		1.38	1.39	1.29	1.28
AUC	0.56	0.60	0.63	0.60	0.62
AUC change†			+0.07		
Proportional variance explained by IRSD: 41.06%					

Table 11. Single and multilevel logistic regression model summaries for obesity (BMI ≥ 30 kg/m^2). ***$p < 0.001$; †Change in Model 2 in relation to Model 0; Model 0—Single level model adjusted for age + sex; Model 1—Single level model adjusted for age + sex; Model 1—null model at SA1 level; Model 2—M1 + individual-level: age + sex; Model 3—Model 1 + Area level: IRSD quintiles of SA1s; Model 4—Model 1 + Model 2 + Model 3.

	FBSL	HbA1c	TC	HDL	ACR	eGFR	Obesity
Model 0: Single level model, adjusted for age and sex							
AIC	103,645	49,897	227,464	123,277	16,596	115,340	242,064
AUC	0.70	0.56	0.63	0.67	0.65	0.88	0.56
Model 1: Null Model, at level 1							
AIC	111,022.8	50,114.5	235,931.6	130,649.7	17,130.0	167,164.8	242,793.2
τ^2	0.101	0.103	0.026	0.071	0.092	0.189	0.115
ICC (%)	3.0	3.0	0.8	2.1	2.7	5.4	3.4
MOR	1.35	1.36	1.16	1.29	1.34	1.51	1.38
AUC	0.61	0.63	0.56	0.60	0.70	0.64	0.60
Model 2: Age and sex adjusted model, at level 1							
AIC	103,066.2	49,690.2	227,254.6	122,700.0	16,585.2	115,257.1	239,122.6
τ^2	0.103	0.106	0.020	0.081	0.073	0.024	0.117
ICC (%)	3.0	3.1	0.6	2.4	2.2	0.7	3.4
MOR	1.36	1.36	1.14	1.31	1.30	1.16	1.39
AUC	0.73	0.64	0.64	0.71	0.69	0.89	0.63
AUC change†	+0.03	+0.08	+0.01	+0.04	+0.04	+0.01	+0.07
PCV	+1.88%	+3.02%	−21.76%	+15.25%	−20.53%	−87.26%	+1.48%
Model 3: IRSD adjusted model, at level 1							
AIC	110,552.5	49,875.3	235,795.4	130,294.3	17,053.0	166,930.0	242,443.7
τ^2	0.034	0.047	0.018	0.030	0.044	0.138	0.071
ICC (%)	1.0	1.4	0.6	0.9	1.3	4.0	2.1
MOR	1.19	1.23	1.14	1.18	1.22	1.43	1.29
AUC	0.60	0.61	0.56	0.59	0.62	0.63	0.60
PCV	−66.41%	−54.82%	−27.81%	−58.05%	−52.88%	−26.84%	−38.76%
Model 4: Age, sex and IRSD adjusted final model, at level 1							
AIC	102,689.6	49,453.3	227,199.2	122,328.3	16,527.2	115,125.7	238,748.4
τ^2	0.044	0.049	0.017	0.034	0.036	0.014	0.069
ICC (%)	1.3	1.5	0.5	1.0	1.1	0.4	2.0
MOR	1.22	1.24	1.13	1.19	1.20	1.12	1.28
AUC	0.72	0.63	0.64	0.70	0.67	0.88	0.62
PCV	−56.33%	−51.91%	−33.27%	−51.37%	−61.19%	−92.79%	−40.30%

Table 12. Summary model fit values and comparison of the multilevel models. †Change in relation to M0; τ^2—Variance; AIC—Akaike Information Criterion; AUC—Area under the receiver operating characteristic curve; ICC—Intra-cluster Correlation Coefficients; MOR—Median Odds Ratios; PCV—proportional change in variance (in relation to Model 1 s).

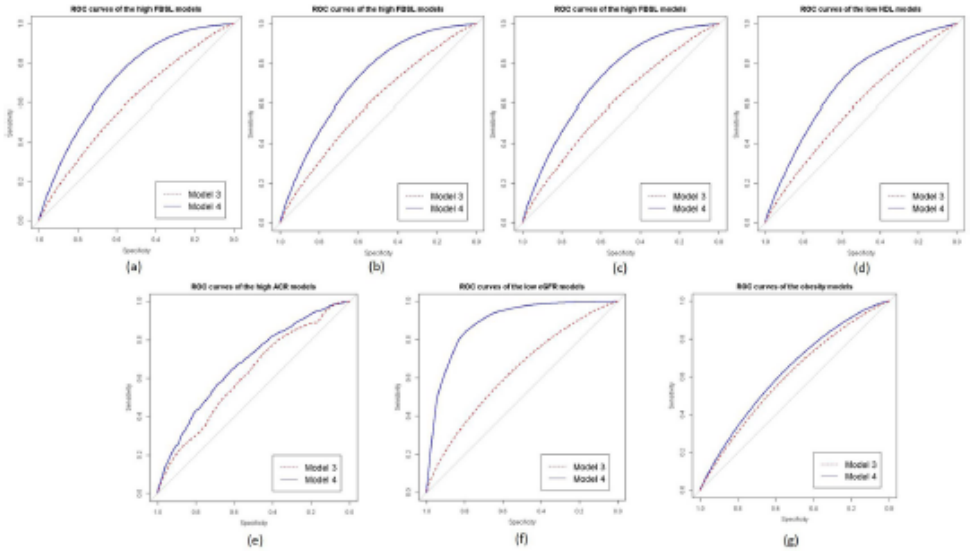

Figure 2. ROC curves of the fitted models (Model 3s and Model 4s) of CMRFs for comparison: (**a**) FBSL models; (**b**) HbA1c models; (**c**) TC models; (**d**) HDL models; (**e**) ACR models; (**f**) FBSL models; (**g**) obesity models; Model 3s—CMRFs adjusted only for IRSD quintiles of areas; Model 4s—Final models of CMRFs adjusted for age + sex + IRSD quintiles of area.

Data availability

Access to, and use of, Southern IML Research (SIMLR) Study data are subject to a License Agreement—Provision of Data (LA) between Southern IML Pathology Pty Ltd (Data Owner) and The University of Wollongong (License Holder), and a Data Access Agreement (DAA) between the License Holder and researchers (Data Users). This process is facilitated by the Illawarra Health and Medical Research Institute (IHMRI) (Data Custodian) through the Southern IML Research Study—Cohort Management Committee (SIMLR—CMC). The Data License does not allow for "public access" to data; however, researcher may access to SIMLR Study data subject to approval by the SIMLR—CMC and an appropriately constituted Australian Human Research Ethics Committee (HREC) as defined in the National Health and Medical Research Council's National Statement on Ethical Conduct in Human Research (2007) (available from https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018). The Data License requires at least one of the research team be affiliated with IHMRI. SIMLR—CMC contact details are: C/o-Associate Professor Kathryn Weston; Southern IML Research Study—Cohort Management Committee; Illawarra Health and Medical Research Institute; Building 32, University of Wollongong, Northfields Avenue, Wollongong NSW 2522, Australia; Phone +61 2 4221 4333; Email: info@ihmri.org.au; Web Link: https://www.ihmri.org.au/research-projects/simlr-cohort-study/.

Received: 27 October 2019; Accepted: 14 July 2020
Published online: 29 July 2020

References

1. Toms, R., Bonney, A., Mayne, D. J., Feng, X. & Walsan, R. Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: a systematic review of the literature. *Int. J. Health Geogr.* **18**, 1 (2019).
2. Toms, R., Mayne, D. J., Feng, X. & Bonney, A. Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia. *PLoS ONE* **14**, e0223179 (2019).
3. Alkerwi, A. *et al.* Geographic variations in cardiometabolic risk factors in Luxembourg. *Int. J. Environ. Res. Public Health* **14**, 648 (2017).
4. Andersen, A. *et al.* Life-course socio-economic position, area deprivation and Type 2 diabetes: findings from the British Women's Heart and Health Study. *Diabetic Med.* **25**, 1462–1468 (2008).
5. Astell-Burt, T., Feng, X., Kolt, G. S., McLean, M. & Maberly, G. Understanding geographical inequities in diabetes: multilevel evidence from 114,755 adults in Sydney, Australia. *Diabetes Res. Clin. Pract.* **106**, e68–e73 (2014).
6. Barker, L. E., Kirtland, K. A., Gregg, E. W., Geiss, L. S. & Thompson, T. J. Geographic distribution of diagnosed diabetes in the US: a diabetes belt. *Am. J. Prev. Med.* **40**, 434–439 (2011).
7. Bonney, A. *et al.* Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users—a cross-sectional study. *PLoS ONE* **10**, e0137261 (2015).
8. Congdon, P. Estimating diabetes prevalence by small area in England. *J. Public Health (Oxf.)* **28**, 71–81 (2006).
9. Cubbin, C. *et al.* Neighborhood deprivation and cardiovascular disease risk factors: protective and harmful effects. *Scand. J. Public Health* **34**, 228–237 (2006).

