

Lipids and dementia: An investigation of their relationship

Luke A McGuinness

University of Bristol

*A thesis submitted for the degree of
Doctor of Philosophy*

2022

Abstract

Background

In the UK, an estimated 800000 people are currently living with dementia and this number is expected to double by 2040. Despite the number of dementia cases and decades of research, there remains much unknown about the pathogenesis and progression of the disease, and, at present, no effective treatment exists to arrest or reverse the cognitive decline associated with the condition. In this context, identification of causal relationships between modifiable targets and dementia risk is central to the development of evidence-based prevention strategies and will be critically important in maintaining the long-term health of the ageing public. Blood lipid levels have been implicated in the aetiology of dementia by genetic linkage and functional cell biology studies, but current epidemiological evidence has yet to reach a consensus on their role in dementia risk.

Lipids and dementia: An investigation of their relationship



Luke A McGuinness

University of Bristol

A thesis submitted for the degree of

Doctor of Philosophy

2022

Abstract

Background

In the UK, an estimated 800000 people are currently living with dementia and this number is expected to double by 2040. Despite the number of dementia cases and decades of research, there remains much unknown about the pathogenesis and progression of the disease, and, at present, no effective treatment exists to arrest or reverse the cognitive decline associated with the condition. In this context, identification of causal relationships between modifiable targets and dementia risk is central to the development of evidence-based prevention strategies and will be critically important in maintaining the long-term health of the ageing public. Blood lipid levels have been implicated in the aetiology of dementia by genetic linkage and functional cell biology studies, but current epidemiological evidence has yet to reach a consensus on their role in dementia risk.

For Brendan McHugh

Acknowledgements

Professional

There are so many people without whom this would not have been possible:

[Check all name spellings]

Sharen O'Keffe Neal Haddaway Martin Westgate

Emily Hennessy

robvis contributors: (there are at least 3)

Annual reviewers: Rupert Payne and Emma Anderson

Those who contributed their time to dual marking papers and code with me,

Thanks to Francesca for sharing advance copy of review of Mendelian randomisation studies, to inform choice of tool.

I am a NIHR Doctoral Research Fellow with grant code **TBC**. The funding source had no role in the design, conduct of the study, collection, management, analysis and interpretation or preparation, review, or approval of the thesis.

Thanks also to all those who raised issues or otherwise provided feedback on the open source tools built as part of this thesis: @abannachbrown, @ailengt, @AJFOWLER, @danielskatz, @earcanal, @Jamesohare1, @jannisborn, @paulsharpeY, @rachaelm-burke, @rdboyes, @seabbs, @vanAmsterdam, @yochannah, and @zaddyzad.

Statistical support: check emails!

All authors who volunteered time and effort to provided

A special thanks goes out to Mark Newbury of the

My student, Georgia Gourlay

Personal

Finally, thanks to parents and Ciara Gardiner.

Canynge Hall, Bristol

1 December 2021

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

COVID Impact Statement

My work was influence by COVID in the following ways:

Contents

List of Figures	xxi
List of Tables	xxiii
List of Abbreviations	xxv
Covering material	1
1 Introduction	3
1.1 Additional ideas	3
1.1.1 General	3
1.1.2 Background	4
1.1.3 medrxivr	4
1.1.4	4
2 Background, Theoretical framework, Aims & Objectives	5
2.1 Lay summary	5
2.2 Introduction	6
2.3 Dementia	7
2.3.1 Definition and underlying pathologies	7
2.3.2 Diagnostic criteria	8
2.3.3 Public health importance	10
2.3.4 Treatments	11
2.3.5 Risk factors	11
2.4 Serum lipids	12

Contents

2.4.1	Lipid fractions	12
2.4.2	Statins	14
2.4.3	Other lipid regulating agents (LRA)	15
2.5	Evidence for the association between blood lipids and dementia	17
2.5.1	Basic science	18
2.5.2	Observational studies	18
2.5.3	Randomised controlled trials	19
2.5.4	Mendelian randomisation	20
2.6	Theoretical framework: Evidence synthesis	21
2.6.1	Inclusion of preprints	22
2.6.2	Triangulating across study designs	23
2.6.3	Individual patient data meta-analysis	24
2.7	Thesis overview	25
2.7.1	Aims and objectives	25
2.7.2	Structure	26
2.8	Outputs from this thesis	28
2.8.1	Contributions to the scientific literature	28
2.8.2	Presentations/Talks	29
2.8.3	Software	30
2.9	Summary	31
2.10	References	32
3	medrxivr: an R package for systematically searching biomedical preprints	33
3.1	Lay summary	33
3.2	Introduction	34
3.3	Development	35
3.3.1	Success criteria	35
3.3.2	Alternative medRxiv/bioRxiv interfaces	36
3.3.3	Early versions	37

3.3.4	Package infrastructure	37
3.4	Usage	38
3.4.1	Installation	38
3.4.2	Importing preprint metadata	39
3.4.3	Performing a search	40
3.4.4	Refining your search	41
3.4.5	Exporting to a bibliography file	42
3.4.6	Downloading the PDFs of relevant records	43
3.5	Discussion	43
3.5.1	Reception and future plans	43
3.5.2	Use case	44
3.5.3	Limitations of <code>medrxivr</code>	45
3.5.4	Role of open source tools in evidence synthesis	46
3.6	Summary	46
3.7	References	47
4	Systematic review of all available evidence	49
4.1	Lay summary	49
4.2	Introduction	50
4.3	Methods	51
4.3.1	Protocol	51
4.3.2	Contributions	51
4.3.3	Search strategy	51
4.3.4	Study selection	53
4.3.5	Validation of screening process	56
4.3.6	Data extraction	57
4.3.7	Risk of bias assessment	59
4.3.8	Analysis methods	63
4.4	Results	66
4.4.1	Initial search and validation of search filters	66

Contents

4.4.2	Screening results	66
4.4.3	Characteristics of included studies	69
4.4.4	Risk of bias	74
4.4.5	Meta-analysis of statin use and hypercholesterolemia	74
4.4.6	Dose response meta-analysis of lipid levels	75
4.4.7	Publication bias	76
4.4.8	Added evidential value of including preprints	77
4.4.9	Risk of bias due to missing evidence	77
4.5	Discussion	77
4.5.1	Summary of findings	77
4.5.2	Previous reviews	79
4.5.3	Inclusion of preprints	80
4.5.4	Reviewing Mendelian randomisations studies	83
4.5.5	Comments on the process	84
4.5.6	Strengths and limitations	86
4.6	Conclusions	88
4.7	References	89
5	Primary analysis of lipid regulating agents and dementia outcomes	91
5.1	Lay summary	91
5.2	Introduction	92
5.3	Methods	93
5.3.1	Study protocol	93
5.3.2	Data source	93
5.3.3	Cohort definition	94
5.3.4	Exposures	95
5.3.5	Outcomes	96
5.3.6	Covariates	97
5.3.7	Missing data	98
5.3.8	Estimation methods	99

5.3.9	Immortal time bias and time-varying treatment indicators	100
5.3.10	Time axis	102
5.3.11	Sensitivity analyses	102
5.4	Results	105
5.4.1	Patient characteristics	105
5.4.2	Missing data	110
5.4.3	Primary analysis	111
5.4.4	Sensitivity analyses	113
5.5	Discussion	121
5.5.1	Summary of findings	121
5.5.2	Interpretation of results	121
5.5.3	Strengths and limitations	124
5.5.4	Enabling easy synthesis of this analysis	125
5.5.5	Conclusions	125
5.6	References	127
6	Individual participant data meta-analysis	129
6.1	Lay Summary	129
6.2	Introduction	129
6.3	Methods	130
6.3.1	Data sources	130
6.3.2	Risk of bias assessment	132
6.3.3	Data cleaning	132
6.4	Discussion	132
6.4.1	Reflections on data access	132
6.4.2	Study within a review	133
7	Discussion	135
7.1	Triangulation across existing evidence and evidence produced in this thesis	135

7.1.1	Overview of triangulation	135
7.2	Qualitative vs quantitative triangulation	136
7.2.1	Researcher positionality statement	136
7.3	Discussion of E-values	136
7.4	Use of consensus panel to bring everything together	136
7.5	Additional topics for the discussion	136
7.6	Summary of findings (and implications for policy makers)	137
7.7	Strengths and Limitations	137
7.8	Reproducible research	137
7.9	Public involvement and engagement	138
7.10	Future work	138
7.11	Overall conclusions	138
8	Bibliography	139
Appendices		
A	By Chapter	155
A.1	Chapter 2	155
A.1.1	Publications beyond the scope of this thesis	155
A.1.2	Involvement of patients and the public	156
A.2	Chapter 4	156
A.2.1	Search strategy	156
A.2.2	Web of Science Databases Searched	160
A.2.3	Code to search preprints	160
A.2.4	MR risk of bias tool	162
A.3	Chapter 5	164
A.3.1	Amendments to protocol	164
A.3.2	Code lists	164
A.4	Chapter 6	164
A.4.1	Section 1	164

Contents

A.4.2	Section 2	164
A.5	Chapter 7	164
A.5.1	Catalogue of failures	164
B	Other Appendix	165
B.1	Software used to create this thesis	165
B.2	Producing risk-of-bias visualisations with <code>robvis</code>	165
B.2.1	Introduction	165
B.2.2	Development	167
B.2.3	Installation	167
B.2.4	Usage	168
B.2.5	Reception and Future Plans	170
B.3	Creating	171
B.4	Copies of papers arising from this thesis	171

List of Figures

2.1	Statin mechanism of action	14
3.1	Role of <code>medrxivr</code> in a systematic review workflow	35
3.2	Overview of <code>medrxivr</code> data sources	40
4.1	PRISMA flow diagram	67
4.2	Geographical distribution of study cohorts	74
4.3	Forest plot showing effect of statins	75
4.4	Funnel plot of results examining the relationship between statins and any dementia	76
4.5	Study designs by year of publication	78
4.6	Growth of preprint repositories over time	82
5.1	Decision tree for assigning dementia subtypes	97
5.2	<code>shortcap</code>	101
5.3	DAG of ideal negative outcome	103
5.4	Attrition of CPRD participants	106
5.5	Results from primary analyses of CPRD data	112
5.6	Complete case vs. imputed data analysis	114
5.7	Results of control outcome analysis	115
5.8	Comparison of different combinations of covariates	116
5.9	Sensitivity analysis	117
5.10	Comparison of pregnancy analysis	118
5.11	Analysis stratified by statin properties	119
5.12	Comparison of code used in this analysis with those used in the Smeeth et al. 2010 paper	120

List of Figures

5.13 Confounding by indication causal diagram	122
B.1 Example risk of bias traffic light plot created using ‘robvis’	168
B.2 Example risk of bias summary plot created using ‘robvis’ and the example ROB2 dataset	169

List of Tables

2.1	Overview of the DSM-5 criteria for dementia and vascular dementia.	9
2.2	Classification of blood lipid levels	13
2.3	Overview of common statins	15
2.4	Summary of available treatments for hyperlipidaemia.	17
2.5	Summary of studies included in this thesis	27
4.1	searchOverview	52
4.2	Ranges for Gwet's AC1	57
4.3	Inter-rater agreement	68
4.4	Inter-rater agreement	68
4.5	studyCharacteristics	71
5.1	Example of CPRD Read code hierarchy	94
5.2	Patient characteristics by drug class	109
5.3	Participants who stopped, switched or added treatments by initial treatment type	110
5.4	Frequency of diagnoses by grouped year of cohort entry	117
5.5	Summary of statin properties (lipophilicity vs hydrophilicity).	119
A.1	searchHits	156
A.2	Summary of Web of Science databases searched.	160
A.3	Mendelian randomisation risk-of-bias assessment tool	162
B.1	Description of the arguments available in the two main 'robvis' functions. 'X' indicates that the option is available for the respective function.	169

List of Abbreviations

API	Application programming interface
AzD	Alzheimer's disease
CIND	Cognitive impairment not dementia
CPRD	Clinical Practice Research Datalink
CRAN	Comprehensive R Archive Network
DOI	Digital object identifier
HDL	High density lipoprotein
IPD	Individual participant data
LDL	Low density lipoprotein
MCI	Mild cognitive impairment
MMSE	Mini Mental State Exam
MoCA	Montreal Cognitive Assessment
MR	Mendelian randomization
NOS	Newcastle-Ottowa Scale
PDF	Portable document format
RCT	Randomised controlled trial
TG	Triglycerides
VaD	Vascular dementia

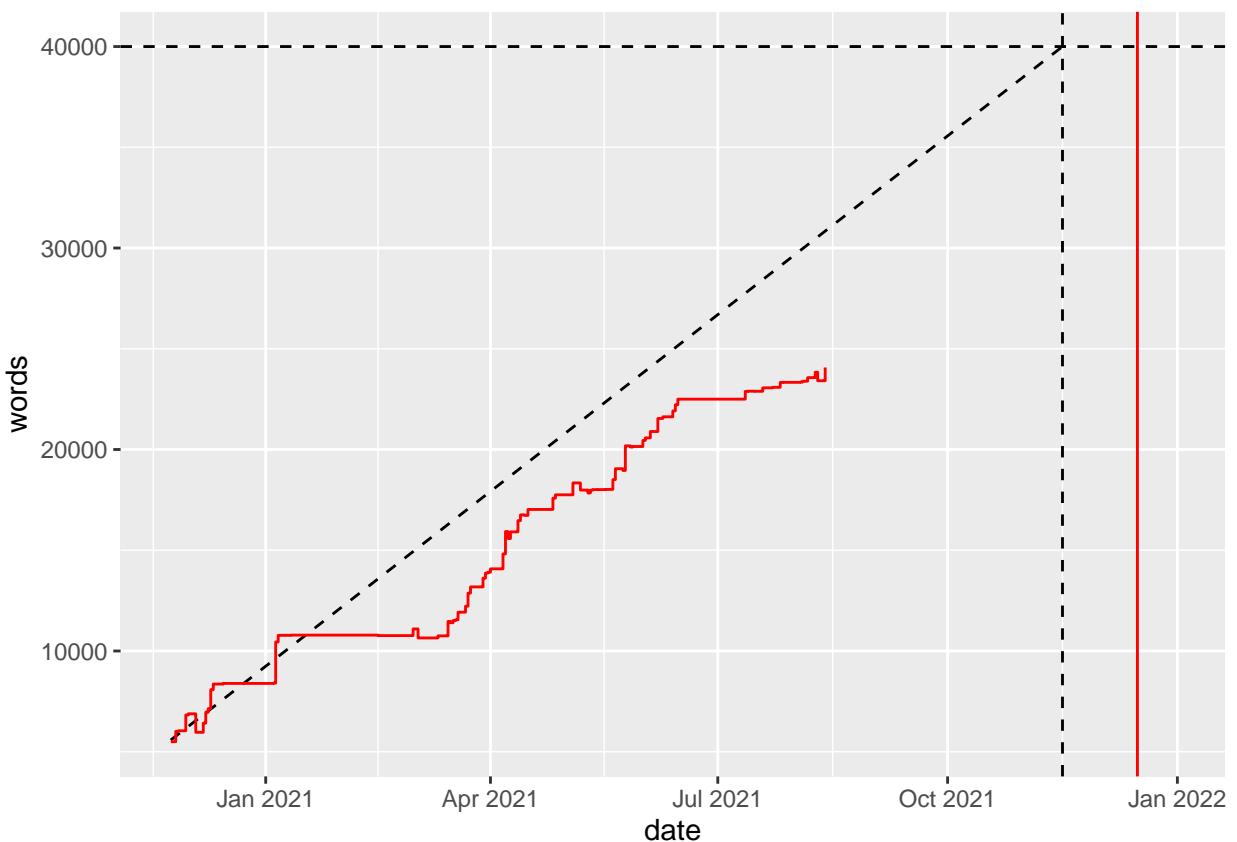
Covering material

Word count: 24069

Days: 262

Words behind: -6925

Words today: 656



1

Introduction

Contents

1.1 Additional ideas	3
--------------------------------	---

1.1 Additional ideas

1.1.1 General

- Explore difference/similarities between the published and unpublished literature, potentially formally, using funnel plots and also by following up preprints to see if they are eventually published. Limit to the date we ran the original search and see if any study included in the study had not been published by the time the review was finished.
- Compare and contrast the codes used to find the AzD cases in the previous study and in this study, potentially with a view to contrasting misclassification between the two.
-

1.1.2 Background

1.1.3 medrxivr

1.1.4

2

Background, Theoretical framework, Aims & Objectives

Contents

2.1	Lay summary	5
2.2	Introduction	6
2.3	Dementia	7
2.4	Serum lipids	12
2.5	Evidence for the association between blood lipids and dementia	17
2.6	Theoretical framework: Evidence synthesis	21
2.7	Thesis overview	25
2.8	Outputs from this thesis	28
2.9	Summary	31
2.10	References	32

2.1 Lay summary

[Note: see Section 2.7.2 for details on the layout of each chapter.]

Around 850,000 people in the UK live with dementia, and by 2040, nearly twice as many will have the condition. Despite many promising candidates, no cure for dementia currently exists, meaning the focus is on finding ways to prevent the

condition. The best way to do this is to find risk factors (characteristics that influence a person's chance of developing a disease) for dementia that we can easily change. Avoiding a risk factor does not guarantee that a person will not develop dementia but makes it less likely. A key risk factor for dementia may be the levels of lipids (fatty substances such as cholesterol) in a person's blood, though not all existing research agrees. The aim of this thesis is to use all available evidence to assess whether blood lipids levels are in fact a risk factor for dementia.

This introductory chapter provides background information on both dementia and blood lipids, and on the potential link between them. It introduces the theory used to frame the research presented here, and then maps the formal aims and objectives of the research project to the relevant chapters of this thesis. Finally, it summarises the outputs (journal articles, presentations and software) that were created as part of this thesis.

2.2 Introduction

This chapter provides an overview of the broad context of this thesis, introducing the core concepts used throughout and providing some background on each. It briefly discusses the underlying pathologies and diagnosis of dementia, its public health importance, and the current state of treatment and prevention research. It then provides background on blood lipids and lipid-modifying treatments, and summarises the types of evidence used to examine the effect of these exposures on dementia outcomes.

The chapter introduces evidence synthesis as the key framework used to guide the research presented in the remaining chapters. Finally, it outlines the aims, objectives and structure of this thesis, and briefly summarises the contributions to the scientific literature that arose from this research.

2.3 Dementia

2.3.1 Definition and underlying pathologies

Defined by the Diagnostic and Statistical Manual of Mental Disorders as a “major neurocognitive disorder”, dementia is a progressive disease which impairs cognitive functions including speech, memory and executive reasoning.^{edition2013diagnostic?} At advanced stage, the condition causes severe behavioral and personality changes,¹ cumulating in reduced motor control that affects patients ability to swallow or breathe.² The condition has several distinct underlying causes, including Alzheimer’s disease and vascular dementia.³

Alzheimer’s disease is the most common cause of dementia, accounting for approximately 60-80% of cases. Characterised by substantial cognitive impairment and difficulty with high level executive function to the extent that it interferes with, it is an insidious disease, within initial onset thought to occur up to 15 years prior to symptomatic presentation.⁴ Much remains unknown about Alzheimer’s pathogenesis, despite research implicating the “amyloid hypothesis”,⁴ as a potential mechanism of disease. Under this hypothesis, the build-up of amyloid plaques (composed mainly of amyloid- β peptide) and neurofibrillary tangles (composed mainly of tau protein) triggers a range of physiological changes, including inflammation and cell death, that result in cognitive impairment.⁴

Vascular dementia (VaD) is the second largest underlying pathology of dementia, accounting for ~10% of cases. Vascular dementia is caused by a range of cerebrovascular disorders, and as a result, presentation of symptoms can vary widely.⁵ Similarly, due to the varied underlying pathophysiology, vascular dementia can onset either quite rapidly following a cerebrovascular event such as a stroke or over a long time-frame due to a series of small infarcts.⁶ Vascular dementia regularly co-occurs in patients with Alzheimer’s disease.⁵ This presentation is described as “mixed” dementia,⁷ and occurs in approximately 25% of cases.³

2.3 - Dementia

The remaining 10-30% of cases are caused by other dementia subtypes (e.g. Lewy body dementia, frontotemporal dementia) or by progression of other neurological diseases (e.g. Parkinson's disease).³

2.3.2 Diagnostic criteria

Dementia is difficult to diagnose, primarily due to its slow onset, in addition to the confusion of initial symptoms with normal ageing.⁴ Dementia is diagnosed on the basis of behavioral and cognitive changes as assessed by an experienced clinician, using one of several diagnostic criteria.

Table 2.1: Overview of the DSM-5 criteria for dementia and vascular dementia.^{edition2013diagnostic?}

Criterion	Major neurocognitive event (previously dementia)
A	Evidence of significant cognitive decline from a previous level of performance in one or more cognitive domains: [*] - Learning and memory - Language - Executive function - Complex attention - Perceptual-motor - Social cognition
B	The cognitive deficits interfere with independence in everyday activities. At a minimum, assistance should be required with complex instrumental activities of daily living, such as paying bills or managing medications.
C	The cognitive deficits do not occur exclusively in the context of a delirium.
D	The cognitive deficits are not better explained by another mental disorder (eg, major depressive disorder, schizophrenia).

* From DSM: Evidence of decline is based on concern of the individual, a knowledgeable informant, or the clinician that there has been a significant decline in cognitive function and a substantial impairment in cognitive performance, preferably documented by standardized neuropsychological testing or, in its absence, another quantified clinical assessment.

One of the most commonly used criteria are those found in the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria (Table 2.1).^{edition2013diagnostic?}

These criteria are outlined in Table 2.1, and form the broad definition of a dementia diagnoses, supported by a detailed patient history, evidence from carers and family members, and objective assessments of cognitive ability using neurocognitive tests.

Many cognitive assessment tools exist for the purpose of informing a diagnoses of dementia,⁸ with two of the best known of these being the Mini Mental State

Exam (MMSE) and Montreal Cognitive Assessment (MoCA) scale. The distinction between these memory scales and diagnostic criteria presented above should be noted. For example, the MMSE is used to provide evidence for part A of the criteria presented in 2.1. Taken alone, it does not indicate the absence or presence of dementia, instead merely indicating cognitive impairment which could be due to another cause (for example, temporary delirium as a result of an infection or surgery).

Differentiating between the underlying causes of a dementia diagnosis is challenging but necessary, as whether the patient has Alzheimer's disease or vascular dementia will affect expected progression and potential treatment options available (see Section 2.3.4). Cause-specific criteria exist for the diagnosis of dementia subtypes. For example, the NINCDS-ADRDA criteria are commonly used to assess patients for Alzheimer's disease,⁹ while vascular dementia is diagnosed using the NINCDS-AIREN criteria.^{roman1993vascular?}

2.3.3 Public health importance

Dementia is quickly becoming a critically important public health issue. Despite the age-specific incidence and prevalence of dementia remaining relatively constant over time,¹⁰ an ageing population is set to create a dementia epidemic, particularly in Westernised countries.¹¹ While approximately 525,000 patients have received a dementia diagnosis, the true number of people currently living with dementia in the UK is thought to be closer to 850,000, with this figure expected to double by 2040.¹² Globally, the prevalence of dementia is expected to reach 75 million by 2030.¹⁰ Dementia is the leading cause of death in the UK, and the only one without a proven cure.

Dementia also has a substantial economic impact. In 2015, the estimated total cost of dementia in England was £24.2 billion. Health care costs alone were £3.8 billion,¹³ Globally, the cost of dementia care is expected to rise to \$1tr by 2030.^{prince2014dementia?}

As such, the urgent need to reduce the burden of dementia, both at the personal and system level, is clear.

2.3.4 Treatments

Developing treatments for dementia is regularly deemed to be one of the hardest markets in the pharmaceutical world, with trials of seemingly promising therapeutics being regularly abandoned due to futility.¹⁴. At present, there are no known curative treatments for dementia, regardless of the underlying cause, though several available therapeutics can help alleviate the symptoms of Alzheimer's disease.

The most common of these are acetylcholinesterase (ACE) inhibitors, which inhibit the degradation of the neurotransmitter acetylcholine by competitively binding the ACE enzyme. Acetylcholine plays a key role in controlling the cholinergic synapses, which are highly concentrated in regions of the brain (such as the neocortex) that control higher level brain functions such as memory and attention.¹⁵ Commonly prescribed ACE inhibitors include donepezil and galantamine.¹⁶ ACE inhibitors increase the availability of the neurotransmitter, and has shown clinical effect in easing the behavioural and memory-related symptoms of Alzheimer's disease.¹⁷ ACE inhibitors represent only a stop-gap treatment, treating the symptoms rather than the underlying pathology which may continue to progress.¹⁸

2.3.5 Risk factors

Given the substantial burden that dementia represents and the absence of any curative therapies, as detailed in the above sections, the assessment of easily modifiable targets for their utility in the prevention of dementia should be prioritized. ^{winblad2016a?}

To date, a substantial amount of research has been produced examining putative risk factors for dementia.¹⁹⁻²¹

The benefits of a prevention-based approach based on addressing these risk factors are well-studied. Reducing the prevalence of the seven most important risk factors for dementia (obesity, hypertension, ^{hughes2020association?} diabetes, smoking, physical

inactivity, and low educational attainment) by 10-20% per decade is estimated to result in a reduction in dementia prevalence of 8-15% by 2050.²²

In this context, lipid levels represent a promising target for preventative treatment, due to the ready availability of lipid-modifying treatments which could be repurposed.^{pushpakom2019a?} Determining whether variations in lipid levels are causative for dementia may prove critical in reducing the future burden of the condition.

This thesis will focus on blood lipids as the primary risk factor of interest. The next section provides an overview of blood lipid fractions and therapeutic interventions that modify them, while Section 2.5 provides an overview of the existing evidence for an association between lipids and dementia outcomes.

2.4 Serum lipids

2.4.1 Lipid fractions

The blood lipid profile contains a range of component parts, or fractions. However, this thesis will only consider the two most important fractions, triglycerides (TG) and cholesterol, which are either absorbed from food (exogenous lipids) or produced internally (endogenous lipids).¹⁹

Triglycerides are the simplest and most common type of lipids found across the body. They are used to store unused calories from food, and to move energy around the body.²³ In contrast, cholesterol is primarily used to create cell walls and certain sex hormones.²⁴ As lipids are not water soluble, within the blood stream, cholesterol is transported in lipoprotein structures of varying densities. Low-density-lipoprotein cholesterol (LDL-c), commonly known as the “bad” cholesterol, transports fat to cells, acting as an energy conveyor. In contrast, High density-lipoprotein cholesterol (HDL-c), transports cholesterol to the liver to be broken down and excreted.¹⁹

In addition to the individual fractions, total serum cholesterol (TC) is a commonly-used summary measure to estimate the total amount of lipid present in the blood.

The measure is derived from measurements of the individual HDL-c, LDL-c and TG levels using the Friedwald formula:²⁵

$$TC \approx LDLc + HDLc + kTG \quad (2.1)$$

where k is 0.20 if measurements are in milligrams per decilitre (mg/dl) and 0.45 if measured in millimole per litre ($mmol/l$).

Widely-used ranges for the acceptable levels of different types of lipids are based on the National Cholesterol Education Program (NCEP)²⁶, and are outlined in Table 2.2.

Table 2.2: Classification of blood lipid levels according to the National Cholesterol Education Program guidelines.²⁶

Fraction	Measure (mg/dL)	Classification
LDL cholesterol	<100	Optimal
	100-129	Near/above optimal
	130-159	Borderline high
	160-189	High
	>190	Very high
HDL cholesterol	<40	Low
	>60	High
Triglycerides	<150	Normal
	150-199	Borderline high
	200-499	High
Total cholesterol	>500	Very high
	<200	Desirable
	200-239	Borderline high
	>240	High

Elevated LDL-c in the bloodstream, a condition also known as hypercholesterolaemia or hyperlipidaemia,²⁷ can lead to atherosclerosis,²⁸ the build-up of fatty deposits in the blood vessels. These deposits constrict blood flow and can lead to vascular complications. Alternatively, part of the deposit can detach from the artery walls, forming a clot that can lead to a heart attack or stroke.²⁸ Globally, the prevalence of elevated cholesterol was estimated by the World Health Organization to be approximately 40%.

2.4.2 Statins

Statins are a commonly prescribed method of lipid regulation.²⁹ Statins inhibit the conversion of 3-hydroxy-3-methylglutaryl-coenzyme-A (HMG-CoA) into mevalonate, by competitively binding with HMG-CoA reductase (HMG-CoA-R). This conversion is a key rate-limiting step in the cholesterol biosynthesis pathway (see Figure ??), enabling statins to reduce effectively the production of LDL cholesterol.

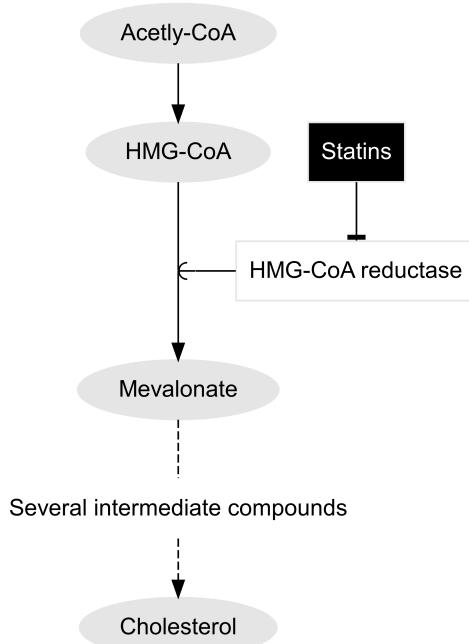


Figure 2.1: Overview of statins mechanism of action, inhibiting HMG-CoA reductase which controls the conversion of HMG-CoA to mevalonate, the rate-limiting step in cholesterol biosynthesis.

Several statin treatments have been widely available for some time (see Table 2.3). Depending on the statin and dosage prescribed, the average reduction in LDL-c concentrations ranges from 15% with low-intensity regimen (e.g. ravastatin 5 mg/day) up to 60% with a high-intensity regimen (e.g. rosuvastatin 80 mg/day).^{30,31} Statins also vary with regard to their lipophilicity (the extent to which they are lipid soluble), affecting their localisation within the body, with hydrophilic statins being concentrated in the liver and lipophilic statins circulating more widely.³² This may create a divide in the pleiotropic affects of statins with differing lipophilicity, particularly given the ability of lipophilic statins to permeate the blood brain barrier.³³

Table 2.3: Overview of commonly-prescribed statins, summarising their approval date (US), properties and lipid-lowering effect.

Name	Brand name	Year approved	Properties	Lipid-lowering effect
Atorvastatin	Lipitor	1996	Lipophilic	+++
Pravastatin	Lipostat	1989	Hydrophilic	+
Rosuvastatin	Crestor	2003	Hydrophilic	++++
Simvastatin	Zocor	1992	Lipophilic	++

2.4.3 Other lipid regulating agents (LRA)

There are several other interventions that can be used to modify a persons lipid profile, which each acting in slightly different ways (Table ??). However, in general, these treatments are either used as adjunct (additional) treatments with statins therapy or are used in situations where statins are contra-indicated or not tolerated.

2.4 - Serum lipids

The most commonly used non-statin therapeutic is ezetimibe,³⁴ which prevents intestinal absorption of cholesterol. However, when used alone, it has a limited LDL-c lowering effect, leading to the creation of combined statin/ezetimibe therapies (both compounds contained in a single pill, as opposed to complimentary treatments).³⁵

Fibrates provide a second example of non-statin therapy. They are used to treat hypertriglyceridaemia by reducing production of triglyceride carrying compounds in the liver. They are commonly used in patients with mixed hyperlipidaemia if treatment with statins has failed to sufficiently control cholesterol levels.

Finally, PCSK9 inhibitors (or PCSK9i) are a relatively new treatment with strong lipid lowering effects, lauded as a potential alternative to statins.³⁶ Their mechanism of action is to bind to and inhibit PCSK9, which breaks down LDL-c receptors on the surface of the liver, thus allowing more LDL-c to be internalised and broken down.

Other therapies targeting triglycerides exist, including nicotinic acids³⁷ and omega-3-fatty acids,³⁸ but they far less effective in LDL-c lowering than the therapies described above.

2.5 - Evidence for the association between blood lipids and dementia

Table 2.4: Summary of available treatments for hyperlipidaemia.

Treatment	Effect	Mechanism of action	Examples
HMG CoA reductase inhibitors (statins)	Lowers LDL-c & TG Raises HDL-c	Inhibits cholesterol biosynthesis pathway in the liver	Atorvastatin, Simvastatin, Pravastatin
Ezetimibe	Lowers LDL-c	Prevents absorption of cholesterol from diet	
Bile acid sequestrants	Lowers LDL-c	Prevent bile acid reabsorption in the gastro-intestinal tract, increasing conversion of cholesterol to bile acids	Colestipol
Proprotein convertase subtilisin kexin 9 (PCSK9) inhibitors	Lowers LDL-c	Bind to PCSK9 protein, preventing it from breaking down LDL receptors on hepatic cells, increasing cholesterol uptake	Evolocumab, Alirocumab

2.5 Evidence for the association between blood lipids and dementia

This section provides an overview of the varying sources of evidence on the relationship between blood lipid levels and dementia risk.

2.5 - Evidence for the association between blood lipids and dementia

2.5.1 Basic science

A role for lipids in the aetiology of the dementia is supported by both genetic linkage studies and functional cell biology studies. The generation of the amyloid plaques found in the brains of Alzheimer's patients is cholesterol dependent,^{39,40} while the most established genetic risk factor for late-onset dementia, apolipoprotein E (ApoE), is involved in cerebral cholesterol transport. Several other genes involved in cholesterol transport have also been found to be associated with increased AD susceptibility.⁴¹⁻⁴³

Despite these results, evidence from the diverse range of epidemiological studies on this topic has been inconclusive.

2.5.2 Observational studies

By far the largest source of evidence on the relationship between comes from observational designs. Several studies have examined the relationships between concentrations of serum lipids (total cholesterol (TC), low density lipoprotein cholesterol (LDL-c), high density lipoprotein cholesterol (HDL-c) and triglycerides) and both Alzheimer's disease and vascular dementia and reported extremely varied results. In some studies, a high serum cholesterol concentration has been found to be associated with an increase in susceptibility to AD,⁴⁴⁻⁴⁸ however others have shown no association,^{49,50,li2005a?,tan2003a?} or a reduced susceptibility.^{51,reitz2004a?}

With regards vascular dementia, decreased levels of HDL-c appear to be associated with increased risk,^{reitz2004a?} while for LDL-c, studies have reported both positive and negative associations.^{52,reitz2004a?}

Several previous systematic review of observational studies examining the effect of lipids⁵³ and lipid-regulating agents^{chu2018b?,poly2020c?} on dementia outcomes have been performed. However, these reviews have several limitations. Many did not consider grey literature sources (see Section 2.6.1). Additionally, many of the

2.5 - Evidence for the association between blood lipids and dementia

reviews of observational studies did not perform any risk-of-bias assessment^{chu2018b?} or used an outdated assessment tool.^{53,poly2020c?}

2.5.3 Randomised controlled trials

In terms of the central research of this thesis, RCTs of statin therapy can be used to provide indirect evidence for the effect of reducing blood LDL-c levels on dementia risk.

However, RCTs may be infeasible if the outcome of interest is one with a long prodromal period, such as dementia (see Section 2.3.1), as they would require extremely long and costly follow-up.⁵⁴ It is no surprise then that the two previous trials providing evidence on the effect of statins on dementia risk, identified by a recent Cochrane review,^{mcguinness2016a?} are in fact trials of statins for the prevention of coronary related outcomes.

While being widely cited, these studies have major limitations that reduce their utility as a source of evidence on the effect of statin treatment on assessing the impact of lipid-lowering treatment on dementia risk. Firstly, there was no clinical cognitive evaluation of patients to determine a dementia outcome. One of the trials, the Prospective Study of Pravastatin in the Elderly (PROSPER) trial,⁵⁵ reported not on dementia outcomes but on the change in cognitive scores over a mean of 3.2 years. As highlighted in Section 2.3.2, a “change in score” alone is insufficient to diagnose a dementia outcome. The second trial, the Medical Research Council/British Health Foundation Protection Study,^{2002?} found no effect of simvastatin on dementia (OR: 1.00, 95%CI: 0.61-1.65), but did not report how the outcome was assessed/recoded within the trial.

Additionally, the two trials did not make any effort to assign an underlying pathology to each case, instead reporting an all-cause dementia outcome. As discussed in Section 2.3.1, the different underlying pathology of dementia have different mechanisms of action, and so it is not guaranteed that the effect of statins would be consistent across them.

2.5 - Evidence for the association between blood lipids and dementia

Both trials were also limited by the relatively short follow-up period examined, expected when the primary outcome of the trials were coronary related conditions rather than dementia.^{55,2002?} The PROSPER trial had a mean follow-up of 3.2 years, while the MRC/BHF Protection Study estimated risk at 5 years of follow-up. Given the long lag time between non-symptomatic onset of dementia and clinical presentation, it is likely that these durations are insufficient to fully capture the onset of dementia. Finally, as they included only patients at high vascular risk, their generalisability to other settings is limited.^{mCGUINNESS2016b?}

2.5.4 Mendelian randomisation

Newer methodological approaches, such as Mendelian randomisation (MR),⁵⁶ have also been used to examine the effect of varying lipid levels on dementia risk in an effort to combat the risk of reverse causation and residual confounding inherent to observational studies.⁵⁷ In brief, MR uses genetic variants that are both strongly associated with the exposure of interest and are independent from potential confounders to strengthen causal inference.⁵⁶ The analytic method relies on several assumptions about the instrumental variable (IV),⁵⁸ namely that:

1. the IV is associated with the exposure of interest (the relevance assumption);
2. the IV and outcome do not share a common cause (the independence assumption); and
3. the IV does not affect the outcome other than via the exposure (the exclusion restriction assumption).

Recent MR studies indicated that genetically determined low levels of LDL-c may cause a reduction in AD risk.^{59,60} However, the effect was attenuated in sensitivity analysis that exclude the region surrounding the ApoE gene, the strongest known risk factor for Alzheimer's disease.⁶¹ Inclusion of ApoE4 variants invalidates the exclusion restriction criteria (Assumption 3, above), as the risk reduction observed may be driven by variants in this region via a pathway independent of

2.6 - Theoretical framework: Evidence synthesis

lipid levels. This was supported by further MR studies where *ApoE4* variants were intentionally excluded.⁶²

Despite the increasing number of MR studies examining this topic, no systematic review of this study design as a source of evidence has been performed.

In summary, multiple sources of evidence exist on the relationship between statins and dementia. In the next section, I introduce the synthesis of diverse sources of evidence as the theoretical framework used in this thesis.

2.6 Theoretical framework: Evidence synthesis

Evidence synthesis is the process of finding and integrating information from several sources to examine a research question.⁶³ A common type of evidence synthesis is a systematic review, either with or without a meta-analysis.⁶⁴

The results of an evidence synthesis exercise can be used to provide a more definitive answer to that question or, failing that, to highlight gaps in the existing evidence base. The ability to identify these gaps is particularly useful in guiding future research to address questions that have yet to be answered.

This thesis seeks to use an evidence synthesis framework to assess the effect of lipids, and treatments that influence lipid levels, on dementia outcomes. Specifically, this thesis considers three concepts within the umbrella term of evidence synthesis:

- Inclusion of preprints
- Triangulation across evidence sources
- Individual patient data meta-analysis

These three elements are expanded on below and are used to frame the research presented in the subsequent Chapters.

2.6.1 Inclusion of preprints

The importance of including grey (or gray) literature in systematic reviews is widely acknowledged. Meta-research studies have demonstrated that systematic reviews excluding grey literature sources overestimate the effect of interventions.^{65–67} Common, well-accepted forms of grey literature include conference abstracts and theses.⁶⁸

A important developing source of grey literature are preprints. Defined by the Committee on Publication Ethics (COPE) as ‘scholarly manuscript[s] posted by the author(s) in an openly accessible platform, usually before or in parallel with the peer review process’⁶⁹, preprints serve several purposes. They are used to establish primacy when submitting to a journal where the peer-review process may take several months,⁷⁰ to rapidly disseminate research findings, as occurred during the COVID-19 pandemic,^{fraser2020a?} and to make available publications that may not have been accepted elsewhere in an attempt to combat publication bias or the “file-drawer” effect.⁷¹

One of the major criticisms of using preprints as an evidence source is that they have not yet undergone formal peer review.^{72,73} However, this approach assigns substantial weight to peer-review as a indicator of “quality”, and is at odds with the acceptance of non-reviewed conference proceedings as an evidence source.^{68,74} The argument for including preprints as an evidence source is further strengthened by results that demonstrate preprinted studies seldom change following peer review. Meta-studies of the concordance between preprinted and published studies showed that results were broadly comparable between the two, indicating that while the numerical results may change, the overall interpretation of the results were consistent in the majority of cases.^{75,76,shi2021a?} This indicates that preprints should be considered a reliable reflection of a given study.

In this thesis, preprints are considered an important source of evidence, in contrast to previous reviews on this topic. However, as with many sources of grey literature,⁷⁴ there are several logistical issues with carrying out systematic searches in preprint

repositories. As such, to enable the inclusion of preprints in the systematic review described in Chapter 4, a new tool addressing these issues is presented in Chapter ??.

2.6.2 Triangulating across study designs

As illustrated in Section 2.5, several diverse epidemiological methods have been used to examine the effect of varying blood lipid levels on dementia risk. However, each method is limited by its own biases. Aetiological triangulation is a developing evidence synthesis method that seeks to exploit these inherent differences in study design, and as a result, in biases.^{lawlor2016a?} If several sources of evidence are available and point towards identical conclusions about an exposure-outcome relationship, and these sources are at risk of unrelated biases, this strengthens our confidence in the result. The ideal scenario is where predicted sources of bias are likely to be in competing directions, strengthen the effect of the exposure and the other to attenuate it.^{lawlor2016a?} As such, triangulating these results can provides us a middle-ground between the competing directions of bias. A triangulation approach can also prove useful in a prospective manner, helping to design new studies that are at risk of different sources of bias to that already available from the published literature.⁷⁷

This thesis seeks to apply a triangulation approach to provide the best available evidence on the effect of lipids, and lipid regulating agents, on dementia outcomes.

All existing evidence, regardless of study design, is first identified by the by the systematic review presented in Chapter 4. Risk-of-bias assessment using a domain-based tool is already a recommended part of the systematic review process, but is particularly important to a triangulation exercise.^{78,79,sterne2019a?} As such, a core component of the review is a comprehensive domain-based risk-of-bias assessment for all included studies.

Finally all evidence, both pre-existing and produced as part of this thesis (Chapter 5 and ??), are triangulated in Chapter ??.

2.6.3 Individual patient data meta-analysis

Individual patient data meta-analyses are commonly held to represent the gold standard in evidence synthesis methodology.^{80,81} IPD methods seek to obtain the raw data from each study identified in a systematic review, rather than basing the meta-analysis on summary results extracted from the literature.⁸⁰

In the context of this thesis, if lipids are found to have a causal role in development of dementia, evidence-based preventative strategies would be best informed by identifying the types of individuals who are most likely to receive benefit from treatment with lipid-modifying agents.⁸²⁻⁸⁴ However, if primary studies do not present results stratified by covariates of interest, meta-analyses of summary-level data on this topic often have limited ability to examine research questions related to exposure-covariate interactions.⁸⁰ In terms of this thesis, patient sex is considered to be of particular interest.^{82,85}

An IPD meta-analysis of lipid levels on dementia outcomes would overcome this limitation of summary-level data, as access to the raw data allows for an analysis that investigates these interactions.^{riley2020?} This approach has the added benefit of allowing a common set of inclusion criteria and statistical model to be applied across all datasets, potentially eliminating some important sources of heterogeneity.⁸⁶

Despite their advantages, IPD meta-analysis are rarely performed.⁸⁷ Factors limiting their uptake include the increased time and effort they require when compared to a summary-level analysis, and the low success rate associated with obtaining the raw data.^{88,nevitt2017a?} The data underlying primary studies are frequently not publicly available,^{89,federer2018a?} and the availability of data “available on request from authors” declines rapidly over time.⁹⁰ Several systematic barriers to open data sharing have been identified^{vanpanhuis2014a?}. Of particular concern for biomedical IPD analyses are legal issues surrounding the sharing of medical data, motivated by concerns around patient privacy.⁹¹

In response to these limitations, new collaborative initiatives have developed to enable rapid access to relevant data in a secure supported workshop. The most import in relation to this thesis is the Dementia Platform UK (DPUK),⁹² which aims to provide access to several dementia-related datasets via a single simplified application process.

I will attempt to obtain the raw data from relevant primary studies identified by the systematic review in Chapter 4. Any data obtained will be combined with that available from the DPUK portal as part of an individual participant data meta-analysis in Chapter ??, enabling the assessment of the effect of lipids on dementia stratified by key variables such as sex.

2.7 Thesis overview

2.7.1 Aims and objectives

The over-arching aim of this thesis is to explore the relationship between blood lipid levels, and by extension treatments that modify blood lipid levels such as statins, and the subsequent risk of dementia and related outcomes

The specific research objectives that this thesis seeks to address are:

- To create a tool that allows for the inclusion of health related preprints in evidence syntheses in a systematic and reproducible manner
- To review all available evidence across multiple diverse study designs to assess the effect of lipids and lipid regulating agents on dementia risk
- To examine whether there is evidence for an effect of lipid-regulating agents on dementia and related outcomes in a large scale population-based cohort, the Clinical Practice Research Datalink (CPRD)
- To meta-analyse raw dementia-related datasets as part of a individual participant data (IPD meta-analysis) to produce evidence on exposure-covariate interactions

2.7.2 Structure

Chapters are self-contained, presenting the methods and results of that specific research project. They are bookended by introductory and discussion sections which place the methods and results in context. Each chapter is prefaced by a “Lay” or plain English summary, developed with input from the Patient and Public Advisory Group (see Section 7.9 for a discussion of the group’s involvement and Appendix A.1.2 for more detail on the group).

- **Chapter 2:** Background information on dementia and blood lipid levels. This chapter provides an introduction to the topics covered in this thesis to non-subject area experts, and discusses the motivation for the remainder of the thesis.
- **Chapter 3:** This Chapter introduces a new tool, `medrxivr`, which was used to developed to allow for systematic searches of the health-related preprint repositories.
- **Chapter 4:** This Chapter describes a comprehensive systematic review and meta-analysis of all available evidence on the effect of blood lipids, and interventions that modified blood lipids, on dementia outcomes.
- **Chapter 5:** This Chapter examines the relationship between lipid-regulating agent use and dementia outcomes in the Clinical Practice Research Datalink, a large primary care electronic health record database.
- **Chapter 6:** This Chapter describes an individual patient data analysis of several longitudinal cohort studies to describe the relationship between blood serum lipids and dementia outcomes, stratified by important covariates such as sex.
- **Chapter 7:** This Chapter integrates the diverse evidence identified by, and produced as part of, this thesis. The overall strengths and weaknesses of this project are discussed in detail, and further avenues of research are suggested.

An overview of how all research studies in this thesis can be found in Table 2.5

Table 2.5: Summary of studies included in this thesis, and used as evidence sources in the triangulation exercise performed in Chapter 7.
 Note, Chapter 3 is intentionally not included in this table, as it describes a tool rather than a research study.

Chapter	Research Question	Exposure/ Intervention	Outcome	Contribution to evidence synthesis framework
Chapter 3	Based on the available evidence; (i) are lipid fractions associated with subsequent dementia risk, stratified by subtype? (ii) Are lipid regulating agents associated with subsequent dementia risk, stratified by subtype?	Lipids (HDL-c, LDL-c, TC, TG),	Dementia, stratified by subtype	Provides overview of existing evidence
Chapter 4	Are lipid regulating agents associated with dementia risk in a large scale electronic health record database?	Seven classes of lipid regulating agents	Dementia, stratified by subtype	Provides additional observational data on vascular dementia (under-represented in the literature)
Chapter 5	Are lipid levels associated with dementia risk in an individual participant data meta-analysis?	Lipids (HDL-c, LDL-c, TC, TG)	Dementia, stratified by subtype	Provides additional evidence from unanalysed datasets

2.8 Outputs from this thesis

The outputs of this thesis are detailed below, and include peer-reviewed papers, presentations, and open-source evidence synthesis tools.

2.8.1 Contributions to the scientific literature

During the course of this thesis, I have made several contributions to the scientific literature. Those arising from or directly related to the contents of this submission are presented below.

McGuinness, L. A., and L Schmidt. (2020) medrxivr: Accessing and searching medRxiv and bioRxiv preprint data in R. Journal of Open Source Software 5.54 2651. DOI: 10.21105/joss.02651

A paper introducing the open-source preprint search tool described in Chapter 3. As is common for journal articles describing software, the paper is intentionally short providing only a broad overview of the tool while extensive documentation is available from the project website (see Section 3.2 for more details).