10. Dragano, N. *et al.* Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany. *BMC Public Health* **7**, 255 (2007).
11. Inoue, Y. *et al.* Neighborhood characteristics and cardiovascular risk among older people in Japan: findings from the JAGES project. *PLoS ONE* **11**, e0164525 (2016).
12. Lawlor, D., Bedford, C., Taylor, M. & Ebrahim, S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J. Epidemiol. Community Health* **57**, 134–140 (2003).
13. Lawlor, D. A., Davey Smith, G., Patel, R. & Ebrahim, S. Life-course socioeconomic position, area deprivation, and coronary heart disease: findings from the British Women's Heart and Health Study. *Am. J. Public Health* **95**, 91–97 (2005).
14. Maier, W. *et al.* Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update'GEDA 2009 and 2010. *PLoS ONE* **9**, e89661 (2014).
15. Mujahid, M. S., Diez Roux, A. V., Borrell, L. N. & Nieto, F. J. Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics. *Obes. Res.* **13**, 1412–1421 (2005).
16. Naimi, A. I., Paquet, C., Gauvin, L. & Daniel, M. Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area. *Int. J. Environ. Res. Public Health* **6**, 3082–3096 (2009).
17. Roux, A. V. D., Jacobs, D. R. & Kiefe, C. I. Neighborhood characteristics and components of the insulin resistance syndrome in young adults: the coronary artery risk development in young adults (CARDIA) study. *Diabetes Care* **25**, 1976–1982 (2002).
18. Silhol, R., Zins, M., Chauvin, P. & Chaix, B. Investigating the spatial variability in incidence of coronary heart disease in the Gazel cohort: the impact of area socioeconomic position and mediating role of risk factors. *J. Epidemiol. Community Health* **65**, 137–143 (2011).
19. Sundquist, K., Eriksson, U., Mezuk, B. & Ohlsson, H. Neighborhood walkability, deprivation and incidence of type 2 diabetes: a population-based study on 512,061 Swedish adults. *Health Place* **31**, 24–30 (2015).
20. Unger, E. *et al.* Association of neighborhood characteristics with cardiovascular health in the multi-ethnic study of atherosclerosis. *Circ. Cardiovasc. Qual. Outcomes* **7**, 524–531 (2014).
21. Valdés, S. *et al.* Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Diabetes study. *Revista Española de Cardiología (English Edition)* **67**, 442–448 (2014).
22. Zhou, M. *et al.* Geographical variation in diabetes prevalence and detection in China: multilevel spatial analysis of 98,058 adults. *Diabetes Care* **38**, 72–81 (2015).
23. Paquet, C. *et al.* Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia. *Int. J. Environ. Res. Public Health* **13**, 519 (2016).
24. World Health Organisation. *The top 10 causes of death*. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (2017).
25. Australian Institute of Health and Welfare. *Indicators of socioeconomic inequalities in cardiovascular disease, diabetes and chronic kidney disease*. https://www.aihw.gov.au/getmedia/01c5bb07-592e-432e-9fba-d242e0f7e27e/aihw-cdk-12.pdf.aspx?inline=true (2019).
26. Australian Institute of Health and Welfare. *Australian burden of disease study : impact and causes of illness and death in Australia 2011*. https://www.aihw.gov.au/getmedia/d4df9251-c4b6-452f-a877-8370b6124219/19663.pdf.aspx?inline=true (2016).
27. Australian Institute of Health and Welfare. *Cardiovascular disease, diabetes and chronic kidney disease: Australian facts: morbidity—hospital care*. https://www.aihw.gov.au/reports/heart-stroke-vascular-disease/cardiovascular-diabetes-chronic-kidney-morbidity/contents/table-of-contents (2017).
28. Australian Institute of Health and Welfare. *Australia's Health 2014*. https://www.aihw.gov.au/reports/australias-health/australias-health-2014/contents/table-of-contents (2014).
29. Clark, C. R. *et al.* Neighborhood disadvantage, neighborhood safety and cardiometabolic risk factors in African Americans: biosocial associations in the Jackson Heart study. *PLoS ONE* **8**, e63254 (2013).
30. Cox, M., Boyle, P. J., Davey, P. G., Feng, Z. & Morris, A. D. Locality deprivation and Type 2 diabetes incidence: a local test of relative inequalities. *Soc. Sci. Med.* **65**, 1953–1964 (2007).
31. Gabert, R., Thomson, B., Gakidou, E. & Roth, G. Identifying high-risk neighborhoods using electronic medical records: a population-based approach for targeting diabetes prevention and treatment interventions. *PLoS ONE* **11**, e0159227 (2016).
32. Keita, A. D. *et al.* Associations of neighborhood area level deprivation with the metabolic syndrome and inflammation among middle-and older-age adults. *BMC Public Health* **14**, 1319 (2014).
33. Larata, B. A. *et al.* Place matters: neighborhood deprivation and cardiometabolic risk factors in the Diabetes Study of Northern California (DISTANCE). *Soc. Sci. Med.* **74**, 1082–1090 (2012).
34. Mobley, L. R. *et al.* Environment, obesity, and cardiovascular disease risk in low-income women. *Am. J. Prev. Med.* **30**, 327.e321-332.e321 (2006).
35. Riva, M., Gauvin, L. & Barnett, T. A. Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998. *J. Epidemiol. Community Health* **61**, 853–861 (2007).
36. Merlo, J., Wagner, P., Ghith, N. & Leckie, G. An original stepwise multilevel logistic regression analysis of discriminatory accuracy: the case of neighbourhoods and health. *PLoS ONE* **11**, e0153778 (2016).
37. Merlo, J. *et al.* Diastolic blood pressure and area of residence: multilevel versus ecological analysis of social inequity. *J Epidemiol Community Health* **55**, 791–798 (2001).
38. Merlo, J., Viciana-Fernández, F. J. & Ramiro-Fariñas, D. Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia, Spain. *Soc. Sci. Med.* **75**, 1477–1487 (2012).
39. Australian Bureau of Statistics. *Australian Statistical Geography Standard (ASGS): Volume 1—Main Structure: Statistical Area Level 1 (SA1)*. https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+(ASGS) (2016).
40. The Royal Australian College of General Practitioners & Diabetes Australia. *General Practice Management of Type 2 Diabetes 2016–2018*. https://doi.org/10.1007/s00125-010-2011-6 (2016).
41. Australian Bureau of Statistics. *Australian Health Survey: Biomedical Results for Chronic Diseases, 2011–2012*. https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.005?2011-12 (2013).
42. National heart foundation of Australia. *Lipid management profile for health professionals*. https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management.
43. National Kidney foundation (USA). *Albumin creatinine Ratio (ACR)*. https://www.kidney.org/kidneydisease/siemens_hcp_acr (2018).
44. World Health Organization. *Obesity: Preventing and managing the global epidemic: Technical Report Series. WHO Technical Report Series, no. 894*. ISBN: 92 4 120894 5 (2000).
45. Australian Bureau of Statistics. *Main Features—IRSD*. https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011.
46. Ghosh, A., Charlton, K. E., Girdo, L. & Batterham, M. Using data from patient interactions in primary care for population level chronic disease surveillance: the Sentinel Practices Data Sourcing (SPDS) project. *BMC Public Health* **14**, 557 (2014).
47. Goldstein, H., Browne, W. & Rasbash, J. Partitioning variation in multilevel models. *Underst. Stat. Stat. Issues Psychol. Educ. Soc. Sci.* **1**, 223–231 (2002).
48. Wagner, P. & Merlo, J. Measures of discriminatory accuracy in multilevel analysis. *Eur. J. Epidemiol.* **28**, 135 (2013).

49. Pepe, M. S., Janes, H., Longton, G., Leisenring, W. & Newcomb, P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am. J. Epidemiol.* **159**, 882–890 (2004).
50. R Core Team. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* https://www.R-project.org/ (2018).
51. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823 (2014).
52. Zeileis, A. & Hothorn, T. Diagnostic checking in regression relationships. R News 2: 7–10. Accessed August 2011. https://CRAN.R-project.org/doc/Rnews/ (https://CRAN.R-project.org/doc/Rnews/) (2002).
53. Robin, X. *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
54. Australian Bureau of Statistics. *Technical Paper: Socio-Economic Indexes for Areas (SEIFA).* https://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/22CEDA8038AF7A0DCA257B3B00116E34/$File/2033.0.55.001%20seifa%202011%20technical%20paper.pdf (2011).
55. Merlo, J., Ohlsson, H., Lynch, K. F., Chaix, B. & Subramanian, S. Individual and collective bodies: using measures of variance and association in contextual epidemiology. *J. Epidemiol. Community Health* **63**, 1043–1048 (2009).
56. Merlo, J., Wagner, P. & Leckie, G. A simple multilevel approach for analysing geographical inequalities in public health reports: The case of municipality differences in obesity. *Health Place* **58**, 102145 (2019).
57. Lu, D. & Tyler, I. Focus on: a proportionate approach to priority populations. Public Health Ontario. https://www.publichealthontario.ca/en/eRepository/Focus_On_Priority_Populations.pdf. Accessed 29 (2016).
58. Marmot, M. & Bell, R. Fair society, healthy lives. *Public Health* **126**, S4–S10 (2012).
59. Rose, G. Sick individuals and sick populations. *Int. J. Epidemiol.* **30**, 427–432 (2001).
60. Barter, P. J., Brandrup-Wognsen, G., Palmer, M. K. & Nicholls, S. J. Effect of statins on HDL-C: a complex process unrelated to changes in LDL-C: analysis of the VOYAGER Database. *J. Lipid Res.* **51**, 1546–1553 (2010).
61. Mooradian, A. D. Dyslipidemia in type 2 diabetes mellitus. *Nat. Rev. Endocrinol.* **5**, 150–159 (2009).
62. Frei, B., Forte, T. M., Ames, B. N. & Cross, C. E. Gas phase oxidants of cigarette smoke induce lipid peroxidation and changes in lipoprotein properties in human blood plasma. Protective effects of ascorbic acid. *Biochem. J.* **277**, 133–138 (1991).
63. Hu, F. B. Sedentary lifestyle and risk of obesity and type 2 diabetes. *Lipids* **38**, 103–108 (2003).
64. Thorp, A. A. *et al.* Deleterious associations of sitting time and television viewing time with cardiometabolic risk biomarkers: Australian Diabetes, Obesity and Lifestyle (AusDiab) study 2004–2005. *Diabetes Care* **33**, 327–334 (2010).
65. Arai, T. *et al.* Increased plasma cholesteryl ester transfer protein in obese subjects. A possible mechanism for the reduction of serum HDL cholesterol levels in obesity. *Arterioscle. Thromb. J. Vasc. Biol.* **14**, 1129–1136 (1994).
66. McNaughton, S. A., Dunstan, D. W., Ball, K., Shaw, J. & Crawford, D. Dietary quality is associated with diabetes and cardiometabolic risk factors. *J. Nutr.* **139**, 734–742 (2009).
67. Williams, E. D. *et al.* Health behaviours, socioeconomic status and diabetes incidence: the Australian Diabetes Obesity and Lifestyle Study (AusDiab). *Diabetologia* **53**, 2538–2545 (2010).
68. Sodjinou, R., Agueh, V., Fayomi, B. & Delisle, H. Obesity and cardio-metabolic risk factors in urban adults of Benin: relationship with socio-economic status, urbanisation, and lifestyle patterns. *BMC Public Health* **8**, 84 (2008).
69. Fraser, L. K., Edwards, K. L., Cade, J. & Clarke, G. P. The geography of fast food outlets: a review. *Int. J. Environ. Res. Public Health* **7**, 2290–2308 (2010).
70. Macdonald, L., Cummins, S. & Macintyre, S. Neighbourhood fast food environment and area deprivation—substitution or concentration?. *Appetite* **49**, 251–254 (2007).
71. Pearce, J., Blakely, T., Witten, K. & Bartie, P. Neighborhood deprivation and access to fast-food retailing: a national study. *Am. J. Prev. Med.* **32**, 375–382 (2007).
72. Walsan, R., Bonney, A., Mayne, D. J., Pat, N., Feng, X. & Toms, R. Serious mental illness, neighborhood disadvantage, and type 2 diabetes risk: a systematic review of the literature. *J. Prim. Care. Community Health* **9**, 2150132718802025 (2018).
73. Buttar, H. S., Li, T. & Ravi, N. Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation. *Exp. Clin. Cardiol.* **10**, 229 (2005).
74. Fiuza-Luces, C. *et al.* Exercise benefits in cardiovascular disease: beyond attenuation of traditional risk factors. *Nat. Rev. Cardiol.* **15**, 731–743 (2018).
75. Chomistek, A. K. *et al.* Healthy lifestyle in the primordial prevention of cardiovascular disease among young women. *J. Am. Coll. Cardiol.* **65**, 43–51 (2015).
76. Xu, H. Comparing spatial and multilevel regression models for binary outcomes in neighborhood studies. *Sociol. Methodol.* **44**, 229–272 (2014).
77. Chaix, B., Merlo, J. & Chauvin, P. Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilisation in France. *J. Epidemiol. Community Health* **59**, 517–526 (2005).
78. Toms, R., Feng, X., Mayne, D. J. & Bonney, A. Role of area-level access to primary care on the geographic variation of cardiometabolic risk factor distribution: a multilevel analysis of the adult residents in the Illawarra—Shoalhaven Region of NSW, Australia. *Int. J. Environ. Res. Public Health* **17**, 4297 (2020).

Acknowledgements

We would like to thank Southern IML Pathology and staff for their generosity in providing data for the SIMLR Cohort Study and ongoing support. In particular we would like to thank Mr Bryan Jones for providing the technical expertise for the data acquisition and helpful comments on the manuscript. Southern IML Pathology are the owners of the data contained within this publication and the Illawarra Health and Medical Research Institute (IHMRI) is the custodian facilitating access to the data. https://www.ihmri.org.au/research-projects/simlr-cohort-study/

Author contributions

R.T. contributed to the study conceptualization, data curation, statistical analyses, and writing of the original draft; D.M. contributed to the study conceptualization, methodology and co-supervision of the project; X.F. contributed to the study methodology and co-supervision of the project; and A.B. contributed to the overall project administration, resources, methodology, and main supervision of the project and study validation. All the authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Material: 4 (Published version of research article 4)

International Journal of
Environmental Research and Public Health

MDPI

Article

Role of Area-Level Access to Primary Care on the Geographic Variation of Cardiometabolic Risk Factor Distribution: A Multilevel Analysis of the Adult Residents in the Illawarra—Shoalhaven Region of NSW, Australia

Renin Toms [1,2,*], Xiaoqi Feng [2,3,4], Darren J Mayne [1,2,5,6] and Andrew Bonney [1,2]

1 School of Medicine, University of Wollongong, Wollongong NSW 2522, Australia; dmay8519@uni.sydney.edu.au (D.J.M.); abonney@uow.edu.au (A.B.)
2 Illawarra Health and Medical Research Institute, Wollongong NSW 2522, Australia; xiaoqi.feng@unsw.edu.au
3 Population Wellbeing and Environment Research Lab (PowerLab), School of Health and Society, Faculty of Social Sciences, University of Wollongong, Wollongong NSW 2500, Australia
4 School of Public Health and Community Medicine, University of New South Wales, Sydney NSW 2033, Australia
5 Illawarra Shoalhaven Local Health District, Public Health Unit, Warrawong NSW 2502, Australia
6 School of Public Health, The University of Sydney, Sydney NSW 2006, Australia
* Correspondence: rmbst288@uowmail.edu.au

Received: 18 April 2020; Accepted: 14 June 2020; Published: 16 June 2020

check for updates

Abstract: *Background:* Access to primary care is important for the identification, control and management of cardiometabolic risk factors (CMRFs). This study investigated whether differences in geographic access to primary care explained area-level variation in CMRFs. *Methods:* Multilevel logistic regression models were used to derive the association between area-level access to primary care and seven discrete CMRFs after adjusting for individual and area-level co-variates. Two-step floating catchment area method was used to calculate the geographic access to primary care for the small areas within the study region. *Results:* Geographic access to primary care was inversely associated with low high density lipoprotein (OR 0.94, CI 0.91–0.96) and obesity (OR 0.91, CI 0.88–0.93), after adjusting for age, sex and area-level disadvantage. The intra-cluster correlation coefficient (ICCs) of all the fully adjusted models ranged between 0.4–1.8%, indicating low general contextual effects of the areas on CMRF distribution. The area-level variation in CMRFs explained by primary care access was ≤10.5%. *Conclusion:* The findings of the study support proportionate universal interventions for the prevention and control of CMRFs, rather than any area specific interventions based on their primary care access, as the contextual influence of areas on all the analysed CMRFs were found to be minimal. The findings also call for future research that includes other aspects of primary care access, such as road-network access, financial affordability and individual-level acceptance of the services in order to gain an overall picture of the area-level contributing role of primary care on CMRFs in the study region.

Keywords: geographic access; cardiometabolic risk factor; geographic variation; multilevel logistic regression; primary care access

1. Introduction

Cardiometabolic risk factors (CMRFs) demonstrate significant variation in geographic distribution within countries globally [1–10]. Higher prevalence and clustering of CMRFs is often reported

for socioeconomically disadvantaged areas [11–20]. Reachability or geographic access to primary care is essential for the individual-level identification and management of CMRFs, especially when considering their chronic nature after detection [21–23]. Therefore, access to primary care may be associated with the geographic variation of CMRFs [24].

Previous studies have reported that access to primary care can play a role in the control and management of certain CMRFs [21,25–28]. The dimensions of access to primary care can be fundamentally conceptualized into (1) physical access (2) affordability and (3) acceptability [29]. Research indicates that the physical access to primary care varies across areas, as the locations of primary care physicians and services tend to be positively correlated with population density [30,31]. There is also evidence that medical consultations were reported less likely to happen when physical access to health care services is lower [21]. In addition, access to adequate treatment facilities were reported to have an inverse association with certain CMRFs, such as hypertension [25,26], end stage renal disease [27] and diabetes mellitus [32]. However, these reports are based on individual CMRFs but consistent evidence across a range of CMRFs may provide a stronger evidence base for healthcare service commissioning across areas.

Evidence regarding the association of CMRFs with primary care access over and above area-level disadvantage may also inform area-level resource allocation of primary care services in disadvantaged areas [24,33]. Therefore, the aims of this study were to: (1) identify the area-level association of individual CMRFs with geographic access to primary care; (2) quantify the general contextual effect of areas on CMRFs; and (3) quantify the geographic variation in CMRFs explained by differences in area-level primary care access, within the Illawarra-Shoalhaven region of Australia.

2. Materials and Methods

A retrospective cross-sectional design and multilevel logistic regression models were adopted to meet the study objectives. The study was approved by the University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical Human Research Ethics Committee (HREC protocol No: 2017/124).

The study focused on the Illawarra-Shoalhaven region of New South Wales (NSW), Australia. This area consists of multiple regional cities, smaller towns and rural areas, including the local government areas of Kiama, Shellharbour, Shoalhaven and Wollongong. The region covers a geographical area of 5615 square kilometres and had a population of 369,469 people at the 2011 Australian Census of Population and Housing [34,35]. The geographic unit of analysis used in this study was the Statistical Area 1 (SA1), which is the smallest statistical output unit of the 2011 Census and which has an average population of 400 people (range: 200 to 800) [36]. The study area encompassed 980 conterminous SA1s [37]. Figure 1 shows the study area showing SA1s and major landmarks of the region.

2.1. Data

The study used three different databases: (1) the CMRF pathology test data; (2) primary care provider data; and (3) the estimated resident populations from the 2011 Australian Census of Population and Housing. The CMRF test data were extracted from the Southern IML Research (SIMLR) Study database. The SIMLR Study database comprises de-identified and internally linked pathology results from a major pathology provider in the study region and provides near-census coverage of the study population [8]. The CMRF test data were extracted for multiple risk factors on the most recent test results, of non-pregnant adults aged 18 years and over, undergoing a laboratory test between 1 January 2012 and 31 December 2017.

The primary care provider data were manually extracted in 2016 from publicly available data sources, including Yellow Pages, White Pages, online general practitioner (GP) appointment booking services and Google search results. The 2011 Australian Census of Population and Housing data were accessed to extract the population denominator data of the study region at SA1 level [34].

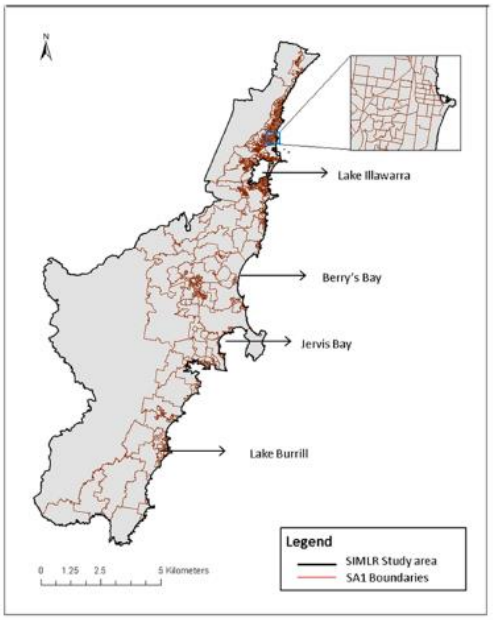

Figure 1. Map of the Illawarra-Shoalhaven region of NSW, Australia, showing SA1 areas and major landmarks.

2.2. Dependent Variable

Dichotomised results of the individual CMRF tests were the dependent variables in this study. The CMRF test results included: fasting blood sugar level (FBSL); glycated haemoglobin (HbA1c); total cholesterol (TC); high density lipoprotein (HDL); urinary albumin creatinine ratio (ACR); estimated glomerular filtration rate (eGFR); and objectively measured body mass index (BMI). The pathology service routinely collects BMI for each of the remaining CMRF tests and thus became available for analyses in this study [12]. However, it should be noted that the data do not include blood pressure readings. Although blood pressure is an important CMRF it is not routinely collected with any of the pathology test samples and thus not available for analyses in this study. During analyses, all the retrieved CMRF test results were dichotomised into higher risk and lower risk values based on established risk classification guideline values. Table 1 shows the CMRF definitions adopted in this study to dichotomies the test results.

Table 1. Definitions of CMRF (cardiometabolic risk factor) test results.

	Higher Risk CMRFs	Definition
1.	High FBSL	FBSL ≥ 7.0 mmol/L [38].
2.	High HbA1c	HbA1c > 7.5% [38].
3.	High TC	TC ≥ 5.5 mmol/L [39].
4.	Low HDL	HDL < 1 mmol/L [40].
5.	High ACR	ACR ≥ 30 mcg/L to mg/L [41].
6.	Low eGFR	eGFR < 60 mL/min/1.73 m^2 [41].
7.	Obesity	BMI ≥ 30 kg/m^2 [42].

CMRFs—cardiometabolic risk factors; FBSL—fasting blood sugar level; HbA1c—glycated haemoglobin; TC—total cholesterol; HDL—high density lipoprotein; ACR—albumin creatinine ratio; eGFR—estimated glomerular filtration rate; BMI—body mass index.

2.3. Independent Variable

Primary care access index calculated for the small areas at SA1-level was the independent study variable. An access index score was calculated for each SA1 using a two-step floating catchment area (2SFCA) method, which balanced both supply and demand of primary care services in the study region.

The 2SFCA method was developed by Luo and Wang in 2003 to measure geographic accessibility of health care services [43]. The method has undergone several enhancements since its inception but essentially consists of two steps underpinned by a gravity model [43,44]. The first step computes a population-to-provider ratio for each primary care service location by aggregating the population size of the SA1s whose centroids (i.e., the geometric centers) are located within a defined spatial buffer distance [45]. The total number of general practitioners working in the primary care service locations within this buffer distance were the numerators for the provider-to-population ratio calculations.

Thus, step 1:

$$R_j = S_j / \sum_i p_i \quad (1)$$

where Sj is the number of general practitioners at location j, pi is the number of adult residents in the SA1s (those SA1 geographic centroids are located within the spatial buffer distance of the primary care locations) and Rj is the population-to-provider ratio for service j [45].

In step 2, a population-to-provider ratio (access score) is computed for each geographic centroid of the SA1s by aggregating all primary care service population-to provider ratios of the primary care services that are located within the same spatial buffer distance [45].

Thus, step 2:

$$A_i = \sum_j R_j \quad (2)$$

where A_i is the access index for population location i.

The resulting access indices were retained as a continuous variable for the analyses. A higher score indicated better geographic access of the SA1s to primary care services.

A spatial radial buffer distance of 30 km was chosen to compute primary care access for SA1s in the study region. In the preliminary stage, sensitivity analyses were performed using 1 km, 16 km and 30 km spatial radial buffer distances. In step 1 2SFCA analyses, the 1 km distance covered only 545 (56%) SA1 centroids in the study region in relation to the primary care provider locations, whereas a 16 km radial buffer distance covered 973 (~99%) and a 30 km radial buffer distance covered 978 (~100%) SA1s' geographic centroids. Therefore, a radial buffer distance of 30 km was chosen to determine the access which was observed to cover the mixed rural, semi-rural and urban distribution of the population in the study region well.

2.4. Covariates

The individual-level variables adjusted at SA1-level were: sex (male and female) and age group (18–29, 30–39, 40–49, 50–59, 60–69, 70–79 and 80+ years). The area-level covariate adjusted at SA1-level was the area-level socioeconomic disadvantage. The Index of Relative Socioeconomic Disadvantage (IRSD) score of the SA1s in the study region based on the 2011 Australian Bureau of Statistics conducted census of population and housing was used as the measurement variable for the area-level socioeconomic disadvantage of the SA1s [37]. The IRSD summarises a range of measures of relative socioeconomic disadvantage of people and/or households within SA1s and includes: level of income; education; employment; family structure; disability; housing; transportation; and internet connection [37]. A higher IRSD score indicated lower levels of disadvantage [37]. The Illawarra-Shoalhaven region has a diverse socio-economic profile, making this landscape useful for area-level population health studies [46].