Hennessy, E. A., Acabchuk, R., Arnold, P. A., Dunn, A. G., Foo, Y. Z., Johnson, B. T., Geange, S. R., Haddaway, N. R., Nakagawa, S., Mapanga, W., Mengersen, K., Page, M., Sánchez-Tójar, A. Welch, V., McGuinness L. A. (2021). Ensuring Prevention Science Research is Synthesis-Ready for Immediate and Lasting Scientific Impact. Prevention Science . DOI: 10.1007/s11121-021-01279-8

The experience of extracting data for the systematic review in Chapter 4 inspired a practical guide for researchers. This piece was co-written with Dr. Emily Hennessy (see Author Declarations in the front materials).

McGuinness, L. A., and Higgins J. P. T. (2020) “Risk-of-bias VISualization (*robvis*): An R package and Shiny web app for visualizing risk-of-bias assessments.” *Research Synthesis Method*). DOI: 10.1002/jrsm.1411

The tool used to visualise the risk-of-bias assessments in Chapter 4 has been published in Research Synthesis Methods. See Appendix B.2 for more details on this tool.

McGuinness, L. A., and Sheppard A. L. 2020. “A Descriptive Analysis of the Data Availability Statements Accompanying Medrxiv Preprints and a Comparison with Their Published Counterparts.” *PLOS ONE* 16(5): e0250887. DOI: 10.1371/journal.pone.0250887

Using the tool described in Chapter ??, I lead a “research-on-research” study to assess the concordance between the openness of data availability statements accompanying a sample of medRxiv preprints and their published counterparts.

For information on additional contributions to the scientific literature not directly related to this thesis, see Appendix A.1.1.

2.8.2 Presentations/Talks

“Identifying and triangulating all available evidence on the effect of blood lipids and statins on dementia outcomes”: Poster presentation, Alzheimer’s Association International Conference 2021.

“medrxivr: A new tool for searching for and retrieving records and PDFs from the medRxiv preprint repository”: Accepted oral presentation abstract, Cochrane Colloquium 2020 (note: event was cancelled due to the COVID-19 pandemic)

“On the shoulders of giants”: advantages and challenges to building on established evidence synthesis packages, using the {robvis} package as a case study: Oral presentation, Evidence Synthesis and Meta-Analysis in R Conference (ESMARConf) 2021.

“RoB 2.0: A revised tool to assess risk of bias in randomized trials”: Webinar, co-presented with Dr. Theresa Moore as part of the Evidence Synthesis Ireland Methods Series.

2.8.3 Software

medrxvir

An R package that allows users to easily search and retrieve bibliographic data from the medRxiv⁹³ and bioRxiv⁹⁴ preprint repositories. See Chapter 3 for more details. Install a stable version of the package from the Comprehensive R Archive Network (CRAN), or alternatively install the development version from GitHub, using:

```
# CRAN version  
install.packages("medrxivr")  
  
# Development version  
devtools::install_github("ropensci/medrxivr")
```

robvis

An R package and associated shiny web application that allows users to easily visualize the results of the risk-of-bias assessments performed as part of a systematic review. See Appendix B.2 for more details. Install a stable version of the package from CRAN, or alternatively install the development version from GitHub, using:

```
# CRAN version  
install.packages("robvis")
```

```
# Development version  
devtools::install_github("mcguinlu/robvis")
```

2.9 Summary

This Chapter has provided background information on the core elements of the central research question, framed the research presented in this thesis in the context of an evidence synthesis framework, and described the contributions of this thesis to the scientific literature.

2.10 References

3

medrxivr: an R package for systematically searching biomedical preprints

Contents

3.1	Lay summary	33
3.2	Introduction	34
3.3	Development	35
3.4	Usage	38
3.5	Discussion	43
3.6	Summary	46
3.7	References	47

3.1 Lay summary

Preprints are copies of academic manuscripts that are posted online in advance of being formally published by an academic journal. They represent an important source of scientific literature. A new software program called `medrxivr` was created to allow researchers to find preprints related to their research in a transparent and reproducible way. Development of this tool was an essential part of this thesis, as preprints represent a key source of information needed for the research reported in future chapters.

3.2 Introduction

Preprints represent an increasingly important source of scientific information (see Section 2.6.1). As a result, repositories of preprinted articles should be considered a distinct but complementary information source when reviewing the evidence base as part of a systematic review. The two key repositories in the health science are bioRxiv, established in 2013,⁹⁴ and medRxiv, which evolved to replace the “Epidemiology” and “Clinical Trial” categories of bioRxiv, which launched in 2019.⁹³

Searching these preprints as part of the systematic review described in Chapter 4 was a necessity, as many of the existing reviews on the topic of lipids and dementia have not considered this important source of grey literature. At the time of writing, however, the bioRxiv and medRxiv websites allow only simple search queries as opposed to the often complex Boolean logic (AND/OR/NOT) that information specialists use to query other major databases.^{96,97} Additionally, the best available extraction mechanism for obtaining references for all records returned by a search were to go through each record, one-by-one, downloading individual citations. As the scale of these preprint databases increase, particularly in light of the massive expansion of the medRxiv repository as a result of COVID, this already time-consuming and error-prone method is no longer feasible.

This chapter outlines the development and key functionality of `medrxivr` (version 0.0.5), a tool created to facilitate the searching of medRxiv and bioRxiv preprints. The factors that necessitated the development of this tool in the context of this thesis are outlined, and the use of `medrxivr` in external projects and by other researchers is discussed. As the majority of work on this aspect of this thesis is represented by lines of code (available at <https://github.com/ropensci/medrxivr>) this Chapter is an intentionally short, high-level summary of the work done on this project. The GitHub repository for the `medrxivr` contains a complete record of the development of this tool, including discussion with other members of the systematic review community.

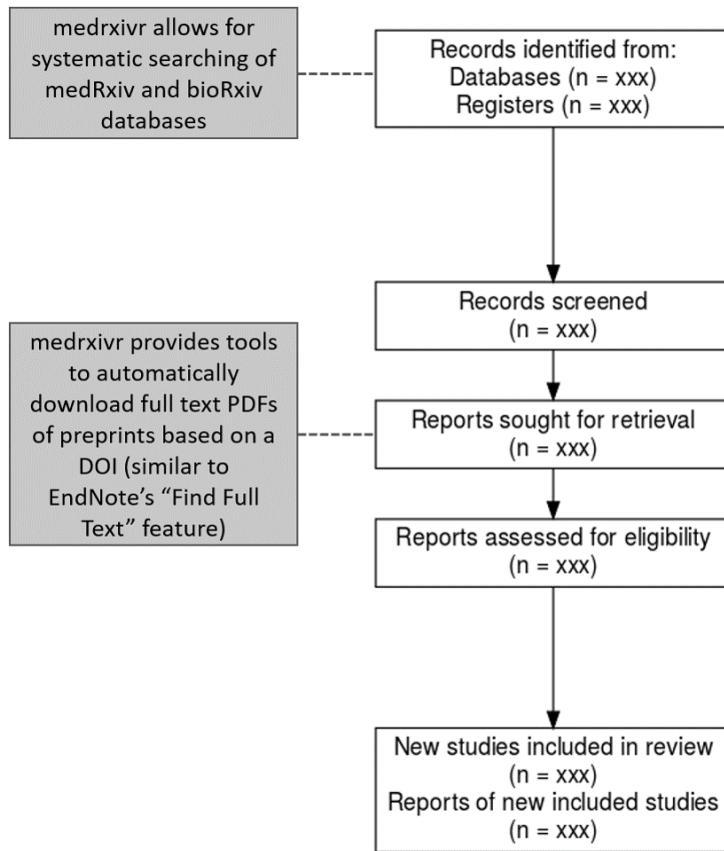


Figure 3.1: Role of `medrxivr` in a systematic review workflow: `medrxivr` allows for systematic searching of biomedical preprints as part of the initial literature searching. Following title and abstract screening, reviewers can then programmatically retrieve a copy of the PDF of included records to facilitate the full-text screening stage (similar to Endnote's "Find Full Text" feature).

3.3 Development

3.3.1 Success criteria

The tool was developed to meet three success criteria,⁹⁸ influenced both by the functionality required to perform systematic searches as part of the review in Chapter 4, discussion with information specialist colleagues, and an informal survey of the evidence synthesis and health librarian communities on Twitter. The criteria were as follows:

1. reliable, reproducible and transparent search functionality, allowing for Boolean (AND/OR/NOT) operator logic;

2. support for bulk export of references returned by the search to a file type that can be readily imported into a reference manager (e.g., *.bib* or *.ris*); and
3. automated retrieval of the full-text PDFs of relevant records, similar to the Find Full Text feature offered by EndNote.

3.3.2 Alternative medRxiv/bioRxiv interfaces

Prior development of this tool, an audit of existing tools for accessing medRxiv and bioRxiv metadata was conducted. While none address the success criteria described above, two of these tools are useful to consider to highlight the additional functionality that `medrxivr` contributes.

The first, a platform called Rxivist,⁹⁹ allows users to search preprints using keywords. However, the core functionality of the Rxivist platform is focused around exploring the number of times a preprint has been downloaded and/or shared on Twitter, to allow researchers to find the most popular papers related to their topic. The search interface¹ does not allow for complex search strategies using Boolean operators and there is no option to batch-export the results of a search.

The second tool, `search.bioPreprint`, allows users to search for terms across a range of preprint servers, including medRxiv and bioRxiv, but also journals which use a post-publication peer-review process such as F1000Research.¹⁰⁰ However, similar to the Rxivist platform, this tool is designed for researchers aiming to keep up to date with recent developments in their fields rather than systematically assess the entirety of the available literature. As such, the platform only returns the most recent 1,000 records by publication date.

Finally, neither tool provides an easy way to programmatically download a copy of the PDF of relevant preprints as part of the preparation for the full-text screening stage of a systematic review.

¹ Available at <https://rxivist.org/>

3.3.3 Early versions

Work on the `medrxivr` tool began in Summer 2019, and initially consisted of a development of set of R scripts to allow for searching medRxiv and bioRxiv as part of the systematic search outlined in Chapter 4. Following interest from other researchers in using the *ad-hoc* web-scraping scripts, additional development work took place in 2019/2020, allowing for improved searching and exporting functionality and the initial version of the `medrxivr` R package was released in February 2020.

Early versions of the tool had a reliance on scraping data directly from the repository website. Web-scraping is a fragile mechanism for extracting data, as it is entirely dependent on consistent website design and underlying code structure remaining unchanged.^{101,102}. In the case of `medrxivr`, constant maintenance work was required to ensure the web-scraping script performed as expected, as the repository website was regularly updated.

However, an Application Programming Interface (API) for the medRxiv and bioRxiv repositories was made public in early 2020 by the institution responsible for managing these preprint repositories, the Cold Springs Harbor Laboratory. This allowed for newer versions of the `medrxivr` package to engage in active “fault prevention” and provide a more robust interface to the data by removing the reliance of web-scraping.¹⁰²

3.3.4 Package infrastructure

The `medrxvir` package was written in R using RStudio,¹⁰³ and followed development best-practice, including detailed documentation, a robust unit testing framework (99% of all code lines within the package are formally tested across multiple

platforms including Windows, MacOS, and Linux), and in-depth code review by two experienced, independent reviewers.

3.4 Usage

The `medrxivr` R package is split into two component parts:

- an interface to the Cold Springs Harbor Laboratory API, which imports medRxiv and bioRxiv metadata into R; and
- a collection of functions for working with the imported metadata, with an explicit focus on searching this data as part of a systematic review or evidence synthesis project.

The standard workflow is to download a copy of all metadata contained in the repository, and then to perform searches on this local copy. This is a workaround as the Cold Springs Harbor Laboratory API does not provide any functionality to search the database.

While the package allows for users to search both medRxiv and bioRxiv, as the process is identical for both, the examples in the sections are limited to the medRxiv repository.

3.4.1 Installation

`medrxivr` has been released to the Comprehensive R Archive Network (CRAN), and can be installed with the following code:

```
install.packages("medrxivr")
```

Alternatively, the development version of the package can be installed from GitHub:

```
# install.packages("devtools")
devtools::install_github("ropensci/medrxivr")
```

3.4.2 Importing preprint metadata

`medrixvr` provides two ways to access medRxiv data, as shown in Figure 3.2. The first, via the `mx_api_content()` function, creates a local copy of all data available from the medRxiv API at the time the function is run.

```
# Get a copy of the database from the live medRxiv API endpoint
mx_data <- mx_api_content()
```

The second, via the `mx_snapshot()` function, provides access to a maintained static snapshot of the database, created each morning at 6am using `mx_api_content()`.

```
# Import a copy of the medRxiv data from the snapshot
mx_data <- mx_snapshot()
```

Allowing users of `medrxivr` to access a maintained snapshot removes any dependency on the API, which can become unavailable during peak usage times. The relationship between the two methods for accessing the data contained in the medRxiv database is summarized in Figure 3.2.

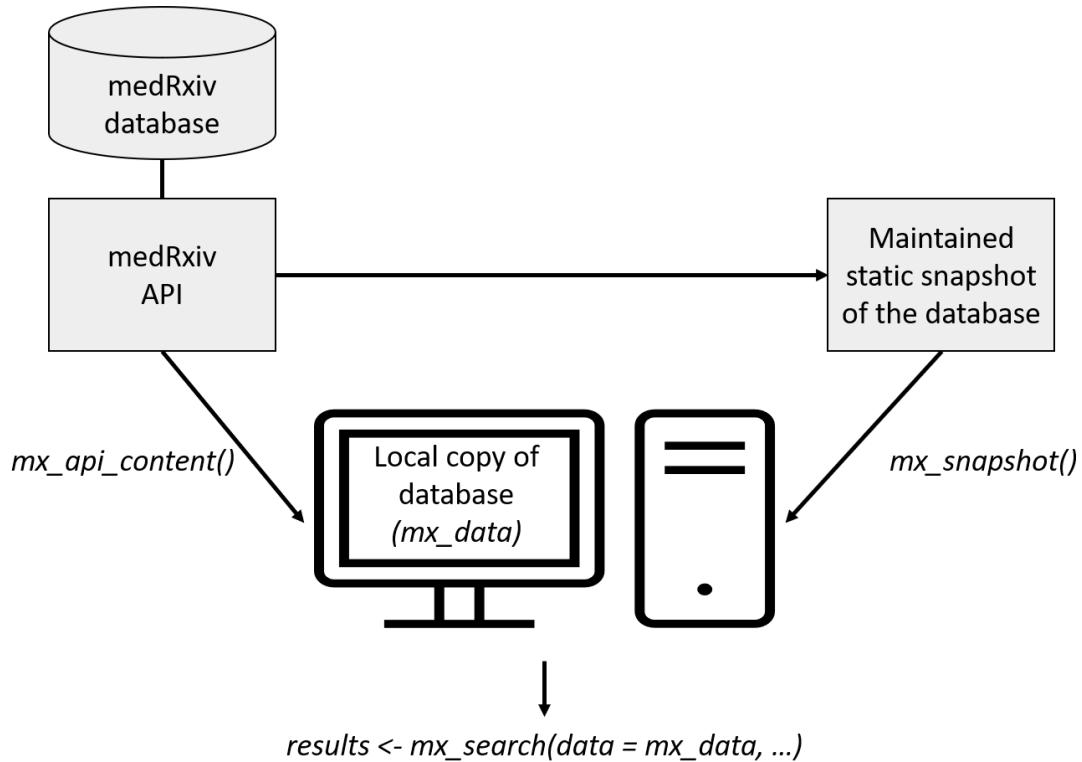


Figure 3.2: Overview of `medrxivr` data sources: Users can either access the API directly via `mx_api_content()`, or can import a maintained snapshot of the database, taken each morning at 6am, via the `mx_snapshot()` function. Note: due to the size of bioRxiv, only a maintained snapshot of the medRxiv repository is available via `mx_snapshot()`.

3.4.3 Performing a search

Once a local copy of the metadata is created, the first step in searching it is to create a search strategy. Search terms to be combined with the OR operator are contained in vectors (`c(...)`), while topics to be combined with the AND operator are contained in lists (`list(...)`).

```
# Create the search query

topic1 <- c("dementia", "alzheimer's") # Combined with OR
topic2 <- c("lipids", "statins")         # Combined with OR

myquery <- list(topic1, topic2)           # Combined with AND
```

For example, when written in standard syntax, the search contained in the `myquery` object above would be: “((dementia **OR** alzheimer’s) **AND** (lipids **OR** statins))”. There is no limit to the number of search terms that can be included in each topic, nor in the number of topics that can be search for. Search terms can also contain common syntax used by systematic reviewers and health librarians, including the use of NEAR statements which allows for identification of co-localised terms, and wild-cards, which allow for alternate spellings, e.g. “randomisation” vs “randomization”.

Once a strategy has been defined, it is passed along with the local copy of the database to the `mx_search()` function.

```
# Run the search

results <- mx_search(mx_data,
                      myquery)
```

3.4.4 Refining your search

An important argument of the `mx_search()` is `report`, which outputs a structured table with each search strategy presented on an individual line and the number of records associated with this strategy. [CITATION NEEDED - PRISMA-S]

```

results <- mx_search(mx_data,
                      myquery, # Combined with OR
                      report = TRUE)

## Found 1 record(s) matching your search.

##
## Total topic 1 records: 224
## dementia: 224
## alzheimer's: 0
##
## Total topic 2 records: 119
## lipids: 90
## statins: 33

```

This allows users to discover which terms in their search are contributing most to the total number of results returned. This is important as part of developing a search strategy,¹⁰⁴ as it allows for the key terms related to each topic to be discovered. It also aids in identifying misspelled or case-sensitive search terms, which will frequently return no results. As an example, in the search presented above, the term “alzheimer’s” returns no records. This is expected, as “Alzheimer” is a proper noun and so should be capitalised, but serves to illustrate the usefulness of the reporting function.

3.4.5 Exporting to a bibliography file

In line with the second success criteria (Section 3.3.1), one of the key features of the `medrxivr` is the ability for users to easily export the results of their systematic search to a reference manager. While it is a seemingly simple request, this is one of the key ways in which `medrxivr` is set apart for other preprint search tools, including the native medRxiv/bioRxiv website search functionality.

For example, the results of our simple search above can be exported to the "medrxiv_export.bib" file using the following code:

```
mx_export(results,
          file = "medrxiv_export.bib",
          report = TRUE)
```

3.4.6 Downloading the PDFs of relevant records

`medrxivr` also allows users to download the full text papers for records that are deemed eligible for full-text screening (see Figure 3.1). `mx_download()` takes the list of included records and saves the PDF for each to a folder specified by the user. This functionality is similar to the “Find Full Text” feature offered by EndNote.

```
mx_download(results, # Search results, less excluded records
            "pdf/") # Directory to save PDFs to
```

3.5 Discussion

3.5.1 Reception and future plans

The tool has been well received by the community (as of December 2021, `medrxivr` has been downloaded more than 4500 times), and several use cases have been reported. It has been used to investigate the role of preprints in the response to the 2019 coronavirus outbreak,¹⁰⁵ perform searches of preprints as part of a systematic review,^{106,107} and examine how data-sharing behaviour is affected by journal policies (see 2.8). `m McGuinness2020c?`

The package was been onboarded into the rOpenSci suite of packages, a collection of “carefully vetted, staff- and community-contributed R software tools that lower

barriers to working with scientific data sources on the web".¹⁰⁸ As part of this process, following rigorous peer-review, an associated article introducing the tool was published by the Journal of Open Source Software.^{mcguinness2020a?} The entire review discussion is publicly available and can be viewed online.¹⁰⁹ The tool has also been well received by the open-source community, demonstrated by the engagement of other developers in contributing to important new functionality and suggesting bug-fixes.

Lobbying of the Cold Springs Harbor Laboratory to develop the API to allow for direct searching of the database has been ongoing. This would negate the current need to download a local copy of the relevant preprint database before searching it, which is currently the rate limiting step for performing searches. For example, as of January 2021, downloading a copy of the bioRxiv database takes approximately an hour.

3.5.2 Use case

In addition to being used to search systematically search health-related preprint servers, as illustrated in the systematic review presented in Chapter 4, `medrxivr` has other uses. For example, I led a descriptive analysis of the change in data availability statements between preprinted and published versions of the same manuscript, stratified by journal data sharing policy access, underpinned by preprint meta-data provided by `medrixvr`.

medRxiv requires that authors complete a data availability statement, but do not police

We found some evidence that data availability statements more frequently described open data on publication when the journal mandated data sharing compared to when the journal did not mandate data sharing.

This study has since been published in PLOS One, and a copy is included in Appendix B.4.

3.5.3 Limitations of medrxivr

While searching of the medRxiv and bioRxiv databases was crucial for the systematic review element of this thesis presented in Chapter 4, there are some important limitations to note here. A key example is that the tool only searches the available metadata of preprint records (the title, abstract and keywords), rather than the full text of preprints, meaning some relevant records might be missed. However, this approach echoes that used by other search platforms such as OvidSP, and while some relevant records may be missed (reduced sensitivity), limiting the search to the metadata fields prevents non-relevant records from being returned (high specificity). A key example of the reduced specificity when searching the full text, identified during development of **medrxivr**, is that a search for “dementia” would return a record where the only occurrence of this term is in the title of one of the references.¹¹⁰

There is also the potential that the cross-section of literature posted on medrxiv/bioRxiv is substantially different from the true grey literature (studies or analyses that are not published for a range of reasons including results that are not deemed “novel” or are not statistically significant).¹¹¹ This is because simply lowering the barriers to publication may well encourage authors to publish “null” results, but due to the effort involved in writing up a distributable manuscript, it is unlikely to completely address the “file drawer” effect.⁷¹ It is likely too early (and likely too methodologically difficult) to tell whether the increased popularity and acceptance of preprint repositories will have any effect of the availability of research that was not considered “publishable” at other venues.

3.5.4 Role of open source tools in evidence synthesis

Part of the motivation for creating the `medrxivr` tool was a belief that the development and distribution of open source scripts and tools should be a fundamental part of evidence synthesis research.^{goldacre2019b?, mckiernan2016c?} In the case of `medrxivr`, it is likely that several other evidence synthesists had written personal scripts that have a similar, or related, functionality - in fact, following development of the tool, I identified one other researcher that has done so (Nicholas Fraser, author of the `rbiorxiv` package, which allows for importing medRxiv metadata into R but does not provide search functionality).^{rbiorxiv?} If these scripts continue to be developed in private and are never shared or publicised, this will inevitably harm to the evidence synthesis community, not only in terms of duplication of time and effort but also due to lost opportunities for collaboration.^{mckiernan2016c?} Creating and sharing well-documented packages, the recognised standard for sharing code in R, represents one way to reduce this inefficiency.¹¹²

3.6 Summary

- In this Chapter, I have introduced a new tool, `medrxivr`, for performing complex systematic searches of the medRxiv and bioRxiv preprint repositories.
- I have outlined the motivation for developing this tool in relation to this thesis - more specifically, that it was used to perform systematic and reproducible searches of a key literature sources used in the comprehensive systematic review described in Chapter 4.
- I have contrasted `medrxivr` with other available interfaces to medRxiv/bioRxiv data to highlight the added functionality it offers. I have also discussed the tools reception to date, its limitations, and the important role of open-source tools like `medrxivr` in evidence synthesis.

3.7 References

4

Systematic review of all available evidence

Contents

4.1	Lay summary	49
4.2	Introduction	50
4.3	Methods	51
4.4	Results	66
4.5	Discussion	77
4.6	Conclusions	88
4.7	References	89

4.1 Lay summary

Systematic reviews are a type of research that aim to use all existing evidence to provide the best answer to an important research question. They do this by finding and combining the results from many related primary research studies. Reviews involve multiple steps including: searching of existing studies; assessment of the studies against predefined inclusion criteria; collection of data from each study; assessment of each study's methods.

This chapter presents a systematic review of primary studies that have examined the relationship between the levels of blood lipids (such as cholesterol and triglycerides),

and treatments that change these levels, and dementia.

There were 127 primary studies that contained information on this relationship. I found that statins reduce the risk of Alzheimer's disease, but had no effect of vascular dementia. Lipids were not associated with any outcome. The methods used in some of the primary studies meant that I was less confident in the accuracy of their results.

The use of the results of this review in subsequent chapters is dicussed.

4.2 Introduction

In this chapter, I describe a comprehensive systematic review of the relationship between blood lipid levels, and treatments that modify them, and the subsequent risk of dementia and related outcomes. This analysis sought to address two specific aims.

Firstly, as discussed in the Introduction to this thesis (Section 2.5), several diverse forms of evidence on the relationship of lipids and dementia exist. These include randomised controlled trials, observational studies of different analytical design (cohort, case-control, etc), and Mendelian randomisation studies. However, based on a scoping review of existing literature, no previous evidence synthesis exercise has attempted to examine the association of lipids/statins with dementia outcomes across these distinct evidence types. Collating these diverse evidence sources is important, as if the observed association between lipids and dementia is constant across them, it increases our confidence in the association. As such, the primary aim of this analysis was to systematically review all available literature, regardless of study design.