2.5. Statistical Analyses

Multilevel logistic regression models were fitted to individual CMRF test data at the SA1 level. For each of the seven CMRFs analysed in this study, five nested models were fit that included fixed effects for access index after adjusting for sex, age and IRSD score; and random effect intercepts for SA1s. In the nested models, Model 1 (M1) was a null model of CMRF at SA1-level; Model 2 (M2) included the area-level study variable (access index) only; Model 3 (M3) included individual-level factors at SA1-level (age and sex) only; Model 4 (M4) included individual and area-level factors (age, sex and IRSD score) at SA1-level; and Model 5 (M5) included M4 variables plus access index. Thus, the final model (M5) estimated the effect of primary care access after adjusting for individual and area-level factors. Odds ratios (ORs) were derived from the exponentials of regression coefficients from fitted models. As the IRSD scores and access index of the SA1s were fitted as mean-centred continuous variables, ORs were expressed per standard deviation unit change in these variables. Statistical significance of the models was evaluated using likelihood ratio tests and a type I error rate of 0.05.

2.6. Model Comparison

Model fit was compared using the Akaike Information Criterion (AIC). The models were also evaluated for: area-level variance (τ^2); proportional change in variance (PCV) in comparison with the null model; intra-cluster correlation coefficient (ICC) of the model; and the median odds ratios (MORs). The ICC and MOR of the models were used to index the between-area variability. A latent variable approach was used to derive the ICC of models [47]. The MOR translates the area-level variance into an easily interpretable OR and is assumed to be statistically independent of the test specific prevalence of the CMRFs [48]. The unique contribution of the primary care access of the SA1s to the area-level variance of CMRF was estimated through the reduction in PCV between M4 and M5.

2.7. Statistical Package

All mapping and geospatial measurements were performed using ArcGIS version 10.4.1 (ESRI Inc. Redlands, CA, USA) [49]. All statistical analyses were performed using R version 3.4.4. (R Foundation for Statistical Computing, Vienna, Austria) [50]. Multilevel models were fit using the glmer function in the lme4 package [51]; and likelihood ratio tests were calculated using the lrtest function in the lmtest package [52]. The glmer function fit the generalized linear mixed model, which incorporates both fixed-effects parameters and random effects in a linear predictor, via maximum likelihood [53,54].

3. Results

A total of 1,132,029 CMRF test results for 256,525 individual residents in the Illawarra-Shoalhaven region between 2012 and 2017 were extracted for analysis. The mean number of tests undertaken per person was 4.4 (SD = 1.8, range = 1–7). After excluding 1162 (1.0%) test results with incomplete details, a total of 1,130,894 tests were retained in the final data set. IRSD scores of the SA1s were the most frequent missing variable, as this was not available for some SA1s in the study region [55]. Available IRSD scores ranged between 446.7 and 1143.7 (mean = 976.7, SD = 98.6) for SA1s, with a higher score indicating lower area-level disadvantage. Table 2 details the individual-level CMRFs risk proportions of the final data set.

For primary care access, a total of 165 primary care service locations with 611 general practitioners were identified in the study area in 2016. The primary care access index of the SA1s in the study region ranged between 0 and 5.41 general practitioners per 1000 people (mean = 2.1, SD = 0.77). Figure 2 illustrates the distribution of the primary care access index within the study region.

Table 2. Frequency and proportion of CMRFs risk classification with gender.

	Cardiometabolic Risk	Test *n*	Higher Risk *n* (%) *	Male *n* (%) *	Female *n* (%) *
1.	High FBSL	193,679	16,280 (8.4)	9289 (4.8)	6991 (3 .6)
2.	High HbA1c	73,885	7927 (10.7)	4448 (6.0)	3479 (4.7)
3.	High TC	194,816	63,422 (32.6)	26,139 (13.4)	37,283 (19.1)
4.	Low HDL	182,237	21,261 (11.7)	15,885 (8.7)	5376 (3.0)
5.	High ACR	50,790	2047 (4.0)	1266 (2.5)	781 (1.5)
6.	Low eGFR	244,166	27,241 (11.2)	12,456 (5.1)	14,785 (6.1)
7.	Obesity	192,455	64,832 (33.7)	29,613 (15.4)	35,319 (18.4)

* The denominators of the percentages are the total number of each CMRF tests.

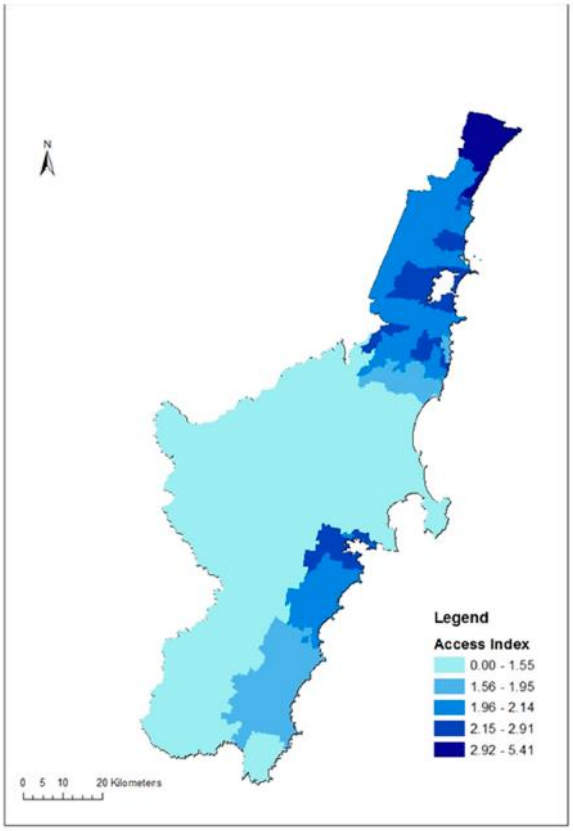

Figure 2. Geographic access to primary care services in the Illawarra-Shoalhaven region of the NSW, Australia.

Multilevel logistic regression models for each CMRF are presented in Tables 3–9. The null models indicated significant geographic variation in the distribution of all CMRFs at the SA1 level. Model 2s showed inverse associations between access index and all CMRFs except TC, which displayed a positive association with the access index. Model 3s adjusted CMRF models for individual-level age

and sex, which accounted for 1.5% (obesity) to 87.3% (eGFR) of unexplained variation in the null model. The general contextual effect of areas over and above their individual composition, such as age and sex, is obtained by a measure of clustering (i.e., ICCs) in the model 3s, which ranged between 0.6–3.4% in the CMRFs models presented. Model 4s demonstrated significant inverse associations between area-level socioeconomic disadvantage and all CMRFs except for TC after adjusting for individual-level factors. Total cholesterol again showed a positive association with area-level socioeconomic disadvantage. In the final models (M5s), the access index was found to be inversely associated with low HDL (HDL < 1 mmol/L) and obesity (BMI ≥ 30 kg/m^2), after adjusting for individual and area-level factors. Including the access index in the final models did not attenuate associations between area-level disadvantage and CMRFs observed in M4s.

Model fit of the nested models of each CMRF were compared using the Akaike Information Criterion (AIC). The AIC estimated the out-of-sample prediction error rates of individual models and thus the relative quality of individual models for a given set of nested models [56]. Reductions in the AIC values were observed for all CMRFs models, except in TC and eGFR, from the null model (M1) to the final model (M5), indicating a better fit of the final models. The AIC for TC and eGFR models indicated M4 was the best fitting model for these CMRFs.

In the null models (M1s), low eGFR demonstrated the most area-level variance and high TC showed the least. The access only models (M2s) showed a reduction in the residual variance of all CMRFs from those of null models. In Model 3s, adjusting for age and sex initially increased the residual variance of FBSL (PCV = +1.9%), HbA1c (PCV = +3.0%), HDL (PCV = +15.3%) and BMI (PCV = +1.5%). In Model 4s, adjusting the CMRFs for individual-level age and sex and area-level disadvantage resulted in major reductions of variance from −33.1% (in TC) to −93.3% (in eGFR). In the final models (M5s), including access index in the models after adjusting for the covariates extended the reduction in variance in all CMRFs, except for TC and eGFR. Including the access index had been observed to increase the variance in the TC and eGFR final models, compared with the lower level model.

Similarly, in the unadjusted models, the MORs, which indicate the odds of having a higher risk CMRF test result for a person from the most, compared to the least, area-level disadvantage, were the highest among eGFR (τ^2 = 0.189; ICC = 5.4%; MOR = 1.51) and the least among TC (τ^2 = 0.025; ICC = 0.8%; MOR = 1.16). The ICCs of CMRFs in all the models were comparatively small (Table 4) in all the models, indicating minimal contextual effect of areas on any of the CMRFs. In the fully adjusted models, the ICCs further reduced and ranged between 0.4% and 1.8% in low eGFR and BMI respectively. Table 10 presents a summary and comparison of the model fit.

Table 3. Multilevel logistic regression model summaries of high FBSL (FBSL ≥ 7.0 mmol/L).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**	
	OR (95% CI)	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**
High FBSL										
Intercept	0.09 (0.09–0.09)	$p < 0.001$	0.09 (0.09–0.09)	$p < 0.001$	0.01 (0.01–0.01)	$p < 0.001$	0.01 (0.01–0.01)	$p < 0.001$	0.01 (0.01–0.01)	$p < 0.001$
Access			0.89 (0.87–0.92)	$p < 0.001$						
Sex: Female					Reference					
Male					1.63 (1.58–1.69)	$p < 0.001$	1.63 (1.58–1.69)	$p < 0.001$	1.63 (1.58–1.69)	$p < 0.001$
Age: 18–29					Reference					
30–39					1.63 (1.40–1.90)	$p < 0.001$	1.65 (1.41–1.92)	$p < 0.001$	1.65 (1.41–1.92)	$p < 0.001$
40–49					3.53 (3.08–4.05)	$p < 0.001$	3.57 (3.11–4.10)	$p < 0.001$	3.57 (3.11–4.10)	$p < 0.001$
50–59					6.77 (5.93–7.72)	$p < 0.001$	6.81 (5.97–7.77)	$p < 0.001$	6.80 (5.97–7.75)	$p < 0.001$
60–69					11.07 (9.72–12.6)	$p < 0.001$	11.07 (9.7–12.6)	$p < 0.001$	11.05 (9.7–12.6)	$p < 0.001$
70–79					13.93 (12.2–15.9)	$p < 0.001$	13.8 (12.1–15.7)	$p < 0.001$	13.8 (12.1–15.7)	$p < 0.001$
80+					12.33 (10.8–14.1)	$p < 0.001$	12.1(10.6–13.9)	$p < 0.001$	12.1(10.6–13.8)	$p < 0.001$
IRSD							0.79 (0.77–0.80)	$p < 0.001$	0.79 (0.77–0.81)	$p < 0.001$
Access									0.98 (0.96–1.00)	0.111
AIC	111,022.8		110,962.3		103,066.2		102,652.6		102,652.0	
Variance	0.101		0.091		0.103		0.040		0.039	
PCV	-		−9.98%		+1.88%		−60.90%		−61.05%	
ICC (%)	3.0		2.7		3.0		1.2		1.2	
MOR	1.36		1.334		1.36		1.209		1.209	
Proportional variance explained by Access to primary care: −0.38%										

AIC—Akaike Information Criterion; FBSL—fasting blood sugar level; ICC—Intra-cluster correlation coefficient; IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 4. Multilevel logistic regression model summaries of high HbA1c (HbA1c > 7.5%).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**	
	OR (95% CI)	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**
High HbA1c										
Intercept	0.12 (0.11–0.12)	$p < 0.001$	0.12 (0.11–0.12)	$p < 0.001$	0.07 (0.06–0.07)	$p < 0.001$	0.07 (0.06–0.08)	$p < 0.001$	0.07 (0.06–0.08)	$p < 0.01$
Access			0.95 (0.92–0.98)	$p < 0.001$						
Sex: Female					Reference					
Male					1.38 (1.3–1.45)	$p < 0.001$	1.39 (1.32–1.45)	$p < 0.001$	1.39 (1.32–1. 45)	$p < 0.001$
Age: 18–29					Reference					
30–39					0.81 (0.68–0.96)	$p < 0.01$	0.81 (0.68–0.96)	$p < 0.01$	0.81 (0.68–0. 97)	$p < 0.01$
40–49					1.24 (1.07–1.44)	$p < 0.001$	1.25 (1.08–1.45)	$p < 0.001$	1.26 (1.08–1. 46)	$p < 0.001$
50–59					1.56 (1.36–1.80)	$p < 0.001$	1.56 (1.36–1.80)	$p < 0.001$	1.57 (1.36–1. 81)	$p < 0.001$
60–69					1.64 (1.43–1.88)	$p < 0.001$	1.64 (1.43–1.88)	$p < 0.001$	1.64 (1.43–1. 89)	$p < 0.001$
70–79					1.64 (1.42–1.88)	$p < 0.001$	1.62 (1.41–1.86)	$p < 0.001$	1.63 (1.42–1. 87)	$p < 0.001$
80+					1.63 (1.41–1.88)	$p < 0.001$	1.62 (1.40–1.87)	$p < 0.001$	1.62 (1.41–1. 87)	$p < 0.001$
IRSD							0.79 (0.77–0.81)	$p < 0.001$	0.79 (0.77–0. 81)	$p < 0.001$
Access									1.00 (0.97–1.03)	0.750
AIC	50,114.5		50,105.9		49,690.2		49,438.2		49,440.0	
Variance	0.103		0.100		0.106		0.048		0.047	
PCV	-		−2.430%		+3.02%		−53.78%		−53.80%	
ICC (%)	3.0		3.0		3.1		1.4		1.4	
MOR	1.36		1.353		1.358		1.231		1.231	
Proportional variance explained by Access to primary care: −0.04%										

AIC—Akaike Information Criterion; HbA1c—glycated haemoglobin; ICC—Intra-cluster correlation coefficient; IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 5. Multilevel logistic regression model summaries of high TC (TC ≥ 5.5 mmol/L).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**	
	OR (95% CI)	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**
high TC										
Intercept	0.42 (0.41–0.43)	$p < 0.001$	0.42 (0.41- 0.43)	$p < 0.001$	0.20 (0.19–0.21)	$p < 0.001$	0.20 (0.19–0.21)	$p < 0.001$	0.20 (0.19–0.21)	$p < 0.01$
Access			1.02 (1.00–1.03)	$p < 0.01$						
Sex: Female					Reference					
Male					0.69 (0.68–0.71)	$p < 0.001$	0.69 (0.68–0.71)	$p < 0.001$	0.69 (0.68–0.71)	$p < 0.001$
Age: 18–29					Reference					
30–39					2.02 (1.91–2.14)	$p < 0.001$	2.01 (1.90–2.13)	$p < 0.01$	2.01 (1.90–2.13)	$p < 0.001$
40–49					3.01 (2.86–3.17)	$p < 0.001$	3.00 (2.85–3.16)	$p < 0.001$	3.00 (2.85–3.16)	$p < 0.001$
50–59					4.08 (3.88–4.29)	$p < 0.001$	4.07 (3.87–4.28)	$p < 0.001$	4.07 (3.87–4.28)	$p < 0.001$
60–69					2.95 (2.80–3.10)	$p < 0.001$	2.95 (2.80–3.10)	$p < 0.001$	2.95 (2.80–3.10)	$p < 0.001$
70–79					1.60 (1.52–1.69)	$p < 0.001$	1.61 (1.52–1.69)	$p < 0.001$	1.61 (1.52–1.69)	$p < 0.001$
80+					1.13 (1.07–1.20)	$p < 0.001$	1.14 (1.07–1.21)	$p < 0.001$	1.14 (1.07–1.21)	$p < 0.001$
IRSD							1.06 (1.04–1.07)	$p < 0.001$	1.06 (1.04–1.07)	$p < 0.001$
Access									1.00 (0.98–1.01)	0.616
AIC	235,931.6		235,927.9		227,254.6		227,193.8		227,195.5	
Variance	0.0255		0.0250		0.020		0.01703		0.01705	
PCV	-		−1.69%		−21.76%		−33.11%		−33.07%	
ICC (%)	0.8		0.8		0.6		0.5		0.5	
MOR	1.16		1.163		1.14		1.133		1.133	
Proportional variance explained by Access to primary care: +0.12%										

AIC—Akaike Information Criterion; TC—total cholesterol; ICC—Intra-cluster correlation coefficient; IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 6. Multilevel logistic regression model summaries of low HDL (<1 mmol/l).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	$p < 0.0001$		$p < 0.0001$		$p < 0.0001$		$p < 0.0001$		$p < 0.0001$	
	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
low HDL										
Intercept	0.13 (0.13–0.13)	$p < 0.001$	0.13 (0.13–0.18)	$p < 0.001$	0.06 (0.06–0.07)	$p < 0.001$	0.06 (0.06–0.07)	$p < 0.001$	0.06 (0.06–0.07)	$p < 0.001$
Access			0.92 (0.90–0.94)	$p < 0.001$						
Sex: Female					Reference					
Male					3.98 (3.85–4.11)	$p < 0.001$	3.98 (3.85–4.11)	$p < 0.001$	3.98 (3.85–4.11)	$p < 0.001$
Age: 18–29					Reference					
30–39					1.11 (1.03–1.20)	$p < 0.001$	1.12 (1.04–1.21)	$p < 0.001$	1.12 (1.04–1.21)	$p < 0.001$
40–49					0.99 (0.92–1.05)	0.658	1.00 (0.93–1.07)	0.957	1.00 (0.93–1.07)	0.947
50–59					0.88 (0.82–0.94)	$p < 0.001$	0.89 (0.83–0.95)	$p < 0.001$	0.88 (0.83–0.95)	$p < 0.001$
60–69					0.82 (0.77–0.88)	$p < 0.001$	0.83 (0.77–0.88)	$p < 0.001$	0.82 (0.77–0.88)	$p < 0.001$
70–79					0.86 (0.80–0.92)	$p < 0.001$	0.85 (0.80–0.91)	$p < 0.001$	0.85 (0.79–0.91)	$p < 0.001$
80+					0.93 (0.86–1.00)	$p < 0.010$	0.92 (0.85–0.99)	$p < 0.010$	0.91 (0.85–0.99)	$p < 0.010$
IRSD							0.81 (0.80–0.82)	$p < 0.001$	0.82 (0.80–0.83)	$p < 0.001$
Access									0.95 (0.93–0.97)	$p < 0.001$
AIC	130,649.70		130,601.4		122,700.0		122,291.9		122,271.4	
Variance	0.07		0.064		0.081		0.031		0.029	
PCV	-		−9.48%		+15.25%		−55.90%		−59.05%	
ICC (%)	2.1		1.9		2.4		0.9		0.9	
MOR	1.289		1.273		1.313		1.183		1.183	
Proportional variance explained by Access to primary care: −6.61%									6.61%	

AIC—Akaike Information Criterion; HDL—high density lipoprotein; ICC—Intra-cluster correlation coefficient; IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 7. Multilevel logistic regression model summaries of high ACR (≥30 mcg/L to mg/L).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	$p < 0.0001$		$p < 0.0001$		$p < 0.0001$		$p < 0.0001$		$p < 0.0001$	
	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
High ACR										
Intercept	0.04 (0.04–0.04)	$p < 0.001$	0.04 (0.04–0.04)	$p < 0.001$	0.02 (0.02–0.03)	$p < 0.001$	0.02 (0.02–0.03)	$p < 0.001$	0.02 (0.02–0.03)	$p < 0.001$
Access			0.91 (0.86–0.96)	$p < 0.001$						
Sex: Female					Reference					
Male					1.75 (1.60–1.92)	$p < 0.001$	1.76 (1.60–1.93)	$p < 0.001$	1.75 (1.60–1.92)	$p < 0.001$
Age: 18–29					Reference					
30–39					1.00 (0.69–1.45)	0.985	1.01 (0.69–1.46)	0.978	1.00 (0.69–1.46)	0.982
40–49					0.69 (0.47–0.97)	$p < 0.01$	0.70 (0.50–1.00)	$p < 0.01$	0.70 (0.50–1.00)	$p < 0.01$
50–59					0.77 (0.56–1.05)	0.101	0.77 (0.56–1.07)	0.115	0.77 (0.56–1.06)	0.115
60–69					0.95 (0.70–1.30)	0.762	0.96 (0.71–1.31)	0.794	0.96 (0.70–1.30)	0.777
70–79					1.55 (1.15–2.10)	$p < 0.001$	1.55 (1.14–2.09)	$p < 0.001$	1.54 (1.14–2.08)	$p < 0.001$
80+					2.74 (2.02–3.71)	$p < 0.001$	2.71 (2.00–3.67)	$p < 0.001$	2.70 (1.99–3.66)	$p < 0.001$
IRSD							0.82 (0.78–0.85)	$p < 0.001$	0.82 (0.79–0.86)	$p < 0.001$
Access									0.97 (0.91–1.02)	0.206
AIC	17,130.0		17,119.9		16,585.2		16,510.8		16,511.2	
Variance	0.092		0.085		0.073		0.028		0.025	
PCV	-		−7.92%		−20.53%		−69.14%		−72.39%	
ICC (%)	2.7		2.5		2.2		0.9		0.8	
MOR	1.34		1.321		1.30		1.175		1.165	
Proportional variance explained by Access to primary care: −10.53%										

AIC—Akaike Information Criterion; ACR–albumin creatinine ratio; ICC—Intra-cluster correlation coefficient;IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 8. Multilevel logistic regression model summaries of low eGFR (<60 mL/min/1.73 m^2).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	$p < 0.0001$		$p < 0.0001$		$p < 0.0001$		$p < 0.0001$		$p < 0.0001$	
	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
Low eGFR										
Intercept	0.11 (0.11–0.12)	$p < 0.001$	0.11 (0.11–0.12)	$p < 0.001$	0.00 (0.00–0.00)	$p < 0.001$	0.00 (0.00–0.00)	$p < 0.001$	0.00 (0.00–0.00)	$p < 0.001$
Access			0.89 (0.86–0.92)	$p < 0.001$						
Sex: Female					Reference					
Male					0.98 (0.95–1.01)	0.208	0.98 (0.95–1.01)	0.258	0.98 (0.95–1.01)	0.248
Age: 18–29					Reference					
30–39					1.66 (1.25–2.20)	$p < 0.001$	1.66 (1.24–2.23)	$p < 0.001$	1.65 (1.22–2.24)	$p < 0.001$
40–49					4.26 (3.35–5.41)	$p < 0.001$	4.27 (3.34–5.50)	$p < 0.001$	4.30 (3.32–5.58)	$p < 0.001$
50–59					12.26 (9.8–15.3)	$p < 0.001$	12.29 (9.73–15.52)	$p < 0.001$	12.28 (9.63–15.66)	$p < 0.001$
60–69					41.8 (33.6–51.8)	$p < 0.001$	41.84 (33.29–52.57)	$p < 0.001$	41.83 (32.97–53.06)	$p < 0.001$
70–79					150.7 (121.3–187.1)	$p < 0.001$	149.69 (119.3–187.9)	$p < 0.001$	149.6 (118.1–189.5)	$p < 0.001$
80+					509.3 (410.1–632.4)	$p < 0.001$	503.19 (400.9–631.6)	$p < 0.001$	503.0 (396.9–637.4)	$p < 0.001$
IRSD							0.90 (0.88–0.91)	$p < 0.001$	0.90 (0.88–0.91)	$p < 0.001$
Access									1.00 (0.98–1.02)	0.925
AIC	167,164.8		167,113.4		115,257.1		115,109.2		115,111.2	
Variance	0.189		0.176		0.024		0.013		0.013	
PCV	-		−6.53%		−87.26%		−93.31%		−93.26%	
ICC (%)	5.4		5.1		0.7		0.4		0.4	
MOR	1.51		1.492		1.16		1.113		1.113	
Proportional variance explained by Access to primary care: (+) 0.63%										

AIC—Akaike Information Criterion; eGFR—estimated glomerular filtration rate; ICC—Intra-cluster correlation coefficient; IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 9. Multilevel logistic regression model summaries of obesity (BMI ≥ 30 kg/m^2).