Secondly, I explicitly sought to include health-related preprint servers as a potential evidence source in this review, as they are infrequently considered by evidence syntheists but may contain important unpublished studies. As an adjunct analysis to the systematic review presented in this chapter, I sought to quantify the additional evidential value of including preprints. This inclusion of preprint serves makes use of the preprint search tool presented in Chapter 3.

The results of this review are used to guide the primary analysis presented in Chapter 5, in addition to forming a key evidence source used in the triangulation exercise presented in Chapter 7.

4.3 Methods

4.3.1 Protocol

A pre-specified protocol for this analysis was registered using the Open Science Framework, and is available for inspection.¹¹⁴ Deviations from this protocol are detailed in the relevant sections.

4.3.2 Contributions

In line with best-practice guidance, secondary reviewers were used to check the accuracy of screening, data extraction and risk-of-bias assessment processes. Due to the scale of the project, this systematic review was performed in conjunction with a team of secondary reviewer. These included Alexandra MacAleenan, Athena Sheppard, Matthew Lee and Lena Schmidt. In addition, Sarah Dawson, and information specialist, provided input to the design of the search strategy.

4.3.3 Search strategy

I systematically search electronic bibliographic databases to identify potentially relevant entries (hereafter referred to as “records”). The search strategy used in each database was developed in an iterative manner using a combination of free text and controlled vocabulary (MeSH/EMTREE)⁶⁸ terms to identify studies which have examined the relationship between blood lipids levels and dementia, incorporating

4.3 - Methods

input from an information specialist. The strategy included terms related to lipids, lipid modifying treatments, and dementia and its sub-types, and was designed for MEDLINE before being adapted for use in the other bibliography databases listed. An outline of the general strategy is presented in the Table 4.1 below and the full search strategies for each database are presented in Appendix A.2.1.

Table 4.1: Summary of systematic search by topic. The full search strategy including all terms and the number of hits per term is included in Appendix A.2.1.

No.	Concept
1	Dementia/Mild Cognitive Impairment
2	Lipids
3	Lipid-modifying treatments
4	1 AND 2
5	1 AND 3
6	4 OR 5
7	Animals NOT (Animals AND Humans)
8	6 NOT 7
9	Observational filter
10	Randomised controlled trial (RCT) filter
11	Mendelian randomisation/Instrumental variable filter
12	OR/ 9-11
13	8 AND 12

For all topics, search queries were comprised of relevant free text & controlled vocabulary terms.

When searching the bibliographic databases, study design filters were employed to try and reduce the screening load. To ensure that the study design filters are not excluding potentially relevant records, a random sample of 500 records identified by the main search but excluded by the filters (defined as “8 NOT 12” in Table 4.1) was screened.

The following databases were searched from inception onwards: Medline, EMBASE,

Psychinfo, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science Core Collection. As the contents of the Web of Science Core Collection can vary by institution, the specific database searched via this platform are listed in Appendix A.2.2. I also search clinical trial registries, for example ClinicalTrials.gov, to identify relevant randomized controlled trials.

In addition, I searched the bioRxiv and medRxiv preprint repositories using the tool developed in Chapter 3 to identify potentially relevant preprinted studies.

Grey literature was also searched via ProQuest, OpenGrey and Web of Science Conference Proceedings Citation Index, while theses were accessed using the Open Access Theses and Dissertations portal. In addition, the abstracts list of relevant conferences (e.g. the proceedings of the Alzheimer’s Association International Conference, published in the journal *Alzheimer’s & Dementia*) were searched by hand. Finally, the reference lists of included studies were searched by hand while studies citing included studies was examined using Google Scholar (forward and reverse citation searching or “snowballing”).

4.3.4 Study selection

Records were imported into Endnote and de-duplicated using the method outlined in Bramer et al. (2016).¹¹⁵ In summary, this method uses multiple stages to identify potential duplicates, beginning with automatic deletion of records matching on multiple fields (“Author” + “Year” + “Title” + “Journal”), followed by manual review of less similar articles (e.g. those matched based on the “Title” field alone).

Following deduplication of records, screening (both title/abstract and full-text) was performed using a combination of Endnote, a citation management tool,¹¹⁶ and Rayyan, a web-based screening application.¹¹⁷ Title and abstract screening to remove obviously irrelevant records was performed primarily by me, with a random sample of excluded records being screened in duplicate to ensure consistency with the inclusion criteria.

4.3 - Methods

Similarly, I completed all full-text screening, with a random ~10% being screened in duplicate by a second reviewer. In addition, any records identified I identified as being difficult to assess against the inclusion criteria were screened in duplicate. Reasons for exclusion at this stage were recorded. Disagreements occurring during either stage of the screening process were resolved through discussion with a senior colleague. A PRIMSA flow diagram was produced to document how records moved through the review.⁷⁸

The criteria used to assess eligibility are presented in the subsequent sections.

Inclusion criteria

I sought to include studies that examine the relationship between blood lipid levels (or any specific lipid fraction, including total cholesterol, HDL, LDL, and triglycerides) and risk of incident dementia/MCI. Eligible study designs include randomized controlled trials and non-randomized observational studies of lipid modifying treatments, longitudinal studies examining the effect of increased/decreased blood lipid levels, and genetic instrumental variable (Mendelian randomization) studies examining the effect of genetically increased/decreased blood lipid levels.

Participants were free (or assumed to be free) of dementia/MCI at baseline. Studies of any duration were included to allow for exploration of the effect of length of follow-up on the effect estimate using meta-regression. No limits were placed on the sample size of included studies.

Eligible studies defined dementia according to recognised criteria, for example the National Institute of Neurological Disorders and Stroke Association-Internationale pour la Recherche en l'Enseignement en Neurosciences (NINDS-AIREN), International Classification of Diseases (ICD), or Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria. For MCI, eligible studies are those that attempted state a definition for diagnoses of MCI (e.g. an adapted version of the

Petersen criteria)¹¹⁸ and create ordinal groups of patients (e.g. no dementia or dementia/MCI/dementia) based on this definition.

No limitations were imposed on publication status, publication date, venue or language, although sufficiently detailed reports of the studies to be able to examine their methods were required for inclusion.

Exclusion criteria

Case-control studies, cross-sectional studies, qualitative studies, case reports/series and narrative reviews were excluded. Studies which present no evidence of attempting to exclude prevalent cases from their analyses were also excluded. Studies that measure change in continuous cognitive measures (e.g. MoCA score) without attempt to map these scores to ordinal groups (e.g. no dementia/MCI/dementia) were excluded. Conference abstracts with no corresponding full-text publication were examined, and I contacted authors to obtain information on the study's status. Studies that are reported in insufficient detail (e.g. only in conference abstracts, new, letters, editorials and opinion) were excluded. Previous systematic reviews were not eligible for inclusion, but their reference lists were screened to identify any potentially relevant articles.

Studies with outcomes not directly related to the clinical syndrome of dementia (e.g., neuroimaging), studies implementing a “multi-domain intervention” where the lipid-regulating agent is included in each arms (e.g. for example, a study examining exercise + statins vs statins alone, but a study examining exercise + statins vs exercise alone would be included), and studies where there was no screening for dementia at baseline except if the sample was initially assessed in mid-life (i.e. below the age of 50) were excluded. Finally, studies using a dietary intervention, for example omega-3 fatty acid enriched diet, were excluded as it is difficult to disentangle the effect of other elements contained within the diet.

4.3 - Methods

Note, this is distinct from studies which delivered a simple tablet-based omega-3 intervention, which would have been eligible for inclusion.

4.3.5 Validation of screening process

Inter- and intra-rater reliability during the screening stage were assessed for a 10% sub-sample of records at the title and abstract screening stage. Intra-rater reliability involved a single reviewer applying the inclusion criteria to the same set of records while blinded to their previous decisions (i.e. assessment of consistency), while inter-rater reliability involved two reviewers independently screening the same set of records (i.e. assessment of accuracy).

Rater reliability was assessed using Gwet's agreement coefficient (AC1).¹¹⁹ This measure of inter-rater reliability was chosen over other methods of assessing inter-rater reliability such as percent agreement (number of agreements divided by total number of assessments), as it accounts for chance agreement between reviewers but does not suffer from bias due to severely imbalanced marginal totals in the same way that Cohen's *kappa* value does.¹¹⁹ Given the small number of included studies in this review as a proportion of the total number screened, this is an important characteristic.

Gwet's AC1 is defined as:

$$AC1 = \frac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}}$$

In reference to a two-by-two table with cells A, B, C and D, it is calculated using the following:

$$AC1 = \frac{\frac{A+D}{N} - e(\gamma)}{1 - e(\gamma)} \quad (4.1)$$

where $e(\gamma)$ is the chance agreement between raters, given as $2q(1 - q)$, where

$$q = \frac{(A + C) + (A + B)}{2N} \quad (4.2)$$

How to interpret agreement co-efficients is widely debated, and while arbitrary cut-off values may mislead,¹²² they provide a useful rubric by which to assess inter-rater agreement. Here, I used guidelines based on a stricter interpretation of the Cohen's *kappa* coefficient,¹²³ presented in Table 4.2.

Table 4.2: Suggested ranges to aid in interpretation of Gwet's AC1 inter-rater reliability metric

Kappa	Interpretation
0 – 0.20	None
0.21 – 0.39	Minimal
0.40 – 0.59	Weak
0.60 – 0.79	Moderate
0.80 – 0.90	Strong
> 0.90	Almost perfect

If this assessment demonstrated issues with the screening process, a larger proportion of records would have been dual-screened.

4.3.6 Data extraction

Data extraction was performed using a piloted data extraction form. Extracted items included: article metadata (year of publication, author list, journal), study characteristics (location, cohort name, exposure, outcomes, outcome criteria used), patient characteristics (age, sex, baseline cognition scores, baseline education scores), and results (exposure-outcome pairing, effect measure, effect estimate, error estimate,

4.3 - Methods

p-value). All data was extracted the data in the first instance and was checked for accuracy by a second member of the review team.

Harmonisation of cholesterol measures

Harmonization of cholesterol measures across studies was performed, as different studies used different methods to quantify exposure, including comparing differing risks in the highest vs lowest quartiles of a lipid, using a binary classification of patients into a hypercholesterolaemia or not, categorizing lipid levels into high, middle, and low groups according to study-defined criteria, and simply treating the exposure as a continuous variable. Where necessary, lipid levels reported in *mmol/L* were converted in *mg/dL* using the following formula:

$$mg/dL = mmol/L \times Z \quad (4.3)$$

where $Z = 38.67$ for total cholesterol, LDL-c and HDL-c, and $Z = 88.57$ for triglycerides. For widely-used categorises of lipids levels on the *mg/dL* scale, see Table 2.2 in Section 2.4.1.

Study-fication

As part of the data extraction process, multiple records resulting from the analysis of the same data were included and grouped into single units, hereafter called studies. This is likely in the advent of multiple papers reporting results on the same cohort, but say, at different time points. Study-fication builds out the most comprehensive accounts of the studies and results from as many published articles were applicable.

This was particularly relevant to preprints and published papers reporting the same study, which were not considered duplicate records. Instead, they were considered as different reports of the same study. This is due to the potential for the published version to offer some information that the preprint did not, and vice versa.

Following up with authors

Where data points required either for the analysis or risk-of-bias assessment but were not reported, the primary authors of the study were contacted. This approach was taken due to the potentially large impact of following up with authors on the results of the review,¹²⁴

This was particularly important for the dose response meta-analysis, where the number of participants and the cut-offs per category were often not reported.

Converting between different exposure measures

The range of effect measures presented by studies (odds ratios, risk ratios, hazard ratios, etc) are not directly interchangeable in the context of systematic review. If the outcome is rare, odds and risk ratios approximate each other. However, hazard ratios provide a very different interpretation, taking into account time-to-event in each treatment group.

Several existing reviews do not distinguish between the types of effect measures and include all existing studies in a single meta-analysis to produce an overall effect. In addition, there is some evidence of manipulation of effect estimates in previous reviews,(e.g. Chou, Sci Reports - at least one study disagrees with) but this is not accurately documented in the review text.

4.3.7 Risk of bias assessment

A key use of the review presented in this chapter is to identify different sources of evidence at risk of a diverse range of biases, and to contrast and compare findings across them (see Section 7.1.1 for an overview of triangulation and Section 2.6.2 for the results of this qualitative analysis). To enable this triangulation

4.3 - Methods

exercise, a detailed and structured risk of bias assessment formed an important part of this review.

There has been a recent movement within the evidence synthesis community from examining *methodological quality* to assessing *risk of bias*,^{79,125} and thus directly evaluating the internal validity of a study. Internal validity is defined here as the absence of systematic error (or bias) in a study, which may influence its results.^{126,127}

This move was prompted by a unclear definition of “methodological quality” which could include facets such as, and challenges in the comparison of results from different tools. As part of this, the community also moved from checklist or score based tools towards domain-based methods, in which different potential sources of bias in a study are assessed in order.

Additionally, results should be assessed at the result (defined as a specific outcome at a specific timepoint) rather than the study level. For example, a study may report on the efficacy of an intervention at 6 months and two years follow-up. missing outcome data that is not an issues at 6 months may introduce bias at 2 year follow-up. In this case, assigning a bias judgement to the study as a whole hides this differential biases for each result.

In this review, domain-based tools were used to assess the risk of bias for each result in each included study. The study design-specific tools are introduced and discussed in more detail in the following sections. In addition, the tool also aim to capture the potential direction of bias for each result. Possible responses included: “Favours experimental”, “Favours comparator”, “Towards null”, “Away from null”, and “Unpredictable”.

Randomised controlled trials

Risk of bias assessment in randomised controlled trials was performed using the domain-based risk-of-bias assessment tool appropriate to the study design.

Randomized controlled trials were assessed using the RoB2 tool.¹²⁸

The tool assess the risk of bias across five domains: Bias arising from the randomization process, Bias due to deviations from intended intervention, Bias due to missing outcome data, Bias in measurement of the outcome, Bias in selection of the reported result. Acceptable judgements include: low risk of bias, some concerns, high risk of bias. Each of the 5 domains contains a series of signalling questions or prompts, which guide the user through the tool. Once a domain-level judgement for each domain has been assigned, an overall judgement, using the same three levels of risk of bias, is assigned to the result.

Non-randomised studies of interventions/exposures

For non-randomised studies of interventions (NRSI), I used the ROBINS-I (Risk Of Bias In Non-randomised Studies - of Interventions) tool.¹²⁵ This tool assess the risk of bias across seven domains: Bias due to confounding, Bias due to selection of participants, Bias in classification of interventions, Bias due to deviations from intended interventions, Bias due to missing data, Bias in measurement of outcomes, and Bias in selection of the reported result. Similar to RoB 2, it has a number of prompting questions per domain, with acceptable judgements including “Low risk”, “Moderate risk”, “Serious risk” and “Critical risk”. Ideally, observational studies should be assessed in reference to an idealised randomised controlled trial. Further, the rare overall judgement of “Low” is considered equivalent to a randomised controlled trial.

For non-randomised studies of exposures (NRSE), I opted to use a preliminary version of the ROBINS-E tool.¹²⁹ The tool is still in development, but while the signalling questions for each domain are yet to be confirmed, the overarching risk-of-bias domains have been finalised. The motivation for this using this tool above other existing published tools such as the Newcastle-Ottawa scale (NOS).¹³⁰ was two-fold. In the first instance, as mentioned in the introduction to this section, using a

4.3 - Methods

domain-based tool has distinct advantages over better-developed checklist-type tools such as the NOS. In addition, using a domain-based tool for non-randomised studies of exposures enabled better comparison with risk-of-bias assessments performed for the other study designs.

Mendelian randomisation studies

At present, no formalised risk-of-bias assessment tool for Mendelian randomization studies is available. Assessment of the risk of bias in Mendelian randomisation studies was informed by the approach used in a previous systematic review of Mendelian randomisation,¹³¹ as identified by a review of risk of bias assessments in systematic reviews of MR studies (advance results from this review were obtained from contact with the authors.). A copy of this tool is available in Appendix A.2.4.

Risk of bias due to missing evidence

A recent shift towards the assessment of missing evidence due to selective non-reporting - as distinct from the selective reporting of a single result from multiple planned - is demonstrated via the forthcoming RoB-ME (Risk of Bias due to Missing Evidence in a synthesis) tool. The tool is in development stages, and as part of this review, I piloted the tool, and provided feedback to the developers.

This additional appraisal marks a departure from the registered protocol, as there was initially no intention to try and examine the risk of bias due to missing evidence. This is because the tool did not exist when the protocol was originally registered.

4.3.8 Analysis methods

An initial qualitative synthesis of evidence was performed, summarising the data extracted from studies across

Where individual studies were deemed comparable, they were incorporated into a quantitative analysis or “meta-analysis”. Meta-analysis provides a summary or pooled effect estimate across studies.

Of note, studies were not combined across different study designs (i.e. RCTs were not combined in a meta-analysis with results from observational studies). The results from each individual analytical approach were summarised, but are compared and contrasted more fully in the triangulation exercise presented in Chapter 7

Standard meta-analysis

Both a fixed-effect and random-effects meta-analysis model was employed to combine the different included studies. The fixed-effect method was implemented as:

$$\theta_i = \mu + u_i \quad (4.4)$$

$$\text{weighted average} = \frac{\sum Y_i(1/SE_i^2)}{\sum(1/SE_i^2)} \quad (4.5)$$

Dose-response analyses

Several of the included studies presented data on multiple categories of lipid levels, but provided an overall effect estimate based on a comparison of only two of these categories (e.g. for example, highest vs lowest quartile). While this allows for easy interpretation of the resulting effect estimate, it ignores any potential non-linear relationships between the exposure and outcome, in addition to discarding useful information contained in the interim groups.

4.3 - Methods

Where possible, data for all exposure groups was extracted. Studies were excluded from this analysis if the number of categories was less than three, if the exposure cut-off points for each category were not presented (e.g. if the study reports splitting participants into quartiles and comparing the highest vs lowest without giving the quartile bands in mmol/L).

A spline model was fitted to allow for a non-linear relationship, for example a U or J-shaped relationship, where low and high levels of an exposure can have an effect versus the “normal” reference dose. Reference doses were defined *a priori* as the cut-off of the “Normal”/“Optimal” categories for each fractions, as detailed in Table 2.2. Under this approach, the reference dose was defined as 200 mg/dL for total cholesterol, 100 mg/dL for LDL-c, 40 mg/dL for HDL-c, and 150 mg/dL for triglycerides.

Due to the requirements for the dose response analysis, studies were excluded from this secondary analysis if they did not provide the require information: cut-off points. Where this was not reported in the study, I contacted the corresponding author to attempt to obtain the required information (see 4.3.6).

When the highest category was open ended (e.g LDL-c ≥ 200 mg/dL), I calculated category midpoint by assuming the width of the higest category was the same as the one immediately below it. Similarly, when, the lowest category was open-ended (e.g LDL-c ≤ 100 mg/dL), I set the lower boundary for this category to zero (though this is unlikely to occur natually, it was difficult to define).

Sensitivity analyses

I conducted a leave-one-out analysis in order to explore the impact of any given study on the results. In addition, where available, I performed meta-regression to assess whether the results varied by age or sex.

Visualisation of results

The use of evidence maps is becoming increasingly common to explore the distribution of research cohorts included in a systematic review. As such, each individual cohort contributing to the evidence base was quantified and visualised on a world map.

Further, given the importance of visualising the potential biases of a result alongside the result itself, a new visualisation tool was designed to allow for “paired” forest plots (as recommended by the ROB2 publication). This tool was developed as an adjunct to this thesis to aid in creating standard risk-of-bias figures,¹³² and the “paired” forest plot functionality grew out of a collaboration with other researchers to design a modular method for creating these plots.¹³³ A summary of this tool is contained in Appendix B.2, and all forest plots presented in this analysis were created using this tool.

Assessment of added value of including preprints

Preprints are a valuable evidence source (see Introduction, Section @ref()) but their inclusion in a systematic review. As part of a study within this larger review, I also explored the additional evidential value of including preprints in the meta-analysis, assessed using the fixed effect weight from a standard meta-analysis.

Additionally, I followed preprints up over time to investigate whether all identified preprints included in the review were subsequently published (in which case preprints provide a snapshot into the future, and a systematic review update would capture these reports) or alternatively, if some preprints were not published, then preprints provide a distinct evidence source.

4.4 Results

4.4.1 Initial search and validation of search filters

Twenty-three thousand, four hundred forty-seven records were identified through database searches.

Of the 500 random records screened to ensure the accuracy of the study design filters, no eligible records were identified. Many of those excluded by the filters were basic science studies, commentaries or educational articles.

4.4.2 Screening results

Following de-duplication, the titles and abstracts of 16,109 records were assessed for eligibility. 387 were deemed potentially eligible, and the full text records for these were accessed and screened. 4.1

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) flow diagram⁷⁸ presented in Figure 4.1, illustrates the movement of articles through the review.

To highlight the contribution of preprint archives to the review, the flow diagram delineates between those records captured through databases searches (presented on the right of the diagram) and those captured by the search tool described in the previous chapter (presented in grey on the left of the diagram).

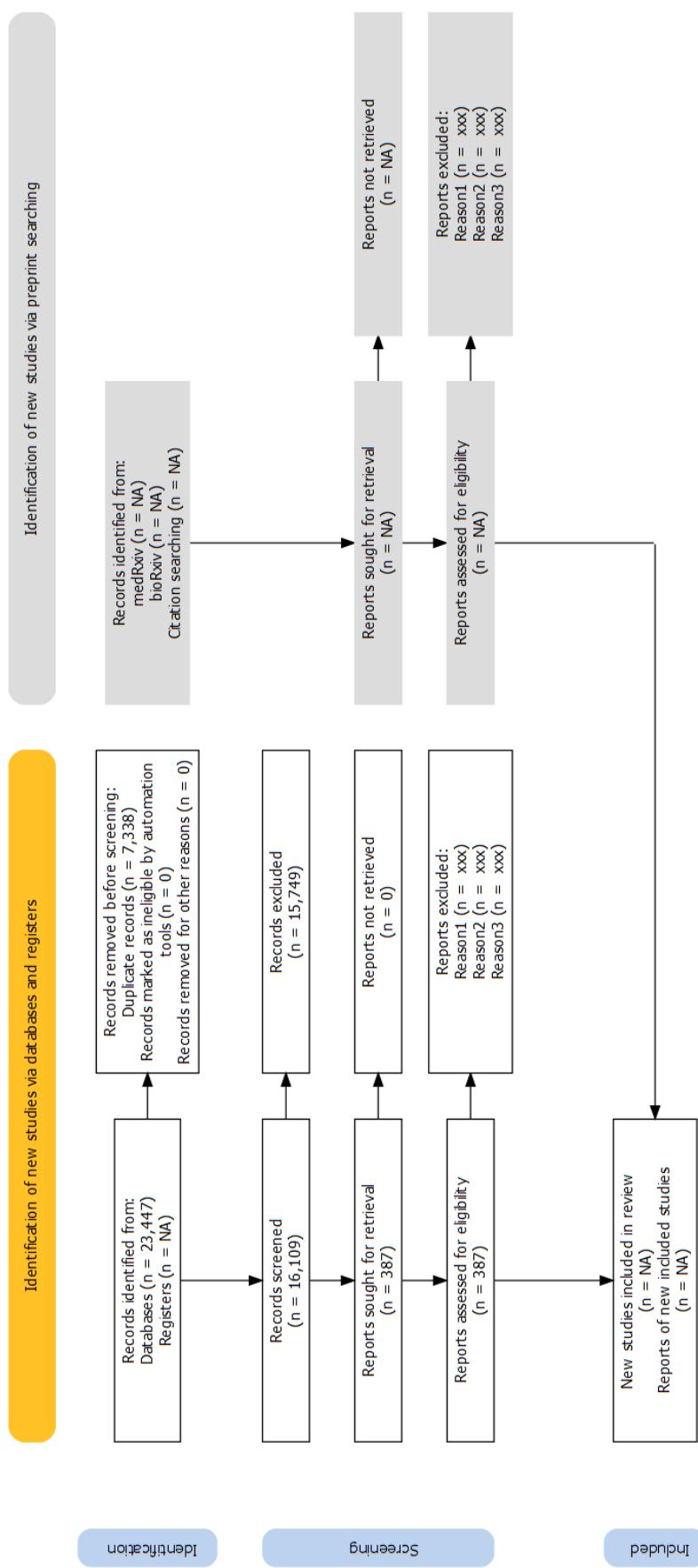


Figure 4.1: PRISMA Flow diagram illustrating how records moved through the systematic review process. The different contributions of databases and preprint servers to the review are indicated.