Variables	Model 1		Model 2		Model 3		Model 4		Model 5	
Significance (LRT)	***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**		***p* < 0.0001**	
	OR (95% CI)	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**	**OR (95% CI)**	***p* Value**
Obesity										
Intercept	0.51 (0.50–0.52)	$p < 0.001$	0.51 (0.50–0.52)	$p < 0.001$	0.25 (0.24–0.26)	$p < 0.001$	0.25 (0.24–0.25)	$p < 0.001$	0.25 (0.24–0.26)	$p < 0.001$
Access			0.88 (0.86–0.90)	$p < 0.001$						
Sex: Female					Reference					
Male					0.99 (0.97–1.01)	0.214	0.99 (0.97–1.01)	0.195	0.99 (0.97–1.01)	0.193
Age: 18–29					Reference					
30–39					1.63 (1.56–1.71)	$p < 0.001$	1.64 (1.57–1.71)	$p < 0.001$	1.64 (1.57–1.71)	$p < 0.001$
40–49					2.20 (2.11–2.29)	$p < 0.001$	2.21 (2.12–2.30)	$p < 0.001$	2.20 (2.12–2.30)	$p < 0.001$
50–59					2.44 (2.34–2.53)	$p < 0.001$	2.45 (2.35–2.54)	$p < 0.001$	2.44 (2.34–2.53)	$p < 0.001$
60–69					2.73 (2.63–2.84)	$p < 0.001$	2.74 (2.63–2.85)	$p < 0.001$	2.72 (2.62–2.83)	$p < 0.001$
70–79					2.44 (2.34–2.54)	$p < 0.001$	2.44 (2.34–2.54)	$p < 0.001$	2.42 (2.33–2.52)	$p < 0.001$
80+					1.46 (1.39–1.55)	$p < 0.001$	1.45 (1.38–1.54)	$p < 0.001$	1.45 (1.37–1.53)	$p < 0.001$
IRSD							0.81 (0.79–0.82)	$p < 0.001$	0.82 (0.80–0.83)	$p < 0.001$
Access									0.93 (0.91–0.95)	$p < 0.001$
AIC	242,793.2		242,686.2		239,122.6		238,731.8		238,680.6	
Variance	0.115		0.099		0.117		0.068		0.062	
PCV	-		−14.20%		+1.48%		−41.21%		−46.19%	
ICC (%)	3.4		2.9		3.4		2.0		1.8	
MOR	1.38		1.350		1.39		1.282		1.268	
Proportional variance explained by Access to primary care: −8.47%										

AIC—Akaike Information Criterion; BMI—body mass index; ICC—Intra-cluster correlation coefficient; IRSD—Index of Relative Socioeconomic Disadvantage; LRT—Likelihood ratio test; Model 1—null model at SA1 level; Model 2—M1 + Primary care access index of SA1s; Model 3—M1 + individual-level: age + sex; Model 4—Model 3 + Area level: Index of Relative Socioeconomic Disadvantage score of SA1s; Model 5—Model 4 + Primary care access index of SA1s; SA1—Statistical area-level 1; MOR—Median odds ratio; PCV—Proportional change in variance.

Table 10. Summary of model fit values and comparison of the models.

		FBSL	HbA1c	TC	HDL	ACR	eGFR	Obesity
Model 1				Null Model				
	AIC	111,022.8	50,114.5	235,931.6	130,649.7	17,130.0	167,164.8	242,793.2
	τ^2	0.101	0.103	0.025	0.071	0.092	0.189	0.115
	ICC (%)	3.0	3.0	0.8	2.1	2.7	5.4	3.4
	MOR	1.36	1.36	1.16	1.29	1.34	1.51	1.38
Model 2				Access Model				
	AIC	110,962.3	50,105.9	235,927.9	130,601.4	17,119.9	167,113.4	242,686.2
	τ^2	0.091	0.100	0.025	0.064	0.085	0.176	0.099
	ICC (%)	2.7	3.0	0.8	1.9	2.5	5.1	2.9
	MOR	1.334	1.353	1.163	1.273	1.321	1.492	1.350
	PCV	−9.98%	−2.430%	−1.69%	−9.48%	−7.92%	−6.53%	−14.20%
Model 3				Sex + Age Adjusted Model				
	AIC	103,066.2	49,690.2	227,254.6	122,700.0	16,585.2	115,257.1	239,122.6
	τ^2	0.103	0.106	0.020	0.081	0.073	0.024	0.117
	ICC (%)	3.0	3.1	0.6	2.4	2.2	0.7	3.4
	MOR	1.36	1.358	1.14	1.31	1.30	1.16	1.39
	PCV	+ 1.88%	+ 3.02%	−21.76%	+15.25%	−20.53%	−87.26%	+1.48%
Model 4				Sex + Age + IRSD Adjusted Model				
	AIC	102,652.6	49,438.2	227,193.8	122,291.9	16,510.8	115,109.2	238,731.8
	τ^2	0.040	0.048	0.017	0.031	0.028	0.013	0.068
	ICC (%)	1.2	1.4	0.5	0.9	0.9	0.4	2.0
	MOR	1.209	1.231	1.133	1.183	1.175	1.113	1.282
	PCV	−60.90%	−53.78%	−33.11%	−55.90%	−69.14%	−93.31%	−41.21%
Model 5				Sex + Age + IRSD Adjusted and Access included Model				
	AIC	102,652.0	49,440.0	227,195.5	122,271.4	16,511.2	115,111.2	238,680.6
	τ^2	0.039	0.047	0.017	0.029	0.025	0.013	0.062
	ICC (%)	1.2	1.4	0.5	0.9	0.8	0.4	1.8
	MOR	1.209	1.231	1.133	1.183	1.165	1.113	1.268
	PCV	−61.05%	−53.80%	−33.07%	−59.05%	−72.39%	−93.26%	−46.19%

AIC—Akaike Information Criterion; τ^2—residual variance; IC—intra-cluster correlation coefficients; MOR—median odds ratio; PCV—proportional change in variance; FBSL—fasting blood sugar level; HbA1c—glycated haemoglobin; TC—total cholesterol; HDL—high density lipoprotein; ACR—albumin creatinine ratio; eGFR—estimated glomerular filtration rate.

4. Discussion

This study aimed to inform area-specific interventions for the prevention and control of CMRFs, based on the primary care access status of the small areas within Illawarra-Sholhaven region of NSW, Australia. After adjusting for the covariates, we found that: a) greater access to primary care was associated with a reduction in the odds of low HDL and obesity but was not associated with high FBSL, high HbA1C, high TC, high ACR and low eGFR; b) the general contextual effect of areas on each of the CMRFs were minimal; and c) the geographic variation of CMRFs specifically explained by primary care access was small and did not demonstrate any attenuating effect on the contribution of area-level disadvantage on the variation of CMRFs in the study region. The results demonstrate that though the probability of low HDL and obesity decreases with increasing primary care access, the low general contextual effects of the areas on each of the CMRFs (i.e., low ICCs of Model 3s, ranges 0.6–3.4%) indicate minimal difference between the small areas after controlling for the study variables. Thus, the findings suggest that preventive interventions should not only be focused on areas with lower primary care access. Rather, interventions should be universal but proportional to the need and risk level of the people for the prevention and control of CMRFs. Primary care access was associated with all CRMFs in unadjusted models but only with low HDL and obesity in models fully adjusted for

individual- and area-level covariates. These findings support the arguments of the possible role of confounders and reverse causality in ecological models [57], which question the previously established associations between primary care access and improved health [58,59]. The study suggests higher odds of being identified with low HDL and obesity with reduced access to primary care. In previous studies, when the relationship between health care service outcomes and travel time was modelled using multilevel logistic regression, it was found that GP consultations were less likely to happen when the travel time was longer, which is more common in rural areas [21]. The current study outcomes are consistent with those findings. However, it should also be noted that the current findings pertain only to the geographical/spatial accessibility of the primary health care services within 30 km distance of an SA1 centroid, rather than their road network access, actual usage and affordability.

The primary care access index, derived from the study region, ranged from 0 to 5.41 general practitioners per 1000 people (mean = 2.1, SD = 0.77). Multiple previous studies have reported inequalities in the geographic access to primary care services, using different enhanced versions of the 2SFCA method [45,60–68]. For example, the spatial accessibility index derived from rural Otago in New Zealand, using the travel time distance, ranged between 1 to 10, where a higher score indicated better access [62]. The accessibility index reported from Thimphu district in Bhutan ranged between 0 and 1, where 1 was the maximum access [69]. The spatial accessibility index of GP accessibility in England has been reported to range between 7.2 (South of England) and 13.3 (in London) [69]. The access map of the study region (Figure 2) clearly shows a polarisation of the higher access indices along the northern and southern ends of the study region, thus a visible inequality in their distribution. The WHO recommends universal access to primary care for all populations, where geographic access is one part of physical access to primary care [70].

Area-level disadvantage explained more geographic variation in CMRFs than area-level access to primary care. Inclusion of the access index in the final model did not demonstrate any reduction in the variance explained by area-level disadvantage on the geographic variation of CMRFs. This finding supports the importance of overall socioeconomic development of areas to reduce CMRF risk. Moreover, the ICC values of the final models were too small to suggest any meaningful area-level difference in the modelled CMRF variables. This would support the call for universal approaches for the prevention and control of CMRFs rather than any targeted area-level approaches, but with a proportional priority to disadvantaged populations in the study region [24,28,71].

This study has to be considered within its limitations. First, the cross-sectional nature of the study does not support causal inference. Second, the CMRF data used in this study are from people already utilising health care service in the study area, so care should be taken in generalising the results to the overall population. The SIMLR database does not include hospital or emergency service based tests. Therefore we believe that the database has a reasonable representation of community dwelling adults in the study region. However, it should be noted that the study sample includes only people who have accessed health care and pathology services, and the omission of those who have not accessed care may have biased our results. Given our population coverage this seem unlikely. Third, the study used a radial buffer distance of 30km for access calculations rather than travel time/distance because proprietary road network data were unavailable for this study. Thus, the patients' actual experiences of seeking physical access in daily life need not exactly reflect the compound measure of access index adopted in this study. Even though the 30km buffer distance helped to include a maximum coverage of the population in relation to the geographic location of the primary care providers, this distance might have also influenced the discriminatory accuracy of the SA1s in the multilevel analyses. In addition, it should also be noted that the access index described in the study pertains only to the geographical reachability, but not to the affordability and acceptability of the available services. Forth, the study did not include blood pressure as a variable, although it is a major CMRF, due to non-availability of data. We were also unable to adjust for ethnicity for the same reason.

The main strength of this study is the use of a large population-derived database comprising a wide range of CMRFs. The research adds to the very few studies which consider multiple CMRF variables

from the same region [18–20,72–74] and is unusual for its hierarchical analysis of the associations between a range of CMRFs and primary care access in a widely dispersed population.

Future research is required to investigate other area-level attributes contributing to the geographic variation of CMRFs in the study region. Our previous research has reported that area-level disadvantage contributes 14.7–57.8% of the geographic variation in CMRFs. The current study extended the previous findings by identifying the specific contribution of area-level primary care access, ranging between 0.0–10.5%. Further area-level analyses are required to identify other factors contributing to the geographic inequality of the CMRFs in the study region.