4.4 - Results

The values of $AC1$ were interpreted against the categories presented in Table 4.2. For the inter-rater reliability, agreement was “almost perfect” ($AC1 = 0.97$, $kappa = 0.54$, Table 4.3). Similarly for intra-rater reliability, agreement was “almost perfect” ($AC1 = 0.99$, $kappa = 0.65$, Table 4.4). The discrepancy between the $AC1$ and $kappa$ coefficients illustrates the sensitivity of $kappa$ to imbalanced marginals, caused in this sample by a heavy imbalance towards exclusion.¹³⁴

Table 4.3: Inter-rater agreement on a subset of records, indicating high accuracy.

		Initial screening decision		
		Exclude	Include	Total
Second reviewer decision	Exclude	1244	9	1253
	Include	26	22	48
	Total	1270	31	1301

Table 4.4: Intra-rater agreement on subset of records, indicating high consistency.

		Initial screening decision		
		Exclude	Include	Total
Same reviewer decision	Exclude	1266	14	1280
	Include	4	17	21
	Total	1270	31	1301

Those records which were excluded in the initial screening, but were included by the second reviewer (n=26, Table @ref(tab: agreeInter-table)) were investigated. This discrepancy between the two reviewers was explained in all cases by differing interpretations of the inclusion criteria, specifically around the definition of cognitive decline versus mild cognitive impairment, and the definition of eligible lipids fractions.

4.4.3 Characteristics of included studies

Following full-text screening, XXXX studies met the criteria for inclusion in the review. Table 4.5 presents a summary of the characteristics of each study.

Table 4.5: studyCharacteristics

Study	Location	Data source	Exposures	Outcomes	Diagnostic criteria
Ancelin 2012	France	Three City Study	Fibrate; Statin	Dementia; AD	DSM-IV; NINCDS-ADRDA
Ancelin 2013	France	Three City Study	TG; LDL-c; HDL-c	Dementia; AD	DSM-IV; NINCDS-ADRDA; NA
Arvanitakis 2008	US	Religious Orders Study	Statin; Statin - Lipophilic; Statin - Hydrophilic	AD	Consortium to Establish a Registry for Alzheimer's Disease (CERAD) NA
Dufouil 2005	France	Three City Study	NA	NA	NA
Kuo 2015	Taiwan	Taiwan National Health Insurance	NA	NA	NA
Li 2010	United States of America	HMO (Group Health Cooperative) Three City Study	Statin TC	AD Dementia; AD	NINCDS-ADRDA DSM-III-R; NINCDS-ADRDA NA
Mielke 2010	France	NA	NA	NA	NA
Rasmussen 2018	Denmark	NA	NA	NA	ICD
Redelmeier 2019	Canada	Ontario Health Insurance Plan	NA	NA	NA
Svensson 2019	Japan	Japan Public Health Centre	NA	NA	NA
Tynkkynen 2016	Finland	FINRISK	NA	NA	NA
Tynkkynen 2018	Multiple	NA	NA	NA	NA
Zamirini 2004	United States of America	Veterans Affairs Medical Center	NA	NA	NA

4.4 - Results

4.4 - Results

4.4 - Results

The majority of studies were non-randomised, with only one included randomised controlled trials (the Heart Protection Study) and a relative small number of Mendelian randomisation studies. Of the non-randomised studies, many were assessments of the relationship with statis use.

Of the studies examining lipid levels, by far the most common fraction assessed was total cholesterol. The association with triglycerides levels was not consider in many papers.

The vast majority of studies examined either all-cause dementia or Alzheimer's disease, with only a small proportion examining vascular dementia.

Many studies using electronic health records as their data source did not accurately report the diagnostic criteria used to identify cases, and of those that did, none sought to validate the accuracy of their defined code lists.

Many of the studies were conducted in electronic health record databases, which have a number of advantages over smaller more traditional cohort studies.

As shown in Figure 4.2, the majority of study cohorts were based in the developed world.

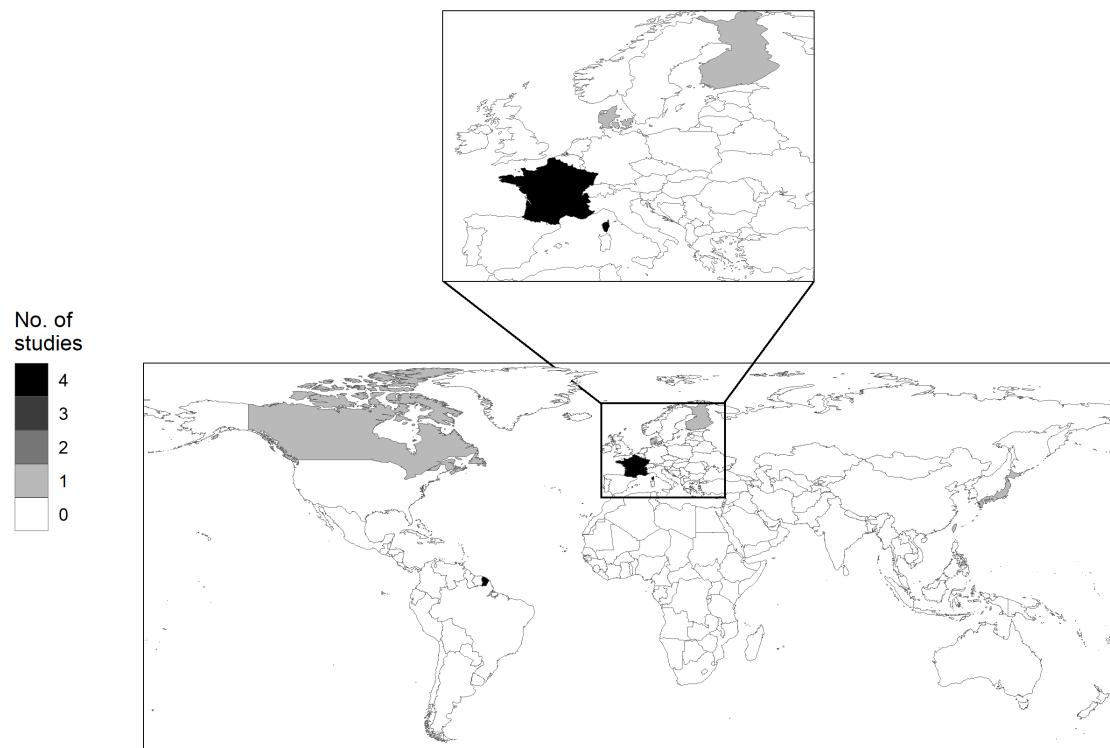


Figure 4.2: Geographical distribution of study cohorts

4.4.4 Risk of bias

Many of the included studies were at high-risk of bias.

The one randomised controlled trial was deemed to be a low risk of bias,

4.4.5 Meta-analysis of statin use and hypercholesterolemia

Statin use was associated with a reduced risk of all-cause dementia (OR: 0.76, 95%CI: 0.66-0.88, Figure 4.3).

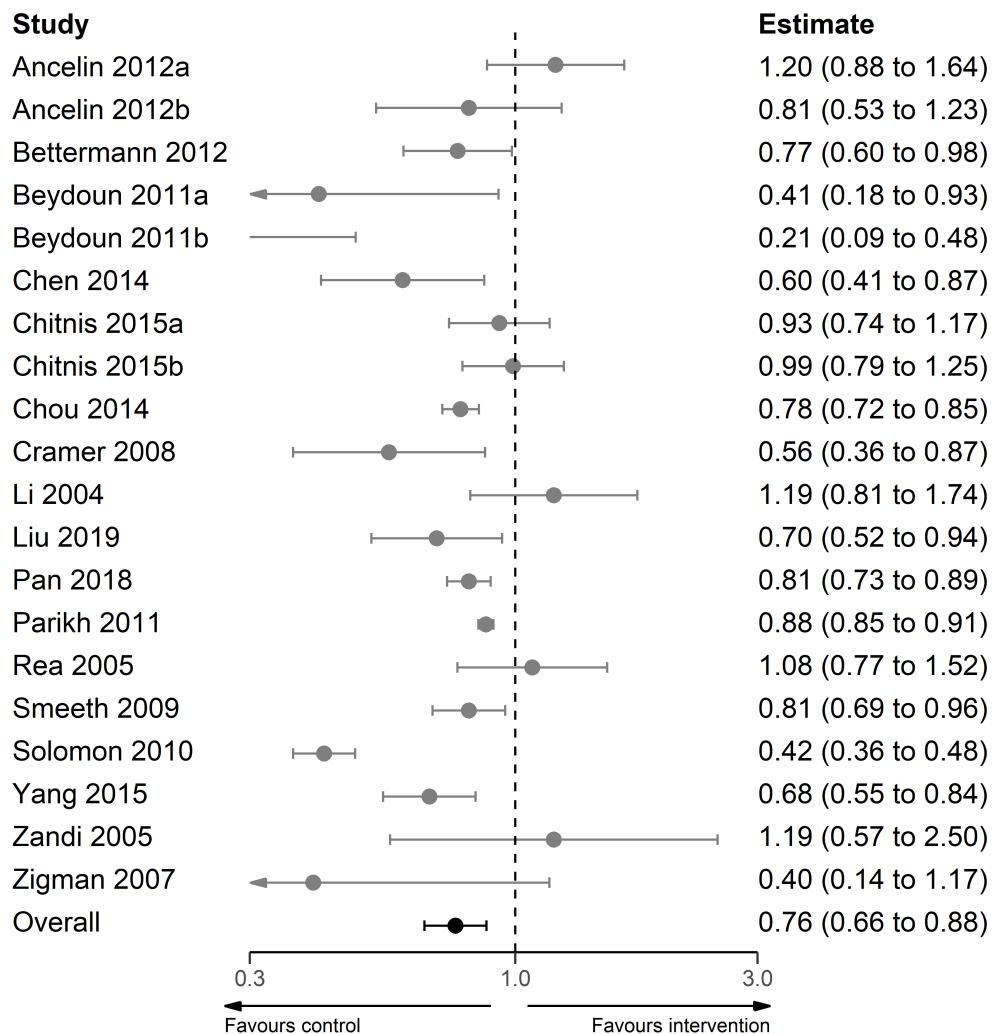


Figure 4.3: Forest plot showing effect of statins:

4.4.6 Dose response meta-analysis of lipid levels

Several studies were excluded from the dose-response meta-analysis, as the number of cases/controls per dose group could not be calculated and the corresponding author for the study did not respond to clarification requests. The results from the dose response analysis can be seen in Figure ??.

4.4.7 Publication bias

There was little evidence of publication bias across the evidence base (Figure 4.4).

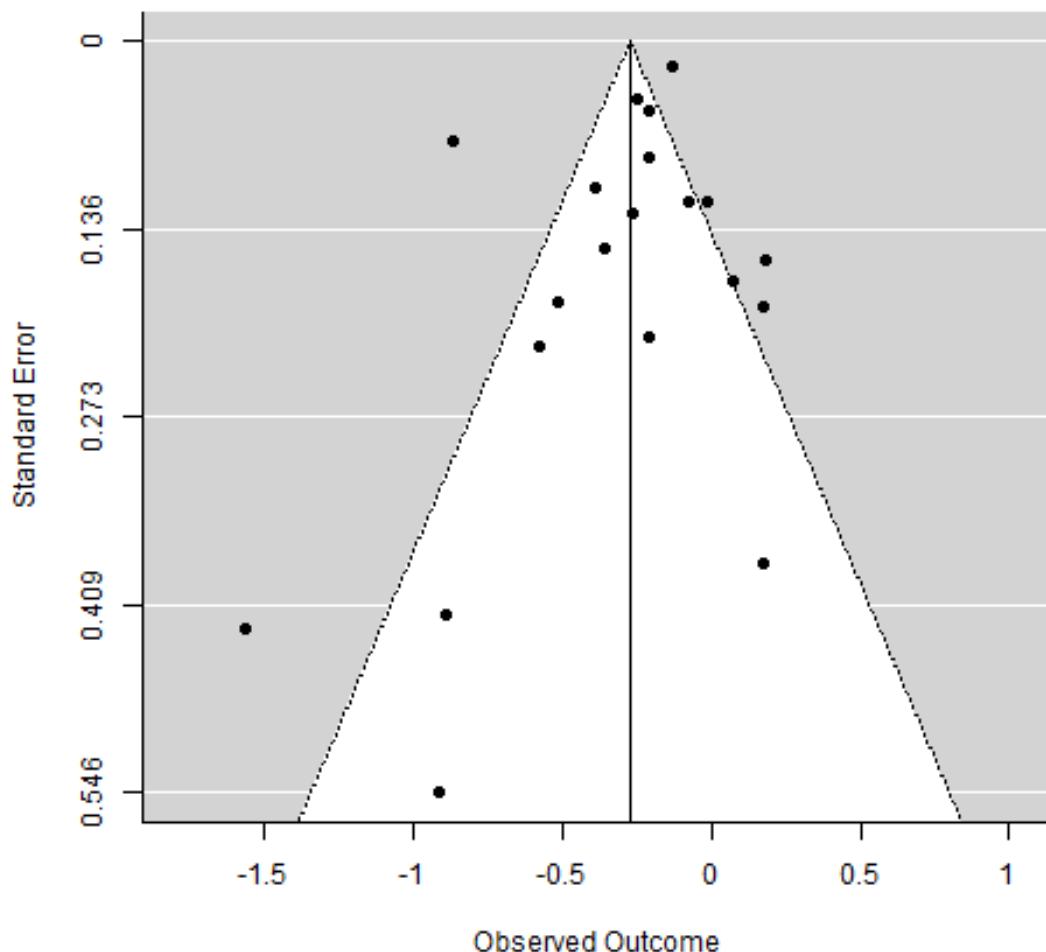


Figure 4.4: Funnel plot of results examining the relationship between statins and any dementia

4.4.8 Added evidental value of including preprints

As show in Figure ??, the number of results returned by the preprint searching, was not substantial (bioRxiv = 256, medRxiv = 0).

4.4.9 Risk of bias due to missing evidence

A key issue identified relating to bias due to missing data was the potential for the search to miss relevant Mendelian randomisation studies which had examined the association between Alzheimer's disease and multiple risk factors. This is discussed further in Section 4.5.4.

4.5 Discussion

4.5.1 Summary of findings

This review has presented a summary of the available evidence on the association between lipids, and treatments that modify lipids such as statins, and the subsequent risk of dementia.

The distribution of evidence between analytical designs is to be expected. Randomised controlled trials of dementia are particularly challenging, as the long follow-up, necessary due to the long latent period of the condition, makes trials logistically challenging and financial expensive. Similarly, Mendelian randomisation is a relatively new study design, and so only appears in the literature in recent years, driven by the growth of dementia GWASs (Figure 4.5).

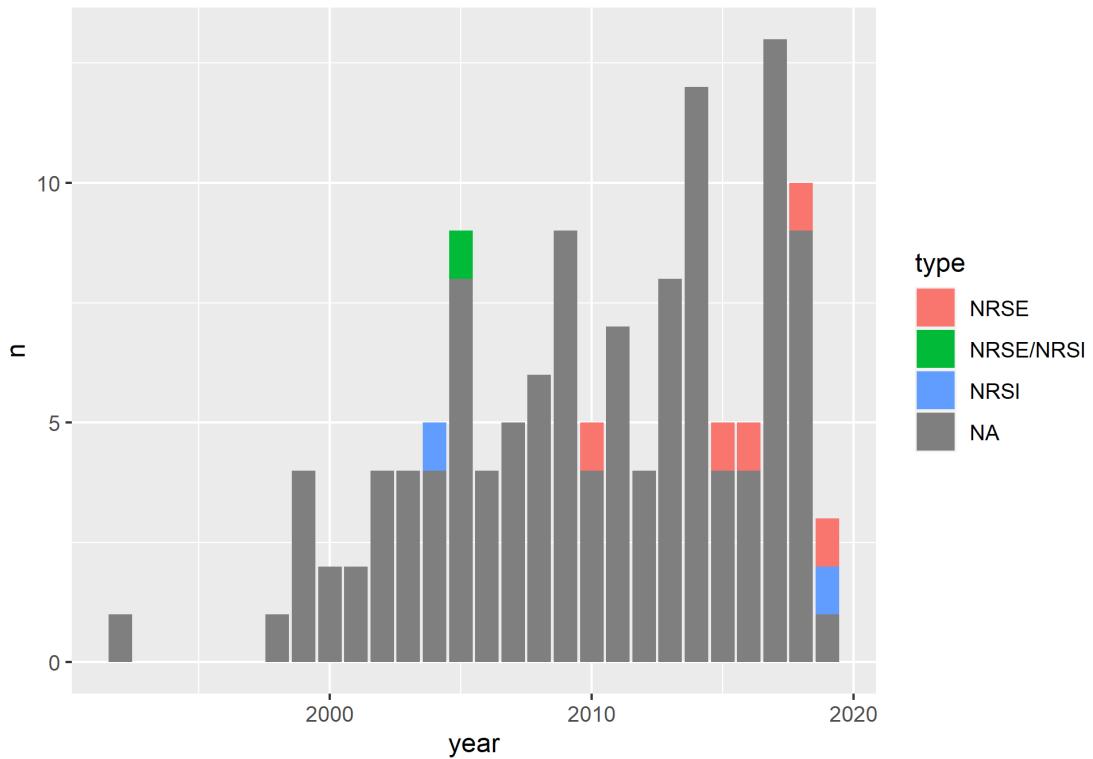


Figure 4.5: Study designs by year of publication -

A common theme across the evidence base was a lack of data on the association of vascular dementia. This is particularly interesting given that lipids and statins are primarily related to vascular disease. There is the potential that studies encountered similar difficulties in address the unexpected results observed in the CPRD analysis in Chapter 5, likely due to confounding by indication, and so may suffer from the “file-drawer effect”. For vascular dementia, few Mendelian randomisation studies examined this outcome, primarily because of the absence (until recently) of genome wide association studies, which are used in two sample summary Mendelian randomisation studies.

One item of particular interest is the attenuation of any effects observed by Mendelian randomisation studies following the adjustment for/exclusion of genetic variation in the Apoe4 gene region. As covered in the discussion, ApoE4 genotype is the major risk factor for Alzheimer’s disease, but is also involved in cholesterol transport.

4.5 - Discussion

This review did not include the commonly cited PROSPER study, which examined the effect of pravastatin on CVD risk and is regularly held up as providing no evidence for an effect of the statin on dementia outcomes. While widely cited and included in the Cochrane review on this topic,¹³⁵ only reported the mean change in a range of cognitive measures (MMSE, Stroop test, Picture-Word Learning test, etc.) over follow-up rather than incident. While this is a useful indicator of general cognitive decline, it is not equivalent to a dementia diagnosis using a , as cognitive tests should feed into a broader diagnostic pathway (see section 2.3.2). As such, this trial did not met the inclusion criteria for this review.

4.5.2 Previous reviews

This section will be completed once Georgia has completed her analysis, and can also be cross-references with the meta-meta analysis published in Brain Sciences recently.

While conducting this review, I identified several previous systematic reviews of this topic.^{136–140} However, this review is the first to use established domain based assessments tools (for example, the RoB 2 tool for randomized controlled trials)¹²⁸ to assess the risk of bias in included studies, and explore the heterogeneity of results across different levels of risk of bias levels. Some previous reviews did assess risk of bias, but used non-domain based assessment tools, such as the Newcastle-Ottowa scale.

However, despite these differences in methodology, the duplication of work across reviews (including this review) is substantial. In retrospect, an alternative approach to conducting a further systematic review from scratch could have been employed. Known as an umbrella review, or review-of-reviews, these studies use other systematic reviews rather than primary studies as the unit of analysis. This approach would have enable more efficient identification of relevant primary studies, to

which the methods which sets this review apart from other published reviews could have been applied.

4.5.3 Inclusion of preprints

As highlighted in Section 2.6.1, this review explicitly sought to synthesize evidence across different publication statuses (preprinted vs. published). Using the tool described in Chapter 3.2, two preprint serves related to health and biomedical sciences were search as part of this review (see Appendix A.2.3 for the code used to search the repositories). There were several relevant preprints captured by the search. The added evidential value of including these preprints was highlighted in Section ??sys-rev-including-preprints).

The small number of studies return by the searches (or the absence of any hits in the medRxiv database - see Figure ??) is due to the timing of the preprint searches. The searches for this review were performed in mid-July 2019, but the first medRxiv preprint was registered on 25th June 2019. As such, at the point it was searched, the medRxiv database contained only a small number of records.

While none of the preprints contributed uniquely to the review, this is again. Consider the example of one of the Mendelian randomisation analyses of the effect of LDL-c on Alzheimer's disease, which found no impact of LDL-c on AD following removal of APOE4. The manuscript was initially published in bioRxiv in July 2017¹⁴¹ which was subsequently published in Nature Communications in January 2018, following peer-review.¹⁴² While this study was captured by both the preprint and published searches in this review, had the searches been run within this window, the preprint would have contributed unique data to the review. This illustrates that while it may not have aided this review, if the aim is to find the current state of the art in the topic area at the time of searching, inclusion of preprints is a necessity.

4.5 - Discussion

Of note, since the start of my review, inclusion of preprints in systematic reviews has

now become widespread due to the role of preprint servers, in particular medRxiv,

as a key evidence dissemination venue during the early stages of the COVID-19

pandemic. However, how well this adoption of preprints will transfer to other

topics, where the speed of research does not put the same focus on preprinted

articles, is currently unknown.

4.5 - Discussion

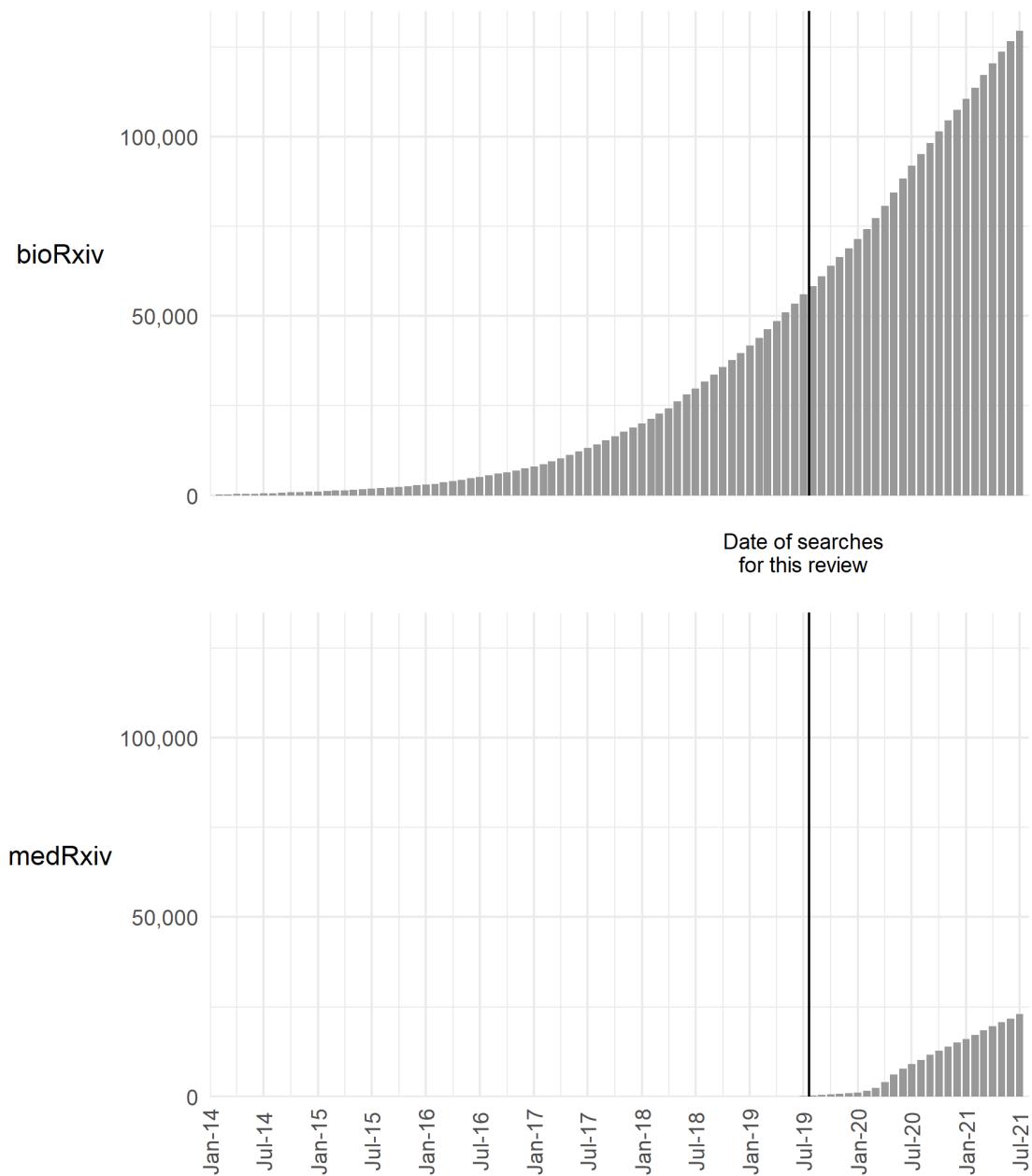


Figure 4.6: Growth of preprint repositories over time - Given the relative sizes of the preprint repositories at the time the searches for this review were conducted (bioRxiv n= 56,007, medRxiv n = 148), the relative number of hits returned by each is expected.

4.5.4 Reviewing Mendelian randomisations studies

Mendelian randomisation is a powerful analytical technique, using natural variation in participants genomes to (assuming the assumptions of the method are valid), though it's inclusion as an acceptable study design in this review was complicated by a number of factors.

Firstly, as this study design is relatively new, particularly when compared to randomised trials or cohort studies, the process and tools for systematically assessing them are not as well developed. A key example of this is in the absence of validated search filters for Mendelian randomisations studies. This limitation is further complicated by the varying terminology used to describe the method, particularly in the early years of it's application.

Additionally, there is currently no widely used risk-of-bias assessment tool for Mendelian randomisation studies. A recent commentary provided a checklist interpreting Mendelian randomisation studies, this guide includes reporting items in their quality checklist. While reporting quality is important, it is a separate consideration to internal validity, as discussed in Section @ref(). Similarly, a previous review of Mendelian randomisation studies used the Q-Genie tool, which was validated to assess the quality of genetic association studies in meta-analysis.¹⁴³ While this tool addresses the studies used, it does not access the additional methodological considerations of the analysis of the Mendelian randomisation analysis itself. For this review, I utilised the best available author-devised tool, sourced on a recent review of systematic reviews of Mendelian randomisation studies.

As a further stumbling block, Mendelian randomisation, particularly when using a two-sample summary data design, is a form of analysis that lends itself to multiple exposure-outcome comparisons. This is particularly relevant to the consideration of bias due to missing evidence. As an example, through snowballing and other measures, I identified at least one relevant Mendelian randomisation study that had not been identified by the search strategy.¹⁴⁴ On review of this paper, the search would not have been expected to find it given the absence of any lipid-related

4.5 - Discussion

keywords in the title and abstract. The study examined the association between lipid fractions with Alzheimer’s disease as one of many risk factors for the condition. Studies such as this can introduce bias into a systematic review, as it is commonly only those risk factors that show a statistically significant result that are reported in the abstract and so are captured by the search. This may bias systematic reviews, including this one, as the analysis of multiple risk factors against a single outcome within a single publication becomes more common. These studies are described as “unknown unknown’s” in the context of the RoB-ME tool, and are particularly challenging (as opposed to an analysis that was insufficiently reported to be included in the statistical analysis, or the “known unknown’s”).

Useful future work to improve the methodology for inclusion of Mendelian randomisation studies in systematic reviews should involve the development of a validated search filter for this study design.^{145,146} Alternatively, in better-resourced reviews, a dedicated search for “risk factors” and “dementia” and “Mendelian randomisation”, followed by manual review of studies that look across multiple risk factors, would be advisable. This was not feasible in the context of this review, given the large number of records to be screened, even when using study design filters (n=16,109). Additionally, the value of methods that supporting the traditional bibliographic database search, such as snowballing (forwards and backwards citation chasing) and communication with relevant topic experts should not be underestimated. Finally, development of a risk-of-bias assessment tool by a panel of methodologists and analysts would be of substantial benefit.

4.5.5 Comments on the process

As part of the reflective element of this thesis, I collated my experiences on performing a systematic review.

Protocol registration

While the protocol was registered on the Open Science Framework largely to allow for the sharing of associated documents such as the proposed search strategy, anecdotal evidence from colleagues and collaborators suggested that the findability of the protocol was limited. In hindsight, I should also have cross-posted the protocol on PROSPERO, the international prospective register of systematic reviews,¹⁴⁷ (which does not allow for the uploading of related files).

People management

This review provide an excellent opportunity to gain experience in managing a team of researchers. However, due to the need for dual screening and data extraction over a long period of time, a number of external researchers became involved in this review. I found the people-management aspect particularly challenging, and in retrospect could definitely have improved the process through better communication of expectations and deadlines.

Open data sharing

As discussed in Section 4.4.6, many primary studies did not report important elements, and so these could not be extracted. This limitation was compounded by the expected low response rate to requests for further information from primary authors (although, in hindsight, the form of contact used (email) has been shown to be less successful in eliciting responses from authors when compared with telephoning¹⁴⁸).

While contacting authors is worthwhile, as it can substantially change the conclusion of a systematic review¹⁴⁹ and is not too costly to systematic reviewers,¹⁵⁰ a far

preferable option is that the authors of primary studies readily deposit all relevant study data at the point of publication.

Based on my experience of extracting data for this review, I co-wrote an guidance article to aid primary prevention scientists in preparing and sharing their data so that it can easily be incorporated into a evidence synthesis exercise, using a trial of mindfulness interventions as an case study.¹⁵¹

Similarly, a substantial amount of time and effort has gone into making the data obtained by this review openly available to other researchers.

4.5.6 Strengths and limitations

Strengths

I believe there are four aspects where this review is distinct from those reviews already available in the published literature (as identified by):

- *Comprehensiveness*: While several reviews of this research topic exist,^{136–139} the overlap between the list of studies included in each is not 100%. As part of this review, I have not only performed a original search of primary literature databases, but have also screened the reference lists of comparable reviews to ensure no study has been omitted. In addition
- *Structured risk of bias assessment*: The majority of the highly cited reviews on this topic either do not formally consider the risk of bias in the observational studies they include or do not use an appropriate domain-based assessment tool (e.g. ROBINS-I/E). This is important area in which this thesis can add value, as based on the risk-of-bias assessments I have performed to date, several primary studies are at high risk of bias and this should be reflected in the findings of any review on this topic.

- *Inclusion of preprints:* Unlike other available reviews and enabled by the tool described in Chapter 3, this review systematically searched preprinted health-related manuscripts as a source of grey literature. As part of this chapter, I plan to examine the extent of the additional information provided to the review by the inclusion of preprints.
- *Contribution to methods work:* A large part of this review was the associated work on improving research synthesis methods. This work is detailed as relevant throughout the Chapter, often referring to additional work detailed in the . In addition this review was used to pilot an upcoming risk of bias tool

Limitations

The primary limitation of this review is that several included studies used data from EHR databases, which come with serious concerns regarding validity^{152 153,154}. Relatedly, several studies which made use of electronic health record database did not report the specific code lists used, potentially introducing substantial heterogeneity between effect estimates. An empirical example of the effect of differing EHR code list is presented as part of the analysis in Chapter 4 (see Section 5.4.4).

In addition, the fact that only a sample of records were dual screened at the title/abstract and full-text stages is a potential limitation, as there is a chance that some eligible records could have been excluded. However, evidence from assessments of inter- and intra-rater reliability indicate that is is not a major concern.

One particular limitation with regards to the risk of bias assessment is the fact that the ROBINS-E assessments were performed without the tool being finalised. This meant that there were no signalling questions to guide the domain-level risk of bias assessment, which may have influenced the accuracy with which domain-level judgements were assigned. However, there is no published empirical evidence supporting the need for signalling questions, and assessment of inter-rater reliability

4.6 - Conclusions

across the different tools did not indicate a specific problem with the ROBINS-E assessments. In fact, low agreement was common across the tools, though this is expected based on the available literature.

One further limitation is the fact that the risk of bias due to missing evidence assessment, combined with some empirical evidence that some studies were missed by the search but contained relevant studies is a definite limitation of this review (see Section @Ref(rev-discussion-MR) above for a fuller discussion of this issue with respect to Mendelian randomisations studies). Unfortunately, this is probably a common limitation across all reviews, based on the way in which increased sensitivity must be balanced with a reasonable workload.

4.6 Conclusions

In this chapter I have presented a comprehensive systematic review of the different sources of evidence available which examined the relationship between lipid levels and dementia use.

This work built on the tool introduced in the preceeding chapter (Chapter ??sys-rev-tools-heading)), and findings from this review are used though out the subsequent chapters: in Chapter 5, summary of the evidence guided the choice of analysis approach, ensuring that the new analysis was at risk of a different source of bias; while in Chapter 6, prospective cohorts identified by the review were contacted in an attempt to obtain individual participant data; finally, the cumulative effect measures calculated here are used as a key source of evidence for the triangulation exercise presented in Chapter ??discussion-heading).

4.7 References

5

Primary analysis of lipid regulating agents and dementia outcomes

Contents

5.1	Lay summary	91
5.2	Introduction	92
5.3	Methods	93
5.4	Results	105
5.5	Discussion	121
5.6	References	127

5.1 Lay summary

Electronic health record (EHR) databases are large collections of patients data, used to manage patient administration and care. Under these systems, whenever a patient attends their GP, their clinical data is recorded in a central database using a standard set of codes. These databases have several advantages over traditional methods of data collection, including the number of people they contain and the length of time for which participants are followed. This is particularly important when studying diseases such as dementia, which may begin to develop in patients

long before symptoms are seen.

This analysis makes use of the Clinical Practice Research Datalink (CPRD), which contains the electronic medical records of more than 3 million people from general practices across the UK. Using this data, the analysis presented in this chapter examined the effect of treatment which lower cholesterol levels, such as statins, on the risk of dementia and related outcomes.

No evidence for an effect of lipid regulating agents effect on the risk of Alzheimer's disease was found, with the exception of a slightly increased risk in those prescribed a certain type of drugs called fibrates. In contrast, I found an increased risk of vascular and other (i.e. non-Alzheimer's) dementia with lipid regulating agent use.

This increased risk is unexpected, and is very likely to be due to the presence of bias in the analysis. This bias, called "confounding by indication", is caused when those who are prescribed a statin are more at risk of vascular dementia for a range of reasons, and so it appears as if statins are harmful. However, despite this limitation, the analysis presented provides an important source of information which will be used in later chapters.

5.2 Introduction

In this Chapter, I present the results of an analysis of a large population-based electronic health record dataset to investigate the relationship between lipid regulating agent (LRA) use and dementia outcomes.

The analysis aims to address important two limitations of the current evidence base as identified by the systematic review presented in Chapter ??.

Firstly, it explicitly examines vascular dementia as an outcome. The systematic review presented in the previous chapter identified an evidence gap around the effect of lipid regulating agents on the risk of vascular dementia. As triangulation

exercises require as many diverse sources of evidence as possible, this analysis provides a source of information on this outcome.

Secondly, and in a similar vein, the analysis intentionally takes a different analytical approach to that most commonly used to examine the effect of statins on dementia (as identified by the systematic review). Specifically, this involved a concerted effort to address immortal time bias though use of a Cox Proportional time-varying treatment indicator. This approach provides a evidence source at risk of a distinct set of biases due to the alternative analytical strategy that will be incorporated into the triangulation exercise presented in Chapter 7.

5.3 Methods

5.3.1 Study protocol

An *a priori* protocol for this study was published,¹⁵⁶ and amendments to this are recorded in Appendix A.3.1.¹⁵⁷

5.3.2 Data source

Previously known as the General Practice Research Database (GPRD), the Clinical Practice Research Datalink (CPRD) is a large population-based, electronic health record (EHR) database.¹⁵⁸ The database has been collecting primary care data from participating practices across England since 1987.^{159,160} It contains the primary care records for more than 10 million primary care patients in England, and is broadly representative of the UK population in terms of age, sex and ethnicity.^{158,161}

To avoid the ambiguity of interpreting free-text clinical notes and to allow for easy analysis of the resulting data, the CPRD primarily collects data using a predefined coding system known as Read codes.¹⁶² All clinical events, included clinical test

results and diagnoses, can be identified by a specific Read code. The codes use a nested approach (see Table 5.1), with the initial characters defining broad diagnostic topics (e.g. Eu... - Mental and behavioural disorders), while subsequent characters provide additional information on the specific condition diagnosed (e.g. Eu001 - Dementia in Alzheimer's disease with late onset).

Table 5.1: Example of CPRD Read code hierarchy, showing how “Dementia in Alzheimer’s disease with late onset” (*Eu001*) is nested under the top-level of “Mental disorders” (*Eu...*). Broad topics are specified using the initial two alpha-numeric characters of the Read code, while subsequent characters are used to define specific conditions and context.

Level	Read code	Term
1	E....	Mental disorders
2	Eu...	Mental and behavioral disorders
3	Eu0..	Organic mental disorder
4	Eu00.	Dementia in Alzheimer’s disease
5	Eu001	Dementia in Alzheimer’s disease with late onset

Lists of relevant codes for each of the index events, exposures and outcomes used in the analysis were created. Each of these are described in more detail in the following sections.

5.3.3 Cohort definition

This analysis included all participants registered at a participating practice between 1 January 1995 and 29 February 2016 who had a flag for “research quality” data. Records pre-dating the 1995 cutoff were included in the original CPRD extract obtained for this analysis. However, these were excluded from the analysis as data quality and reliability is thought to be higher after this date.¹⁶³ Additionally,

individuals with less than 12 months of continuous records prior to cohort entry were excluded, making the effective start date of the cohort 1 January 1996.

Participants were included in the study cohort if their record contained any of the following index events: a Read code for a diagnosis of hypercholesterolemia or related condition; a Read code for prescription of a lipid-regulating agent (such as statins); a total cholesterol test result of >4 mmol/L; or an LDL-c test result of >2 mmol/L. These index events allowed me to define a population of participants who were either at risk of hypercholesterolemia, as indicated by the elevated total or LDL cholesterol test results, or had already been diagnosed with it, as indicated by a diagnostic code/related prescription.

All events of interest were identified using predetermined code lists, which are available for inspection from the archived repository accompanying this analysis (data/code availability is discussed in Section 5.5.4).

The index date for a participant was defined as the date where the first relevant code or test result was recorded on their clinical record, and participants were followed up until the earliest of: an outcome of interest; death; end of follow-up (29 February 2016); or last registration date with their GP practice. Participants were removed from the sample if they were less than 40 years of age, had less than 12 months of “research quality” data, were simultaneously prescribed more than one lipid-regulating agent (due to the difficult of assigning these to a single exposure group), or were diagnosed with an outcome of interest before or on the date of the index event (i.e. had less than one full day of follow-up).

5.3.4 Exposures

We considered seven lipid-regulating drug classes based on groupings in the British National Formulary (BNF),¹⁶⁴ namely: statins, fibrates, bile acid sequestrants, ezetimibe, nicotinic acid groups, ezetimibe and statin (representing one treatment

5.3 - Methods

containing both drugs, rather than the two classes being prescribed concurrently), and omega-3 fatty acid groups.

A participant's drug class was assigned based on their first recorded prescription, and any drug switching was ignored in an effort to mimic an intention-to-treat approach. We did however examine how often the initial drug class altered according to one of three criteria:

- **stopped:** defined as the last prescription of the primary class being followed by at least six months of observation;
- **added:** defined as a second drug class being prescribed before the last prescription of the initial class; and
- **switched:** defined as a second drug class being prescribed after the last prescription of the initial class.

5.3.5 Outcomes

We considered five outcomes as part of this analysis: probable Alzheimer's disease, possible Alzheimer's disease, vascular dementia, other dementia, and a composite all-cause dementia outcome. When two or more outcomes were coded in a participant's clinical record, a decision tree was used to differentiate between them (see Figure 5.1). The diagnosis date of the outcome was determined by the first record of a relevant code.

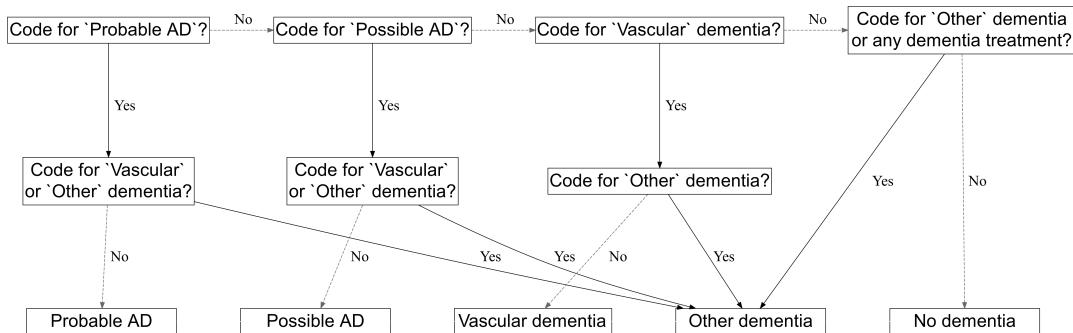


Figure 5.1: Decision tree for assigning dementia subtypes, based on the presence of Read codes in the patient's record. Note that an outcome of "Probable" or "Possible" Alzheimer's disease (AD) requires the absence of any vascular outcome codes.

Participants were censored at the earliest of date of: diagnosis; death; transfer out of the study cohort; or end of follow-up (29th February 2016).

5.3.6 Covariates

A range of additional variables were included in the analysis, intended to address the different distributions of potential confounding variables between those who were prescribed an lipid-regulating agent and those who were not.

Demographic covariates adjusted for included age and gender. Age was calculated at date of entry into the cohort. Socioeconomic status was proxied using the Index of Multiple Deprivation (IMD) 2010, which draws on seven domains (income; employment; education, skills and training; health and disability; crime; barriers to housing and services; living environment) to create an overall deprivation score for each of 32844 statistical geography areas in England. To help preserve patient privacy, IMD score is only available from the CPRD in twentiles, with 1 indicating

5.3 - Methods

the least deprived and 20 indicating the most deprived. Smoking and alcohol use was determined at index, and participants were categorised as current, former, or never users of each.

Body mass index (a summary measure calculated as $weight/height^2$), baseline total cholesterol and baseline LDL cholesterol measures were obtained, using the last recorded value prior to the index date. A variable indicating grouped year of entry into the cohort (<2000, 2000-2004, 2005-2009, >2010) was included to allow for changes in prescribing trends across the lifetime of the cohort. To assess healthcare utilisation, I adjusted for the average annual number of consultations between the beginning of a patients data and their entry into the cohort.

Finally, presence of a range of related conditions at baseline were accounted for, including cardiovascular disease, coronary bypass surgery, coronary artery disease, peripheral arterial disease, hypertension, chronic kidney disease, and Type 1 and Type 2 diabetes. In addition to adjusting for these covariates individually, a Charlson co-morbidity index (CCI) score was calculated for each participant. The CCI is a weighted index that uses presence and severity of a number of conditions to enable adjustment for the general health of a participant in terms of their mortality risk.¹⁶⁵ Inclusion of this index allowed me to attempt to adjust for the general health of patients included in the analysis.

Codelists for all covariates can be found in the archived data repository accompanying this analysis (see Section 5.5.4).

5.3.7 Missing data

Missing data are a recognised issue in electronic health records databases,¹⁶⁶ given that they contain administrative data, collected primarily for the purposes of patient management and care rather than academic research.

In this analysis, missing data were handled using a multiple imputation approach. Variables with missing observations were identified, and 20 imputed datasets were created.¹⁶⁷ Nominal variables with missing values were modelled using multinomial logistic regression, while continuous variables were modelled using linear regression. As per best practice, all variables used in the analytic model, including the outcome, were included in the imputation model.¹⁶⁸ Imputation was performed using the MICE (Multiple Imputation by Chained Equations) command in STATA¹⁶.

Missing data was only considered an issue for variables where a numerical test result was expected (e.g. BMI), or where a code existed for the absence of the condition (e.g. categorical smoking status). This approach was necessary, as absence of a code for other treatments or conditions (e.g. statin use or dementia) was assumed to indicate absence of the treatment/condition, rather than being considered missing.¹⁶⁶

Using multiple imputation to handle missing data is an alternative to a “complete case” approach,¹⁶⁹ where participants missing any covariate are dropped from the dataset. As a sensitivity analysis, I preformed and compared the results of both methods,¹⁷⁰ to investigate the impact of multiple imputation on the results.

5.3.8 Estimation methods

A Cox proportional hazards (PR) model was used to estimate the effect of statins on dementia outcomes. Cox PR models are defined, in general terms, as:

$$h(t) = h_o(t) \times \exp(b_1x_1 + b_2x_2 + \dots + b_px_p) \quad (5.1)$$

where:

- t is the survival time;
- $h(t)$ is the hazard function; and

- x_1, x_2, \dots, x_p are the covariates which determine the hazard function, while b_1, b_2, \dots, b_p are the coefficients for each covariate.
- $h_o(t)$ is the baseline hazard - when all x_i are zero, the $\exp()$ function resolves to 1.

As the values of b_i increases, value of $\exp(b_i)$ (i.e. the hazard ratio) increases in tandem, thus reducing the survival time, t .

A Cox PR model was chosen for this analysis as it inherently accounts for the length of time participants spend in each exposure group. Using this approach, time-at-risk can be properly attributed, thus mitigating the impact of immortal time bias. This is discussed in detail in the following section.

5.3.9 Immortal time bias and time-varying treatment indicators

Immortal time bias covers two distinct but related types of bias. The first presentation, the selection bias aspect (Panel A, Figure 5.2), occurs when time prior to the exposure is excluded leading to the exposed and control groups being followed up from different time points.¹⁷¹ For example, if the unexposed group are followed from a cholesterol test result, while the exposed group is followed from date of LRA prescription, any events that occur in the exposed group prior to the exposure event will be inappropriately excluded from the analysis.

The second presentation of immortal time bias is as a type of misclassification bias (Panel B, Figure 5.2). It occurs when the exposure time prior to the exposure date, and any events occurring within it, is inappropriately assigned to the exposed group. This second presentation appears to be common in the existing literature, as several of the studies included in the systematic review presented in Chapter 4 were identified as being at risk of immortal time bias following formal risk of bias assessment using the ROBINS-I tool (see Section 4.4.4).

5.3 - Methods

This analysis is primarily concerned with the second presentation of immortal time bias, as all participants were followed from a common index date (earliest of: date of raised cholesterol test results; hypercholesterolemia diagnosis; or LRA prescription). To address the potential for this second form of immortal time bias in the analysis, I employed a time-varying indicator of treatment status to correctly allocate time-at-risk to the exposed and unexposed groups.¹⁷¹

Under this approach, all patients start in the unexposed group, & contribute time-at-risk until they are prescribed a lipid regulating agent and move into the exposed group. Note, patients for whom prescription of a lipid regulating agent was the index event only contribute time to the exposed group (i.e. they enter the cohort and move into the exposed group on the same day).

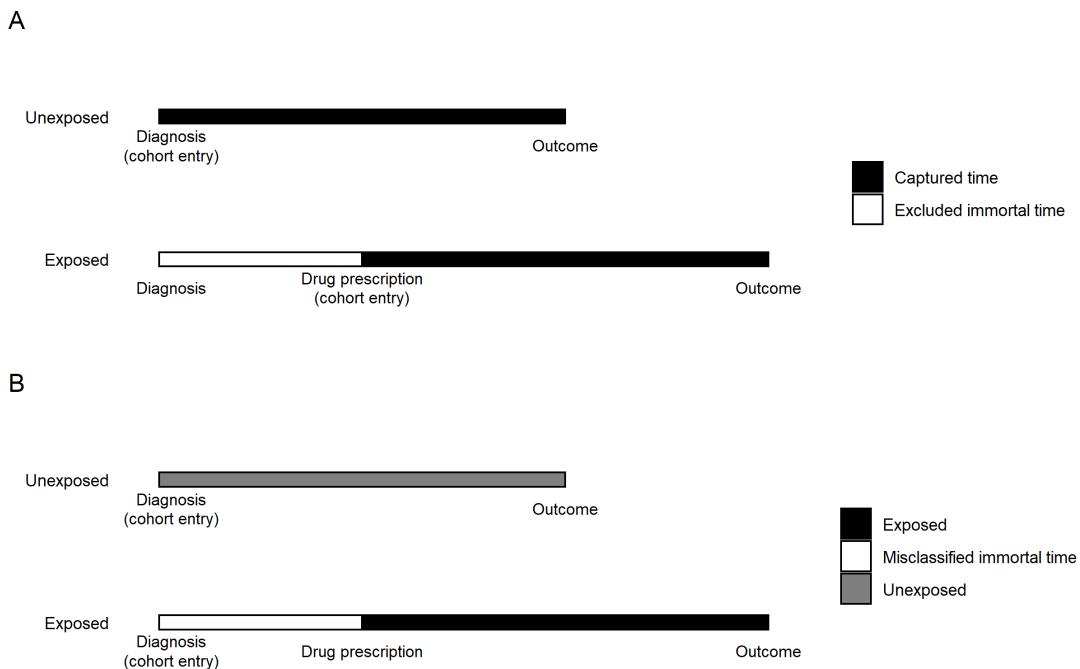


Figure 5.2: Diagram illustrating the two presentations of immortal time bias, as a selection bias (Panel A) and a misclassification bias (Panal B).