5. Conclusions

The findings of the study suggest that adults residing in areas that have a poor primary care access are more likely to be identified with low HDL and obesity. However, the specific contribution of area-level primary care access was small when compared to the contribution of area-level disadvantage. The finding supports the importance of overall socioeconomic development of areas to reduce CMRF risk, while supporting universal approaches for the prevention and control of CMRFs which are proportional to the need and disadvantage level of the individuals. Future research including other aspects of primary care access such as road-network access, financial affordability and acceptance of the services might help to provide an overall picture of the contributing role of primary care in the study region.

Author Contributions: Conceptualization, R.T., X.F., D.J.M., and A.B.; methodology, R.T., D.J.M., A.B., and X.F.; software, R.T., and D.M.; validation, A.B., X.F., and D.J.M.; formal analysis, R.T.; investigation, R.T.; resources, R.T., and A.B.; data curation, D.M., and R.T.; writing—original draft preparation, R.T.; writing—review and editing, A.B., D.J.M., X.F.; visualization, R.T.; supervision, A.B., X.F., and D.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The Australian Government Research Training Program Scholarship covers the university tuition fees of the PhD candidate who is the principal author in this study (RT).

Acknowledgments: We would like to thank Southern IML Pathology and staff for their generosity in providing data for the SIMLR Cohort Study and ongoing support. In particular we would like to thank *Bryan Jones* for providing the technical expertise for the data acquisition and helpful comments on the manuscript. Southern IML Pathology are the owners of the data contained within this publication and the Illawarra Health and Medical Research Institute (IHMRI) is the custodian facilitating access to the data. https://www.ihmri.org.au/research-projects/simlr-cohort-study/.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alkerwi, A.; Bahi, I.E.; Stranges, S.; Beissel, J.; Delagardelle, C.; Noppe, S.; Kandala, N.-B. Geographic variations in cardiometabolic risk factors in Luxembourg. *Int. J. Environ. Res. Public Health* **2017**, *14*, 648. [CrossRef] [PubMed]
2. Astell-Burt, T.; Feng, X.; Kolt, G.S.; McLean, M.; Maberly, G. Understanding geographical inequities in diabetes: Multilevel evidence from 114,755 adults in Sydney, Australia. *Diabetes Res. Clin. Pract.* **2014**, *106*, e68–e73. [CrossRef] [PubMed]
3. Barker, L.E.; Kirtland, K.A.; Gregg, E.W.; Geiss, L.S.; Thompson, T.J. Geographic distribution of diagnosed diabetes in the US: A diabetes belt. *Am. J. Prev. Med.* **2011**, *40*, 434–439. [CrossRef] [PubMed]
4. Congdon, P. Estimating diabetes prevalence by small area in England. *J. Public Health* **2006**, *28*, 71–81. [CrossRef] [PubMed]
5. Lawlor, D.; Bedford, C.; Taylor, M.; Ebrahim, S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J. Epidemiol. Community Health* **2003**, *57*, 134–140. [CrossRef]
6. Paquet, C.; Chaix, B.; Howard, N.J.; Coffee, N.T.; Adams, R.J.; Taylor, A.W.; Thomas, F.; Daniel, M. Geographic clustering of cardiometabolic risk factors in metropolitan centres in France and Australia. *Int. J. Environ. Res. Public Health* **2016**, *13*, 519. [CrossRef]

7. Toms, R.; Bonney, A.; Mayne, D.J.; Feng, X.; Walsan, R. Geographic and area-level socioeconomic variation in cardiometabolic risk factor distribution: A systematic review of the literature. *Int. J. Health Geogr.* **2019**, *18*, 1. [CrossRef]
8. Toms, R.; Mayne, D.J.; Feng, X.; Bonney, A. Geographic variation in cardiometabolic risk distribution: A cross-sectional study of 256,525 adult residents in the Illawarra-Shoalhaven region of the NSW, Australia. *PLoS ONE* **2019**, *14*, e0223179. [CrossRef]
9. Valdés, S.; García-Torres, F.; Maldonado-Araque, C.; Goday, A.; Calle-Pascual, A.; Soriguer, F.; Castaño, L.; Catalá, M.; Gomis, R.; Rojo-Martínez, G. Prevalence of obesity, diabetes and other cardiovascular risk factors in Andalusia (southern Spain). Comparison with national prevalence data. The Di@bet. es study. *Rev. Española Cardiol.* **2014**, *67*, 442–448.
10. Zhou, M.; Astell-Burt, T.; Bi, Y.; Feng, X.; Jiang, Y.; Li, Y.; Page, A.; Wang, L.; Xu, Y.; Wang, L. Geographical variation in diabetes prevalence and detection in China: Multilevel spatial analysis of 98,058 adults. *Diabetes Care* **2015**, *38*, 72–81. [CrossRef]
11. Andersen, A.; Carson, C.; Watt, H.; Lawlor, D.; Avlund, K.; Ebrahim, S. Life-course socio-economic position, area deprivation and Type 2 diabetes: Findings from the British Women's Heart and Health Study. *Diabet. Med.* **2008**, *25*, 1462–1468. [CrossRef] [PubMed]
12. Bonney, A.; Mayne, D.J.; Jones, B.D.; Bott, L.; Andersen, S.E.; Caputi, P.; Weston, K.M.; Iverson, D.C. Area-level socioeconomic gradients in overweight and obesity in a community-derived cohort of health service users—A cross-sectional study. *PLoS ONE* **2015**, *10*, e0137261. [CrossRef] [PubMed]
13. Cubbin, C.; Sundquist, K.; Ahlén, H.; Johansson, S.E.; Winkleby, M.A.; Sundquist, J. Neighborhood deprivation and cardiovascular disease risk factors: Protective and harmful effects. *Scand. J. Public Health* **2006**, *34*, 228–237. [CrossRef] [PubMed]
14. Dragano, N.; Bobak, M.; Wege, N.; Peasey, A.; Verde, P.E.; Kubinova, R.; Weyers, S.; Moebus, S.; Möhlenkamp, S.; Stang, A. Neighbourhood socioeconomic status and cardiovascular risk factors: A multilevel analysis of nine cities in the Czech Republic and Germany. *BMC Public Health* **2007**, *7*, 255. [CrossRef]
15. Lawlor, D.A.; Davey Smith, G.; Patel, R.; Ebrahim, S. Life-course socioeconomic position, area deprivation, and coronary heart disease: Findings from the British Women's Heart and Health Study. *Am. J. Public Health* **2005**, *95*, 91–97. [CrossRef]
16. Maier, W.; Scheidt-Nave, C.; Holle, R.; Kroll, L.E.; Lampert, T.; Du, Y.; Heidemann, C.; Mielck, A. Area level deprivation is an independent determinant of prevalent type 2 diabetes and obesity at the national level in Germany. Results from the National Telephone Health Interview Surveys 'German Health Update' GEDA 2009 and 2010. *PLoS ONE* **2014**, *9*, e89661. [CrossRef]
17. Mujahid, M.S.; Diez Roux, A.V.; Borrell, L.N.; Nieto, F.J. Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics. *Obes. Res.* **2005**, *13*, 1412–1421. [CrossRef]
18. Naimi, A.I.; Paquet, C.; Gauvin, L.; Daniel, M. Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area. *Int. J. Environ. Res. Public Health* **2009**, *6*, 3082–3096. [CrossRef]
19. Roux, A.V.D.; Jacobs, D.R.; Kiefe, C.I. Neighborhood characteristics and components of the insulin resistance syndrome in young adults: The coronary artery risk development in young adults (CARDIA) study. *Diabetes Care* **2002**, *25*, 1976–1982. [CrossRef]
20. Unger, E.; Diez-Roux, A.V.; Lloyd-Jones, D.M.; Mujahid, M.S.; Nettleton, J.A.; Bertoni, A.; Badon, S.E.; Ning, H.; Allen, N.B. Association of neighborhood characteristics with cardiovascular health in the multi-ethnic study of atherosclerosis. *Circulation* **2014**, *7*, 524–531. [CrossRef]
21. Hiscock, R.; Pearce, J.; Blakely, T.; Witten, K. Is neighborhood access to health care provision associated with individual-level utilization and satisfaction? *Health Serv. Res.* **2008**, *43*, 2183–2200. [CrossRef] [PubMed]
22. Schmidt, M.I.; Duncan, B.B.; e Silva, G.A.; Menezes, A.M.; Monteiro, C.A.; Barreto, S.M.; Chor, D.; Menezes, P.R. Chronic non-communicable diseases in Brazil: Burden and current challenges. *Lancet* **2011**, *377*, 1949–1961. [CrossRef]
23. Weinberger, M.; Oddone, E.Z.; Henderson, W.G. Does increased access to primary care reduce hospital readmissions? *N. Engl. J. Med.* **1996**, *334*, 1441–1447. [CrossRef] [PubMed]
24. Kirby, J.B.; Kaneda, T. Neighborhood socioeconomic disadvantage and access to health care. *J. Health Soc. Behav.* **2005**, *46*, 15–31. [CrossRef]

25. Atallah, A.; Inamo, J.; Larabi, L.; Chatellier, G.; Rozet, J.; Machuron, C.; De Gaudemaris, R.; Lang, T. Reducing the burden of arterial hypertension: What can be expected from an improved access to health care? Results from a study in 2420 unemployed subjects in the Caribbean. *J. Hum. Hypertens* **2007**, *21*, 316–322. [CrossRef]
26. Kotchen, J.M.; Shakoor-Abdullah, B.; Walker, W.E.; Chelius, T.H.; Hoffmann, R.G.; Kotchen, T.A. Hypertension control and access to medical care in the inner city. *Am. J. Public Health* **1998**, *88*, 1696–1699. [CrossRef]
27. Occelli, F.; Deram, A.; Génin, M.; Noël, C.; Cuny, D.; Glowacki, F. Mapping end-stage renal disease (ESRD): Spatial variations on small area level in northern France, and association with deprivation. *PLoS ONE* **2014**, *9*, e110132. [CrossRef]
28. Walsh, M.G.; Zgibor, J.; Songer, T.; Borch-Johnsen, K.; Orchard, T.J.; Investigators, A.D. The socioeconomic correlates of global complication prevalence in type 1 diabetes (T1D): A multinational comparison. *Diabetes Res. Clin. Pract.* **2005**, *70*, 143–150. [CrossRef]
29. World Health Organization. Gender, Equity and Human Rights: Accessibility. Available online: https://www.who.int/gender-equity-rights/understanding/accessibility-definition/en/ (accessed on 3 November 2019).
30. Angier, H.; Likumahuwa, S.; Finnegan, S.; Vakarcs, T.; Nelson, C.; Bazemore, A.; Carrozza, M.; DeVoe, J.E. Using geographic information systems (GIS) to identify communities in need of health insurance outreach: An OCHIN practice-based research network (PBRN) report. *J. Am. Board Fam. Med.* **2014**, *27*, 804–810. [CrossRef]
31. Larkins, S.; Gupta, T.S.; Evans, R.; Murray, R.; Preston, R. Addressing inequities in access to primary health care: Lessons for the training of health care professionals from a regional medical school. *Aust. J. Prim. Health* **2011**, *17*, 362–368. [CrossRef]
32. Harris, M.I. Racial and ethnic differences in health care access and health outcomes for adults with type 2 diabetes. *Diabetes Care* **2001**, *24*, 454–459. [CrossRef] [PubMed]
33. Pattenden, S.; Casson, K.; Cook, S.; Dolk, H. Geographical variation in infant mortality, stillbirth and low birth weight in Northern Ireland, 1992–2002. *J. Epidemiol. Community Health* **2011**, *65*, 1159–1165. [CrossRef] [PubMed]
34. Australian Bureau of Statistics. Census Data. Available online: https://www.abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011?opendocument&navpos=280 (accessed on 3 November 2019).
35. Australian Bureau of Statistics. Australian Statistical Geography Standard (ASGS): Census Dictionary. 2011. Available online: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2901.0Chapter23102011 (accessed on 3 November 2019).
36. Australian Bureau of Statistics (ASGS). Australian Statistical Geography Standard (ASGS): Volume 1—Main Structure: STATISTICAL AREA LEVEL 1 (SA1). Available online: https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+ (accessed on 3 November 2019).
37. Australian Bureau of Statistics. Main Features—IRSD. Available online: https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2033.0.55.001main+features100052011 (accessed on 3 November 2019).
38. The Royal Australian College of General Practitioners & Diabetes Australia. General Practice Management of Type 2 Diabetes 2016–2018. Available online: https://static.diabetesaustralia.com.au/s/fileassets/diabetes-australia/5d3298b2-abf3-487e-9d5e-0558566fc242.pdf (accessed on 3 November 2019).
39. Australian Bureau of Statistics. Australian Health Survey: Biomedical Results for Chronic Diseases, 2011-12. Available online: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0052011-12 (accessed on 3 November 2019).
40. National Heart Foundation of Australia. Lipid Management Profile for Health Professionals. Available online: https://www.heartfoundation.org.au/for-professionals/clinical-information/lipid-management (accessed on 3 November 2019).
41. National Kidney Foundation (USA). Albumin Creatinine Ratio (ACR). Available online: https://www.kidney.org/kidneydisease/siemens_hcp_acr (accessed on 3 November 2019).
42. World Health Organization. *Obesity: Preventing and Managing the Global Epidemic: Technical Report Series*; WHO Technical Report Series, no. 894; World Health Organization: Geneva, Switzerland, 2000; ISBN 924-120-8945.
43. Luo, W.; Wang, F. Measures of spatial accessibility to health care in a GIS environment: Synthesis and a case study in the Chicago region. *Environ. Plan. B* **2003**, *30*, 865–884. [CrossRef]
44. Li, Y.; Vo, A.; Randhawa, M.; Fick, G. Designing utilization-based spatial healthcare accessibility decision support systems: A case of a regional health plan. *Decis. Support Syst.* **2017**, *99*, 51–63. [CrossRef]