5.3.10 Time axis

As part of a Cox proportional hazard model, there is the option to use either absolute time in cohort or participants age as the time scale of interest.¹⁷²⁻¹⁷⁴ A model using age as the time axis inherently accounts, or adjusts, for participants age as a potential confounder of the exposure-outcome relationship. The main analyses presented all used age as the time axis.

5.3.11 Sensitivity analyses

The primary analysis examined the effect of a lipid regulating agent on dementia risk, stratified by outcome and drug class. To assess the robustness of the results, a number of sensitivity analyses were performed. These are described in the following sections.

Control outcomes

In addition to the primary outcomes of interest (described in Section 5.3.5), I extracted data on two additional control outcomes. The inclusion of control outcomes in observational analyses are a useful technique to assess the strength of uncontrolled confounding,¹⁷⁵ and these outcomes are usually classed as either “negative” or “positive” outcomes. Negative outcomes are those without a likely causal path between the exposure and outcome (see Figure 5.3 for a directed acyclic graph describing an ideal negative outcome).

Conversely, positive control outcomes are those with a known causal association with the exposure of interest, ideally sourced from large well conducted randomised controlled trials. Positive control outcomes are useful in observational epidemiology, as if the analysis can reproduce a known result for the control outcome, confidence in the result for the outcome of interest is increased.

Due to the wealth of data available on statins as a lipid regulating agent, control outcomes were chosen in reference to this drug class. Muscular backpain was chosen as a negative control outcome in this analysis. Despite observational analyses suggesting a link between statins and muscle pain (as opposed to more serious complications such as myopathy),¹⁷⁶ systematic reviews of the adverse events of statin use³⁰ and N-of-1 trials explicitly exploring the effect of statin use on muscle pain¹⁷⁷ have found no evidence of an effect. As such, if no evidence of an effect of statins on backpain is found in the analysis, then confidence in the results for the dementia outcomes is increased.

Additionally, incident ischemic heart disease was included as a positive control outcome, given the well-established protective effect of lipid-lowering treatment, via statins, on the risk of this condition.³⁰ Similar to the backpain outcome, if the analysis strategy can reproduce this known protective association, this will provide evidence that potential confounders have been sufficiently adjusted for.

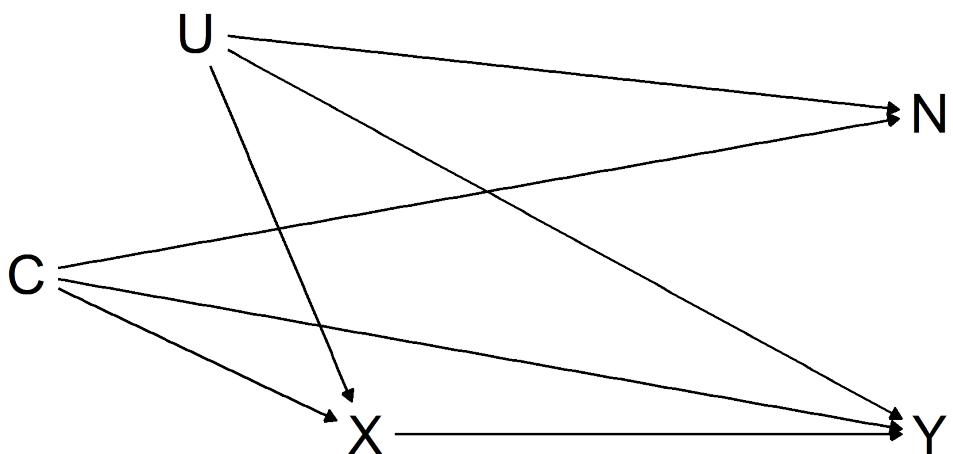


Figure 5.3: Causal diagram showing relationship between exposure X , outcome Y , confounders (measured C and unmeasured U) and an ideal negative outcome N . Note the absence of any arrow between X and N . In this scenario, any association observed between X and N is due to the presence of uncontrolled confounders U (assuming C has been adjusted for).

Impact of additional covariates

To observe the effect of adjusting for additional covariates, I ran a model unadjusted except for age (captured via the time axis in the Cox model) and gender, and compared the results with the full adjusted model.

Sensitivity cohorts

Two sensitivity cohorts were also created. The first stratified by year of entry into the cohort in an attempt to assess for time period effects. The second removed participants who may have been pregnant (coded as under 55 and female) to assess the robustness of the estimates, as statins are contraindicated in pregnancy.¹⁷⁸

Statin properties

As detailed in the introduction, the properties of statins may be important in their effect, based on the ability of lipophilic statins to cross the blood brain barrier (see Section 2.4.2 in the Introductory Chapter).

To explore whether any observed effect for the statin analysis varied by statin property, a sensitivity analysis was performed, stratifying by statin lipophilicity.

Impact of dementia code lists

As part of an exploratory analysis of the effect of the choice of code lists on the analysis, I created an alternative Alzheimer's disease and non-Alzheimer's dementia outcome using code lists from a previous study published in 2009 by Smeeth *et al.*¹⁷⁹

This previous study used a propensity matching approach to estimate the effect of statins on a range of outcomes, and a protective effect of statin use on all-cause

5.4 - Results

dementia (HR: 0.81, 95%CI: 0.69-0.96) and non-AD dementia (HR: 0.82, 95%CI: 0.69-0.97), but no evidence of an effect on AD (HR: 0.81, 95%CI: 0.49-1.35).

The code lists used were obtained through correspondence with the authors of that study, and are available for inspection (see Section 5.5.4).

5.4 Results

5.4.1 Patient characteristics

Of the 3,179,733 participants included in the extract, 1,684,564 met the inclusion criteria (Figure 5.4), with a total follow-up of 10,835,685 patient years at risk.

5.4 - Results

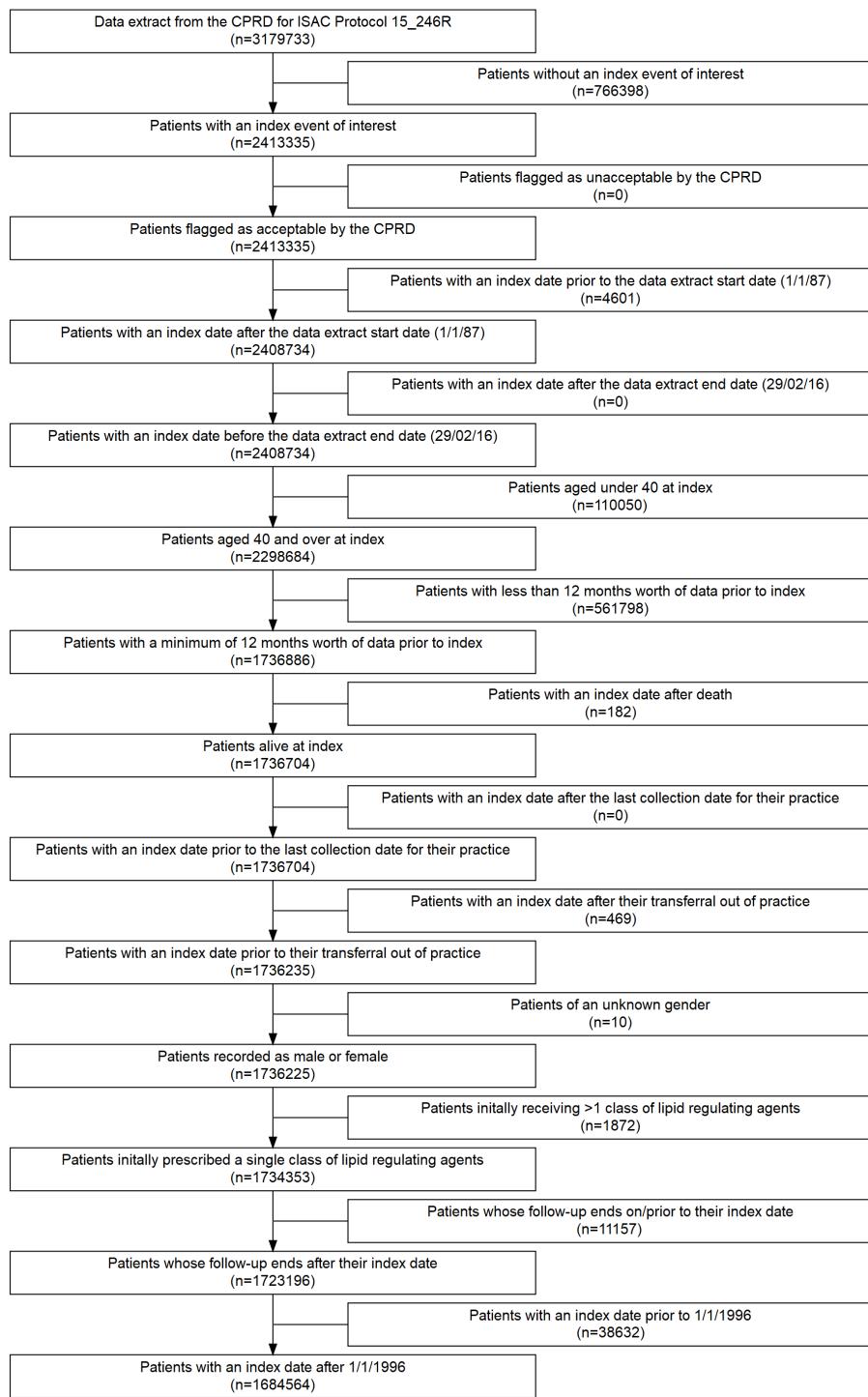


Figure 5.4: Attrition of CPROD participants as the eligibility criteria were applied. Most attrition was due to the absence of an index event of interest.

5.4 - Results

The median participant age at index was 57 years (Inter quartile range (IQR):48-67) and participants were followed up for a median of 5.9 years (IQR:2.7-9.7). During follow-up, an all-cause dementia diagnosis was recorded for 41,830 patients (12,647 probable AD, 9,954 possible AD, 8,466 vascular dementia, 10,763 other dementia). The distribution of baseline characteristics across the seven drug classes can be seen in Table 5.2.

5.4 - Results

5.4 - Results

Table 5.2: Patient characteristics by drug class. Summary statistics are presented as “% (N)” unless otherwise specified in the variable name.

	Whole Sample	None	Statins	Bile acid sequestrants	Ezetimibe	Ezetimibe & Statins	Fibrates	Nicotinic acid groups	Omega-3 Fatty Acid Groups
Sample size (N)	1,684,564	1,087,704	585,528	5,396	763	127	3,889	165	992
Year of cohort entry (median)	2006	2007	2004	2005	2004	2005	2001	2001	2005
Female	53.0% (893,174)	56.2% (610,950)	47.1% (276,043)	66.4% (358,85)	54.5% (416)	52.8% (67)	38.6% (1500)	55.2% (91)	52.6% (522)
Age at cohort entry (median)	57	54	62	57	60	57	58	62	56
CAD	0.4% (7133)	0.1% (589)	1.1% (6465)	0.1% (6)	0.9% (7)	0.0% (0)	1.4% (53)	0.0% (0)	1.3% (13)
CBS	0.3% (5699)	0.1% (682)	0.8% (4926)	0.1% (4)	0.4% (3)	0.0% (0)	2.0% (78)	0.0% (0)	0.6% (6)
CVD	2.1% (34899)	1.1% (11619)	3.9% (22977)	1.6% (86)	2.6% (20)	2.4% (3)	4.4% (170)	4.2% (7)	1.7% (17)
Charlson (ever > 0)	30.6% (516,135)	25.1% (272,642)	40.7% (238,403)	42.5% (2292)	41.7% (318)	24.4% (31)	50.8% (1976)	43.6% (72)	40.4% (401)
IMD-2010 (median) (mean/SD)	9	8	9	8	9	13	10	10	10
Consultation rate (mean/SD)	5.4 (5.4)	5.0 (5.0)	6.2 (6.1)	8.6 (7.4)	7.4 (6.6)	4.8 (4.3)	7.1 (6.2)	9.2 (7.8)	8.0 (8.0)
Alcohol (ever)	85.9% (1,447,151)	86.6% (941,648)	84.7% (496,110)	82.8% (4468)	84.0% (641)	87.4% (111)	82.9% (3223)	83.0% (137)	82.0% (813)
Smoking (ever)	51.1% (861,355)	47.1% (511,826)	58.6% (343,074)	55.2% (2978)	57.5% (439)	60.6% (77)	60.2% (2341)	52.7% (87)	53.7% (533)
BMI (mean/SD)	27.0 (5.3)	26.7 (5.2)	27.7 (5.3)	26.8 (5.8)	28.1 (5.7)	28.1 (4.9)	29.0 (5.2)	26.4 (5.0)	26.9 (5.5)
PAD	0.7% (12613)	0.4% (4039)	1.4% (8424)	0.9% (47)	0.9% (7)	0.8% (1)	1.9% (75)	6.1% (10)	1.0% (10)
Hypertension	16.0% (269,804)	11.5% (124,604)	24.4% (143,101)	12.8% (692)	23.9% (182)	25.2% (32)	25.8% (1002)	21.2% (35)	15.7% (156)
Total cholesterol (mean/SD)	5.7 (10.1)	5.5 (6.4)	6.2 (15.3)	5.3 (1.3)	7.1 (26.5)	6.7 (1.5)	6.4 (5.6)	5.4 (1.5)	5.6 (1.6)
LDL cholesterol (mean/SD)	3.6 (4.9)	3.4 (5.3)	4.0 (3.7)	3.1 (1.0)	3.9 (1.1)	4.2 (1.0)	3.3 (1.8)	3.4 (0.9)	3.2 (1.0)
CKD	0.1% (1295)	0.1% (740)	0.1% (545)	0.1% (6)	0.1% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.3% (3)
Type 1 Diabetes	0.2% (4037)	0.1% (785)	0.5% (3196)	0.3% (14)	1.0% (8)	0.8% (1)	0.8% (31)	0.6% (1)	0.1% (1)
Type 2 Diabetes	2.9% (48557)	1.1% (11797)	6.1% (35941)	2.3% (123)	5.4% (41)	4.7% (6)	15.8% (614)	4.2% (7)	2.8% (28)

Abbreviations:

LRA - Lipid regulating agent; IMD - Index of Multiple Deprivation; BMI - Body Mass Index; CAD - Coronary Arterial Disease; CBS - Coronary Bypass Surgery; CVD - Cardiovascular disease; PAD - Peripheral arterial disease; CKD - Chronic Kidney Disease; SD - Standard deviation.

5.4 - Results

A substantial majority (98.1%) of participants prescribed a lipid-regulating agent were prescribed a statin. I excluded the “Ezetimibe and statins” and “Nicotinic acid groups” classes from subsequent analysis based on the extremely small number of participants in these groups (N=127 & N=168 respectively; see Table 5.2). Note that the “Ezetimibe and statins” treatment group represent those prescribed a single treatment containing both ezetimibe and statins, rather than those where the two treatments were prescribed concurrently.

Table 5.3: Participants who stopped, switched or added treatments by initial treatment type.

	Whole Sample	Statins	Bile acid seques-trants	Ezetimibe	Ezetimibe & Statins	Fibrates	Nicotinic acid groups	Omega-3 Fatty Acid Groups
Stopped	6.9% (115899)	19.1% (111798)	56.1% (3028)	19.7% (150)	12.6% (16)	12.3% (478)	44.8% (74)	35.8% (355)
Added	1.6% (27441)	4.4% (25990)	3.6% (192)	19.0% (145)	3.9% (5)	21.6% (841)	3.6% (6)	26.4% (262)
Switched	0.9% (14935)	2.0% (11996)	11.3% (612)	34.6% (264)	64.6% (82)	44.0% (1713)	45.5% (75)	19.5% (193)

Definitions:

Stopped - last prescription of the primary drug class followed by at least six months of observation with no further prescriptions; Added - second drug class prescribed before the last prescription of the initial class; Switched - second drug class being prescribed after the last prescription of the initial class.

The stopping, addition and switching of drug classes was common across all drug classes (Table 5.3).

5.4.2 Missing data

Full covariate information was available for 450,234 participants (26.7%). Six key variables had some missing data: IMD 2010 score was missing for 625,788 participants (37.1%), because it is only recorded for English practices; alcohol status was missing for 269,526 participants (16%); smoking status was missing for 84,424 participants (5%); BMI, or a calculated BMI from height and weight measurements,

5.4 - Results

was missing for 266,672 participants (15.8%); baseline total cholesterol was missing for 119,675 participants (7.1%); and baseline LDL cholesterol was missing for 787,289 participants (46.7%).

5.4.3 Primary analysis

Crude rate ratios were calculated for each outcome and class of interest and are presented in Table @ref(tab:)

The results of the primary analysis using the fully adjusted Cox proportional hazards model with participant age as the time scale are presented for each drug/outcome combination in Figure 5.5.

For each outcome, the overall “Any drug” estimate was driven by the statin subgroup, based on its large size relative to the other drug classes.

5.4 - Results

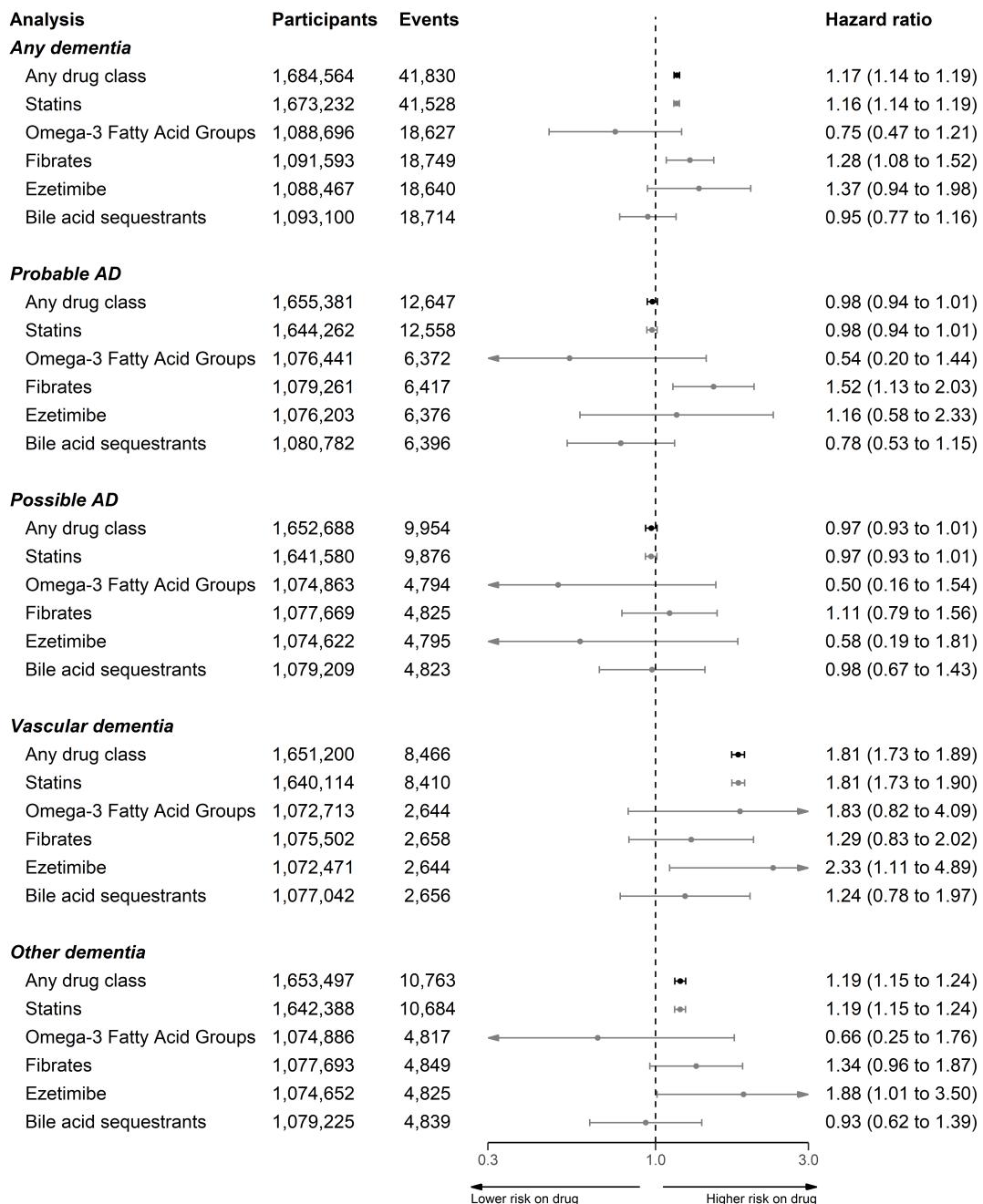


Figure 5.5: Results from primary analyses of CPRD data using the fully adjusted model and participant age as the time scale.

Alzheimer's disease

No evidence was found for an effect of lipid-regulating agents on probable (HR:0.98, 95%CI:0.94-1.01) and possible (HR:0.97, 95%CI:0.93-1.01) Alzheimer's disease

when compared with no treatment, with the sole exception of fibrates on probable Alzheimer's disease (HR:1.28, 95%CI:1.08-1.52).

Non-Alzheimer's disease dementias

In contrast to the findings for Alzheimer's disease outcomes, lipid-regulating agents were associated with an increased risk of a subsequent diagnosis of vascular dementia (HR:1.81, 95%CI:1.73-1.89) or other dementia (HR:1.19, 95%CI:1.15-1.24). Again this effect was driven mainly by the statin subgroup, but there was some evidence that ezetimibe was associated with an increased risk of vascular (HR:2.33, 95%CI:1.11-4.89) and other (HR:1.88, 95%CI:1.01-3.5) dementia.

All-cause dementia

For the composite all-cause dementia outcome, I found treatment with a lipid-regulating agent was associated with a slightly increased risk (HR:1.17, 95%CI:1.14-1.19), but the magnitude of the association was not as extreme as that observed for the vascular dementia subgroup. There was also some evidence that fibrates were associated with increased risk of all-cause dementia (HR:1.28, 95%CI:1.08-1.52).

5.4.4 Sensitivity analyses

The results of the series of sensitivity analyses performed are described in the following sections.

Complete case versus imputed data

In almost all cases, the use of imputed data resulted in a marginal attenuation of the effects observed when using a complete cases analysis. It should be noted that due to the large amount of missing data (e.g. 787,289 participants (46.7%)

5.4 - Results

were missing a baseline LDL cholesterol measure), the number of participants included in the complete case analysis was substantially smaller than that included when using imputed data. In this case, though the overall position of the effect estimates does not change substantially when using the imputed dataset, there is a noticeable gain in power.¹⁶⁷

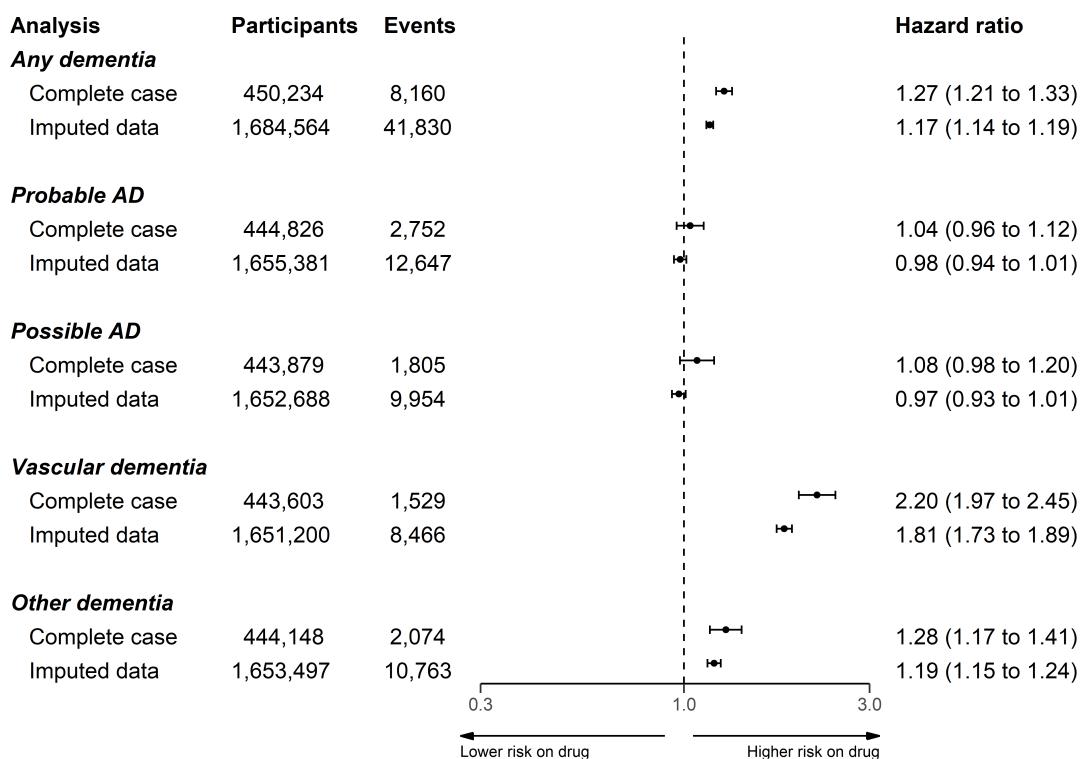


Figure 5.6: Comparison of complete case versus imputed data analyses.