45. McGrail, M.R. Spatial accessibility of primary health care utilising the two step floating catchment area method: An assessment of recent improvements. *Int. J. Health Geogr.* **2012**, *11*, 50. [CrossRef] [PubMed]
46. Ghosh, A.; Charlton, K.E.; Girdo, L.; Batterham, M. Using data from patient interactions in primary care for population level chronic disease surveillance: The Sentinel Practices Data Sourcing (SPDS) project. *BMC Public Health* **2014**, *14*, 557. [CrossRef] [PubMed]
47. Goldstein, H.; Browne, W.; Rasbash, J. Partitioning variation in multilevel models. *Underst. Stat.* **2002**, *1*, 223–231. [CrossRef]
48. Szmaragd, C.; Leckie, G. *Module 6: Regression Models for Binary Responses R Practical*; Centre for Multilevel Modelling: Bristol, UK, 2011.
49. Environmental Systems Research Institute (ESRI). *ArcGIS 10.4.1*; ESRI Inc.: Redlands, CA, USA; Available online: https://www.esri.com/ (accessed on 3 November 2019).
50. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria; Available online: https://www.R-project.org/ (accessed on 3 November 2019).
51. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *arXiv* **2014**, arXiv:1406.5823.
52. Zeileis, A.; Hothorn, T. Diagnostic Checking in Regression Relationships. R News 2002: 7–10. Available online: http://CRAN.R-project.org/doc/Rnews/ (accessed on 25 August 2019).
53. Bolker, B. R Documention: Fitting Generalized Linear Mixed-Effects Models. Available online: https://www.rdocumentation.org/packages/lme4/versions/1.1-23/topics/glmer (accessed on 3 November 2019).
54. Liu, Q.; Pierce, D.A. A note on Gauss—Hermite quadrature. *Biometrika* **1994**, *81*, 624–629.
55. Australian Bureau of Statistics. Technical Paper: Socio-Economic Indexes for Areas (SEIFA). Available online: https://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/22CEDA8038AF7A0DCA257B3B00116E34/$File/2033.0.55.001%20seifa%202011%20technical%20paper.pdf (accessed on 3 November 2019).
56. Aho, K.; Derryberry, D.; Peterson, T. Model selection for ecologists: The worldviews of AIC and BIC. *Ecology* **2014**, *95*, 631–636. [CrossRef]
57. Gulliford, M.C. Availability of primary care doctors and population health in England: Is there an association? *J. Public Health* **2002**, *24*, 252–254. [CrossRef] [PubMed]
58. Shi, L.; Starfield, B. The effect of primary care physician supply and income inequality on mortality among blacks and whites in US metropolitan areas. *Am. J. Public Health* **2001**, *91*, 1246–1250. [CrossRef]
59. Shi, L.; Starfield, B.; Politzer, R.; Regan, J. Primary care, self-rated health, and reductions in social disparities in health. *Health Serv. Res.* **2002**, *37*, 529–550. [CrossRef] [PubMed]
60. Bagheri, N.; Holt, A.; Benwell, G.L. Using geographically weighted regression to validate approaches for modelling accessibility to primary health care. *Appl. Spat. Anal. Policy* **2009**, *2*, 177. [CrossRef]
61. Bissonnette, L.; Wilson, K.; Bell, S.; Shah, T.I. Neighbourhoods and potential access to health care: The role of spatial and aspatial factors. *Health Place* **2012**, *18*, 841–853. [CrossRef] [PubMed]
62. Gruen, R.L.; Weeramanthri, T.S.; Knight, S.E.; Bailie, R.S. Specialist outreach clinics in primary care and rural hospital settings (Cochrane Review). *Community Eye Health.* **2006**, *19*, 31. [PubMed]
63. Kanuganti, S.; Sarkar, A.K.; Singh, A.P.; Arkatkar, S.S. Quantification of accessibility to health facilities in rural areas. *Case Stud. Transp. Policy* **2015**, *3*, 311–320. [CrossRef]
64. Mobley, L.R.; Root, E.; Anselin, L.; Lozano-Gracia, N.; Koschinsky, J. Spatial analysis of elderly access to primary care services. *Int. J. Health Geogr.* **2006**, *5*, 19. [CrossRef]
65. Munoz, U.H.; Källestål, C. Geographical accessibility and spatial coverage modeling of the primary health care network in the Western Province of Rwanda. *Int. J. Health Geogr.* **2012**, *11*, 40. [CrossRef]
66. Shah, T.I.; Milosavljevic, S.; Bath, B. Measuring geographical accessibility to rural and remote health care services: Challenges and considerations. *Spat. Spatio Temporal Epidemiol.* **2017**, *21*, 87–96. [CrossRef]
67. Wang, F.; Luo, W. Assessing spatial and nonspatial factors for healthcare access: Towards an integrated approach to defining health professional shortage areas. *Health Place* **2005**, *11*, 131–146. [CrossRef]
68. Wang, X.; Yang, H.; Duan, Z.; Pan, J. Spatial accessibility of primary health care in China: A case study in Sichuan Province. *Soc. Sci Med.* **2018**, *209*, 14–24. [CrossRef] [PubMed]
69. Bauer, J.; Müller, R.; Brüggmann, D.; Groneberg, D.A. Spatial accessibility of primary care in England: A cross-sectional study using a floating catchment area method. *Health Serv. Res.* **2018**, *53*, 1957–1978. [CrossRef] [PubMed]

70. Evans, D.B.; Hsu, J.; Boerma, T. *Universal Health Coverage and Universal Access*; SciELO Public Health: São Paulo, Brazil, 2013.
71. Rose, G. Sick individuals and sick populations. *Int. J. Epidemiol.* **1985**. [CrossRef] [PubMed]
72. Clark, C.R.; Ommerborn, M.J.; Hickson, D.A.; Grooms, K.N.; Sims, M.; Taylor, H.A.; Albert, M.A. Neighborhood disadvantage, neighborhood safety and cardiometabolic risk factors in African Americans: Biosocial associations in the Jackson Heart study. *PLoS ONE* **2013**, *8*, e63254. [CrossRef]
73. Gabert, R.; Thomson, B.; Gakidou, E.; Roth, G. Identifying high-risk neighborhoods using electronic medical records: A population-based approach for targeting diabetes prevention and treatment interventions. *PLoS ONE* **2016**, *11*, e0159227. [CrossRef]
74. Keita, A.D.; Judd, S.E.; Howard, V.J.; Carson, A.P.; Ard, J.D.; Fernandez, J.R. Associations of neighborhood area level deprivation with the metabolic syndrome and inflammation among middle-and older-age adults. *BMC Public Health* **2014**, *14*, 1319. [CrossRef]

Supplementary Material: 5 (Media/Press release of the research)

What your address tells us about your health

28/08/2020

Researchers investigate how cardiometabolic risk varies within the Illawarra-Shoalhaven region

In partnership with Southern.IML Pathology, researchers at the Illawarra Health and Medical Research Institute (IHMRI), UOW's School of Medicine, UNSW School of Population Health and the Illawarra Shoalhaven Local Health District, have completed a four-year research project looking at the area-level geographic variation in cardiometabolic risk.

PhD Candidate Renin Toms has led the research which was conducted throughout the Illawarra-Shoalhaven region. Ms Toms was driven to undertake the research given the prevalence of cardiovascular disease, which is a leading cause of death in Australia and around the world.

Cardiometabolic risk factors (CMRF) can predispose and worsen cardiovascular disease amongst people and include conditions such as diabetes, chronic kidney disease, cholesterol, and obesity. The prevalence of these risk factors varies geographically and whilst many Australians enjoy better health than other countries, our health is not distributed equally.

"*We wanted to delve deeper than existing national data, to understand how cardiometabolic risk varies by areas within our region. We looked at what influence area disadvantage and primary health care access have on risk, so that we can better plan for health services in the future,*" said Ms Toms.

The research analysed data on individuals within small geographic areas used by the Australian Bureau of Statistics called Statistical Areas Level 1 (SA1). SA1s typically have a population size of 200 to 800 persons and the Illawarra-Shoalhaven region covers a total of 980 SA1s, making it a versatile region to perform such a study. Data included objectively measured routine pathology tests between 2012 and 2017 and reviewed a range of CMRF variables including fasting blood sugar level, total cholesterol and body mass index. The de-identified data were generously supplied by Southern.IML Pathology.

"*This is the first study of its kind in the region, as it considers multiple CMRF variables, is based on population derived data over an extended period and reports on both single (individual) and multilevel (area) analyses*," said Ms Toms.

Outcomes of the research confirmed the relationship between socioeconomically disadvantaged areas and prevalence of CMRF. Over and above age or sex, living in a disadvantaged neighbourhood proportionally increases a person's likelihood of falling into the higher risk CMRF category.

"*Our research suggests that in order to prevent and control cardiometabolic risk factors in this region, individuals should have equal access to health resources, but these should be available at a higher level in areas with greater need and disadvantage*," Ms Toms said.

Media contact
Emily O'Keefe, Marketing & Communications Manager

P: 02 4221 3620

e-mail: eokeefe@uow.edu.au

The research results were published in Scientific Reports (July 2020), International Journal of Environmental Research and Public Health (January 2020), PlosOne (January 2019) and International Journal of Health Geographics (December 2019). The research was approved by the University of Wollongong and the Illawarra Shoalhaven Local Health District (ISLHD) Research Ethics Committee and was supervised by Professor Andrew Bonney from UOW and Associate Professor Xiaoqi Feng from UNSW and Dr Darren Mayne from Public Health Unit, Illawarra Shoalhaven Local Health District. The researchers would like to acknowledge the generous support of Southern.IML Pathology for the research.


thank you