Control outcomes

Following the primary analysis, the fully adjusted model was used to estimate the effect of treatment with a statin on the two control outcomes of back pain (negative) and ischemic heart disease (positive). The results of this analysis are presented in Figure 5.7.

For the negative control, there was some evidence that treatment with a statin was associated with an increased risk of back pain (HR: 1.04, 95%CI: 1.03-1.05), suggesting there may be some residual confounding. However, statin prescription was also associated with a substantially increased risk of ischemic heart disease (HR: 1.62, 95%CI: 1.59-1.64).

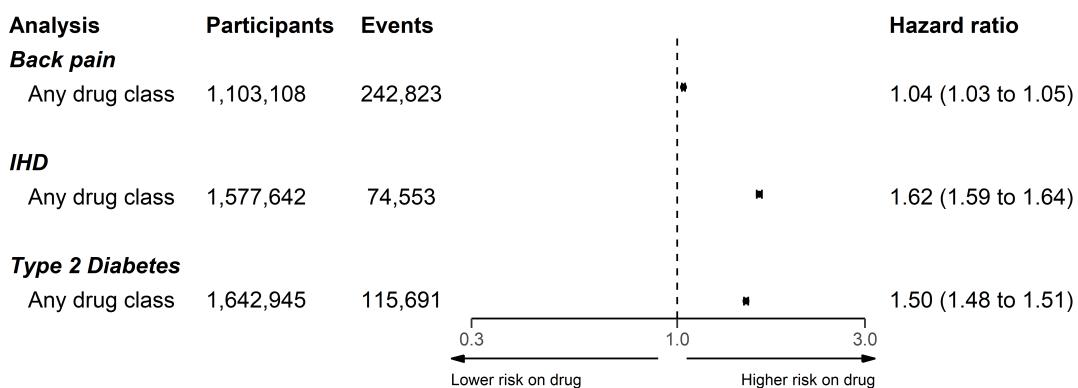


Figure 5.7: Results of control outcome analysis.

Impact of additional covariates

The results of three models adjusted for age only, age and sex, and the full covariates respectively, are presented in Figure 5.8.

These models were used to estimate the impact of adjustment for additional covariates. Note that obtaining a completely unadjusted model is not possible, as age was used in the Cox model as the time scale.

Adjustment for additional covariates beyond age and sex had a limited impact on the observed effect estimates, with the exception of the Probable AD outcome. In this case, adjustment for the full set of covariates attenuated to the null the protective effect observed when adjusting only for age and sex.

5.4 - Results

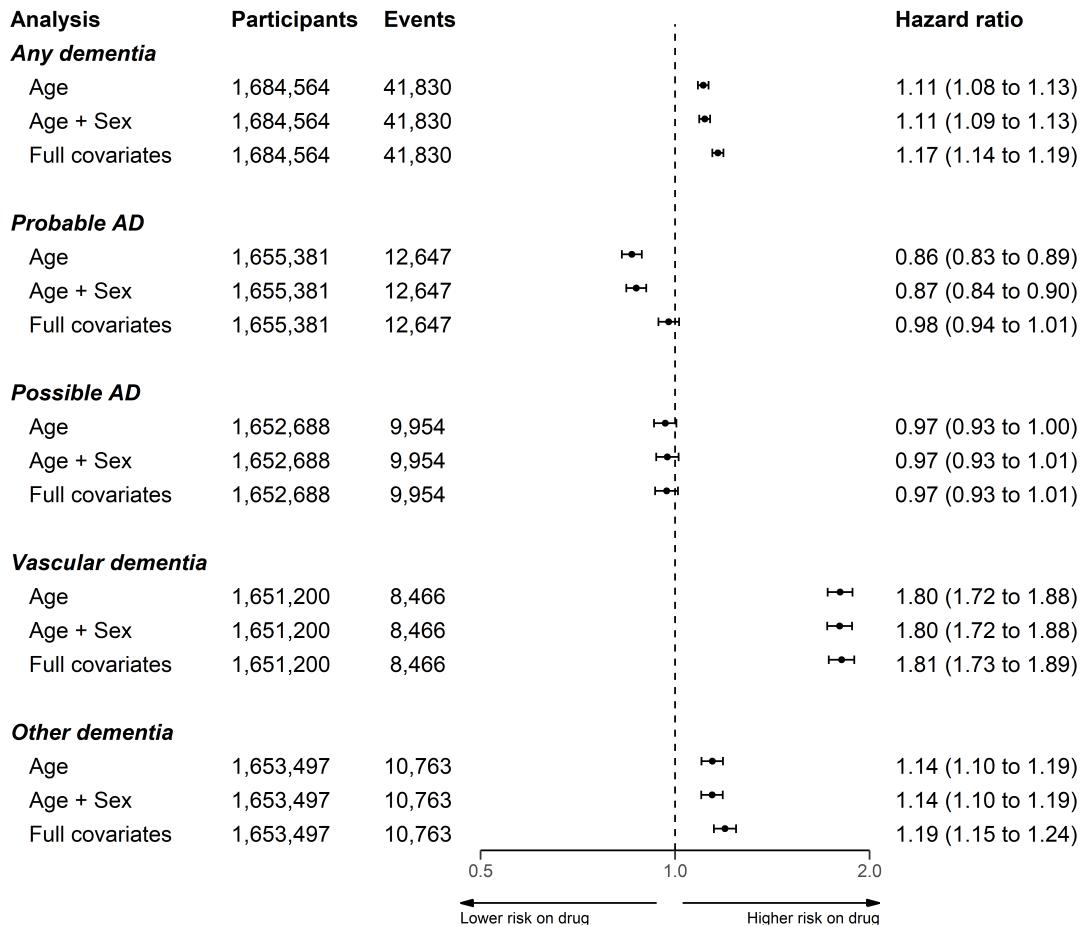


Figure 5.8: Results of three models adjusting for a different set of covariates.

Sensitivity cohorts: Entry year

When stratifying based on year of entry to the cohort, I observed no variation in risk

by time period in any subgroup except for probable Alzheimer's disease (Figure 5.9).

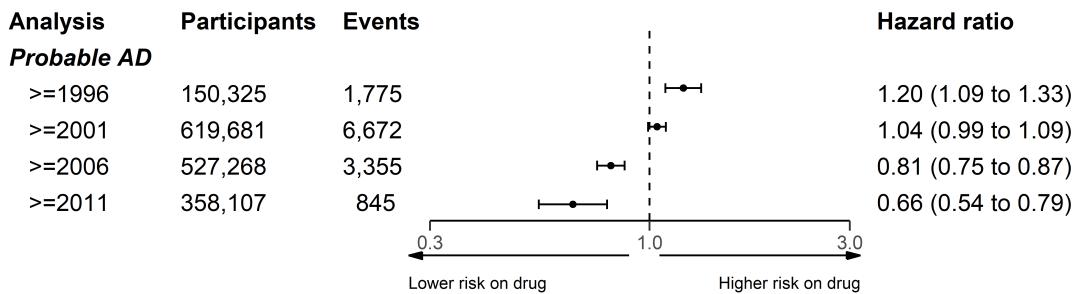


Figure 5.9: Analysis of any lipid regulating agent on Probable AD outcome, stratified by grouped year of cohort entry.

On the assumption that this variation could be caused by changes in the frequency of codes used to define Probable AD in the cohort, I performed a *post-hoc* investigation of the frequency of each diagnoses stratified by year of entry (Table 5.4). While the total frequency of any dementia outcome declines in more recent strata, likely due to the limited follow-up inherent to these groups, the decline in frequency is relatively constant across all dementia subtypes.

Table 5.4: Frequency of diagnoses by grouped year of cohort entry

Year of cohort entry	No dementia	Probable AD	Possible AD	Vascular dementia	Other dementia	Total
>=1996	148550 (95.9%)	1775 (1.1%)	1677 (1.1%)	1345 (0.9%)	1585 (1.0%)	154932
>=2001	613009 (96.3%)	6672 (1.0%)	5711 (0.9%)	4857 (0.8%)	6073 (1.0%)	636322
>=2006	523913 (98.1%)	3355 (0.6%)	2169 (0.4%)	1890 (0.4%)	2506 (0.5%)	533833
>=2011	357262 (99.4%)	845 (0.2%)	397 (0.1%)	374 (0.1%)	599 (0.2%)	359477
Total	1642734 (97.5%)	12647 (0.8%)	9954 (0.6%)	8466 (0.5%)	10763 (0.6%)	1684564

Sensitivity cohorts: Pregnancy

In the second sensitivity cohort, removing participants aged 55 and under at index from the analysis had minimal effect on the effect estimates (Figure 5.10).

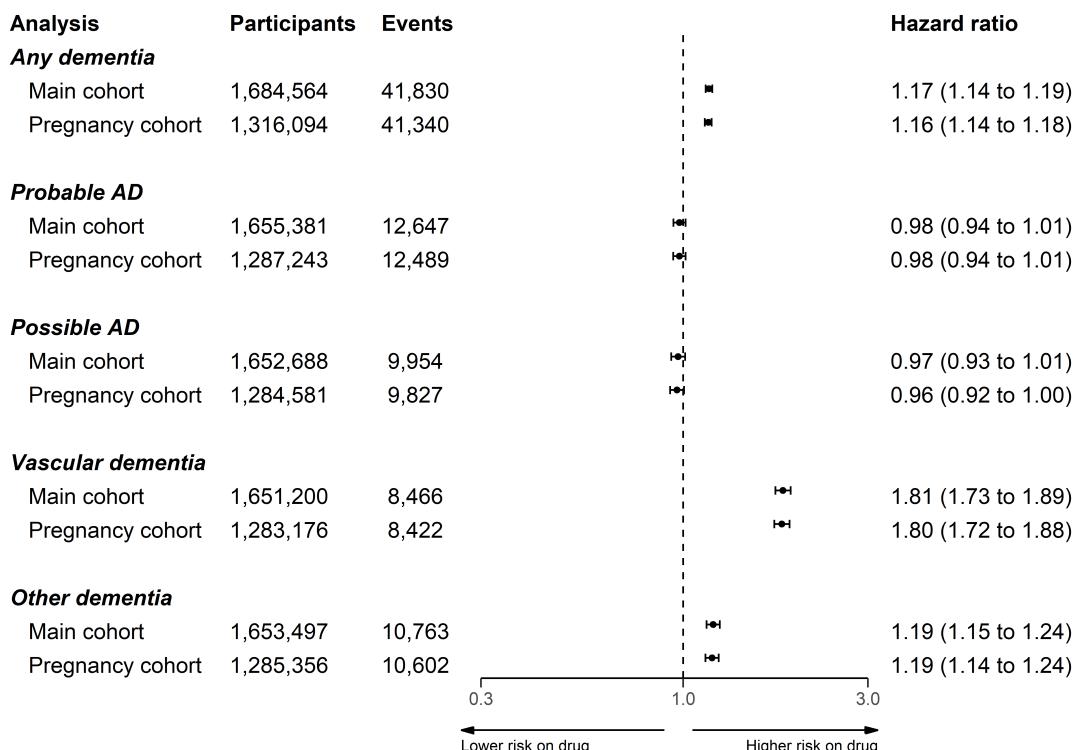


Figure 5.10: Comparison of analysis using main cohort and a cohort with potentially pregnancy women removed.

Statin properties

In the cohort, statins with lipophilic properties were much more frequently prescribed than hydrophilic statins (Table 5.5). Additionally, there is evidence for a increasing tendency to favour hydrophilic statins in recent years, as the proportion of lipophilic statins prescribed fell from 18.2% in 1996-2000 to <1% in 2011-2016.

5.4 - Results

Table 5.5: Summary of statin properties (lipophilicity vs hydrophilicity) by grouped year of prescription.

Prescription Year Group	Hydrophilic	Lipophilic	Total
>=1996	7037 (18.2%)	31531 (81.8%)	38568
>=2001	21427 (10.3%)	187018 (89.7%)	208445
>=2006	3566 (1.6%)	217726 (98.4%)	221292
>=2011	1115 (0.9%)	119035 (99.1%)	120150

When stratifying by statin properties, hydrophilic statins were less harmful in the any, vascular and other dementia outcomes compared to lipophilic statins (Figure 5.11). Additionally, in the AD outcomes, hydrophilic statins had a small protective effect, compared to the absence of evidence for an effect for lipophilic statins.

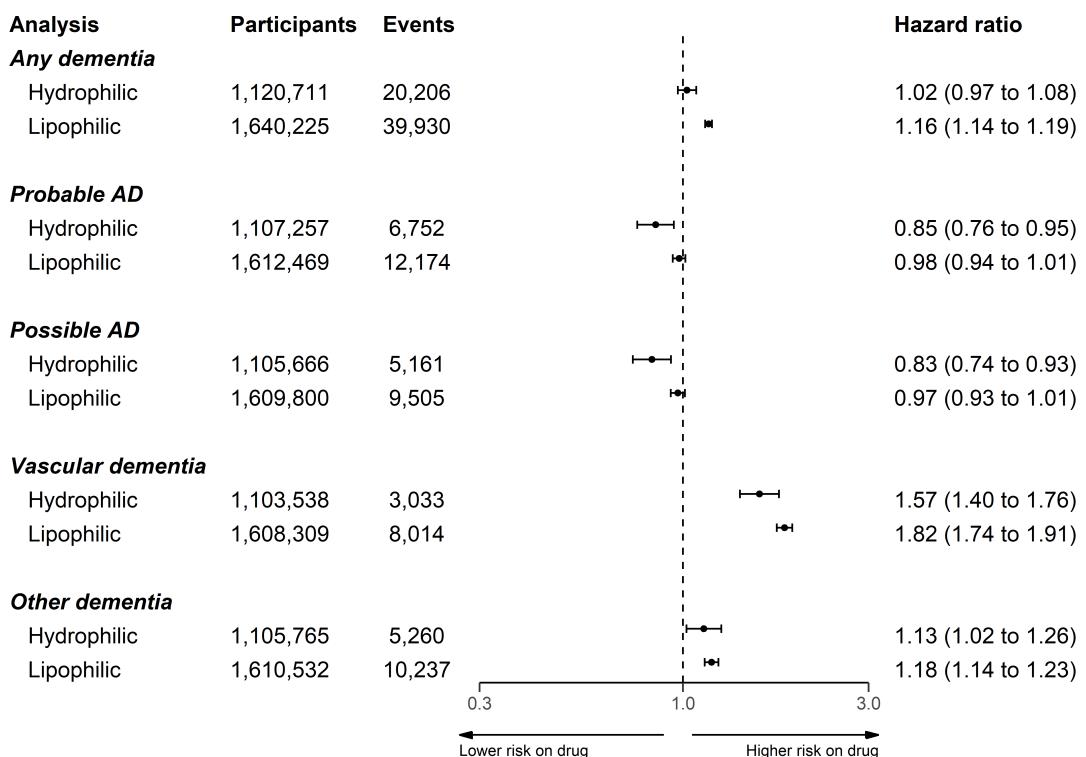


Figure 5.11: Analysis stratified by statin properties (hydrophilic vs lipophilic)

Comparing codelists

When using the Smeeth et al code lists to define dementia outcomes, effect estimates of HR: 1.19 (95%CI: 1.07-1.32) and HR: 1.33 (95%CI: 1.26-1.42) were obtained for the Alzheimer's disease and non-Alzheimer's dementia outcomes respectively.

However, comparison of the results using the two sets of code lists was deemed less useful following a comparison of the codes used. While all of the codes used to define Alzheimer's in the Smeeth paper are included in the Probable Alzheimer's code-list (see Figure 5.12), I included several additional codes used to define this outcome (including, for example, "Eu00013: [X]AD disease type 2"). Additionally, several of the codes used to define "Possible Alzheimer's" in this analysis are included in the "Other dementia" code list used by Smeeth.

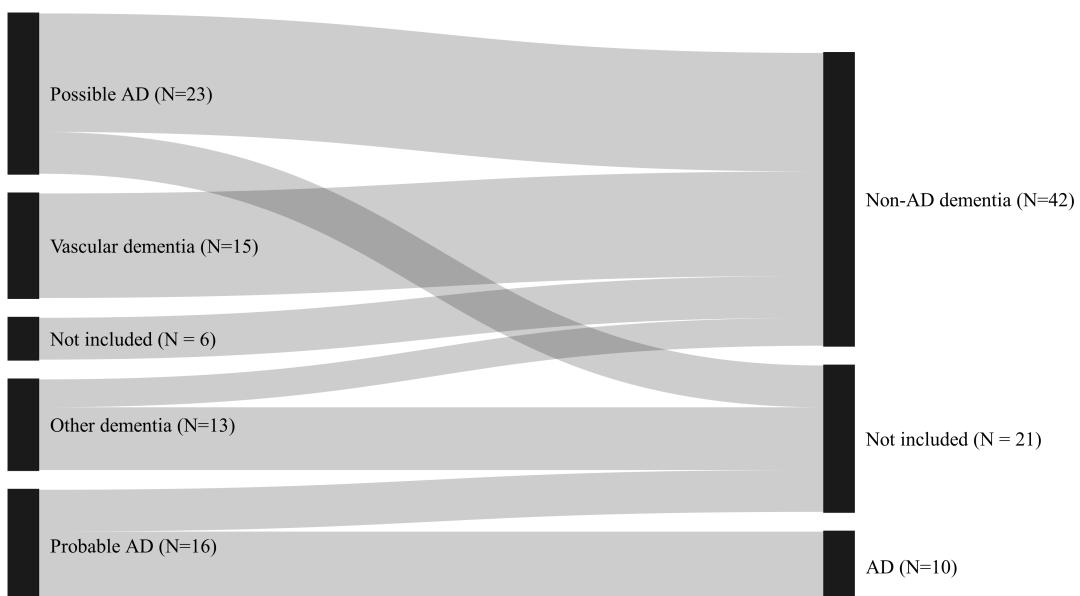


Figure 5.12: Sankey diagram comparing the codes used in this analysis with those used in the Smeeth *et al* paper.¹⁷⁹ The outcomes and number of codes contributing to each are presented (the Smeeth *et al* outcomes are on the right hand side of the figure). The joining lines showing the overlap between the categories in the two analyses.

This analysis serves to illustrate the importance of the code lists chosen to

define the outcomes of interest, particularly if they are competing (e.g. AD vs non-AD dementia).

5.5 Discussion

5.5.1 Summary of findings

Lipid-regulating agents showed no evidence of an effect on probable and possible Alzheimer's disease when compared with no treatment, but were associated with increased risk of an all-cause dementia, vascular dementia and other dementia diagnosis. The effect observed in each case was driven by the statin subgroup, which included a substantial majority of participants. For the other drug classes, no association was found with any outcome, with two exceptions being that ezetimibe was associated with increased risk of vascular and other dementia, while fibrates were associated with increase risk of all-cause dementia and probable Alzheimer's disease.

The effect estimates were robust to the exclusion of potentially pregnant participants, and for all outcomes except Probable AD, no variation across grouped year of entry was observed. When looking at the statin subgroup alone, statin properties appeared to have a modifying effect, with hydrophilic statins being less harmful in the any, vascular and other dementia outcomes compared to lipophilic statins.

5.5.2 Interpretation of results

This section will expand on a potential explanation for the observed results detailed above. However, as the comparison of evidence across different sources is the aim of the triangulation exercise presented in later chapters, the section will not provide a detailed comparison with other published literature, except where needed to illustrate a methodological point. For a comparison of the result presented above

with those from the systematic review (Chapter 4) and the individual patient data analysis (Chapter 6), see Section 7.

A likely explanation for the observed increased risk of vascular and other dementia with lipid regulating agent use is residual confounding by indication. While the term has been used to describe different source of bias in epidemiological analyses,¹⁸⁰ “confounding by indication” is used here to described the role of risk factors that both prompt treatment (in this case statins) and increase the risk of the outcome (in this case vascular dementia), thus causing a distorted positive association between the treatment and outcome (see Figure 5.13). In causal inference language, statins and dementia are said to be d-connected as there is an open “backdoor” path between them via the uncontrolled confounders.¹⁸¹

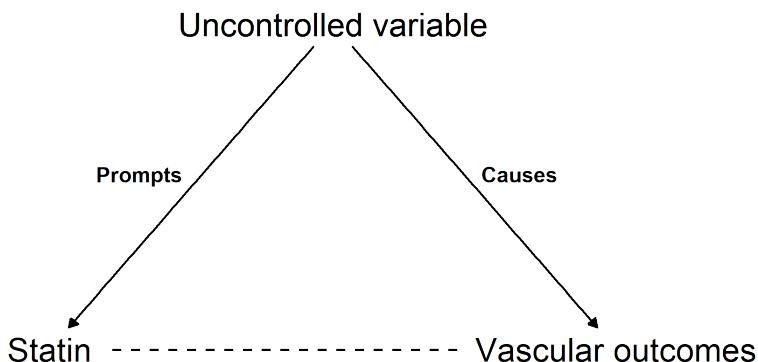


Figure 5.13: Causal diagram (directed acyclic graph) illustrating confounding by indication.

In the case of this analysis, the confounding variable (or, more likely, variables) would prompt prescription of statins (or another lipid regulating agent) but also represent a vascular risk factor that contributes to the development of the vascular dementia.

Supporting evidence for this interpretation comes from a variety of sources, including the results of the control outcome analyses. The slight harmful effect for the backpain outcome is substantially smaller than that observed for the ischemic heart disease outcome, indicating that the majority of the uncontrolled confounding is likely related to vascular factors. This is supported by the increasingly harmful

5.5 - Discussion

effect moving from Probable/Possible AD to other dementia to vascular dementia, indicating that the confounding by indication likely increases as the proportion of vascular outcomes increase. This is supported by the decision tree for assigning outcomes in the presence of greater than one dementia code, where the Alzheimer's disease outcomes require a "pure" condition, and the presence of any vascular or other dementia codes excludes participants from this group (Figure 5.1).

A review of other available literature suggests that this observation (a harmful effect of lipid regulating agents on vascular-related outcome) is not unusual. Using a conventional epidemiological technique, a previous analysis also found an increased risk of coronary heart disease (analogous to ischemic heart disease) in those taking statins (HR: 1.31, 95%CI: 1.04-1.66).¹⁸² Following control for confounding by indication through use of a trial emulation analysis, the expected protective effect of statins was observed.

Conditioning entry into the study on being either "at-risk" or already diagnosed with hypercholesterolemia was employed in a pre-emptive attempt to mitigate confounding by indication, but evidence from the control outcomes suggests this was unsuccessful. I also adjusted for several additional potential baseline variables. However, important confounding variables for which I have not adjusted could include genetic factors. A recent preprint of a study in the UK Biobank demonstrated that an Alzheimer's disease polygenic risk score was associated with an increased risk of vascular dementia, and also with an increased frequency of self-reported raised cholesterol levels, a diagnosis of hypercholesterolaemia, and a history of taking lipid-regulating agents such as statins or ezetimibe.¹⁸³

On a more general note, the unexpected increase in vascular dementia risk with statin use is particularly interesting given the absence of vascular dementia in the published literature, as highlighted in the previous chapter (see Section 4.4.7). It is possible that previous research identified a similar effect to this analysis, and via a publication bias mechanism, these results did not make it into the evidence base.

In summary, despite best efforts to account for strong confounding, participants prescribed a lipid regulating agent likely have a different vascular risk profile than those who are not, which may induce a spurious harmful association between lipid regulating agent use and vascular dementia outcomes.

5.5.3 Strengths and limitations

The primary strength of this analysis compared to others available in the literature is the relative size of the CPRD and length of follow-up. Having reviewed the other studies identified by the systematic review in Chapter 4, this analysis of 1,684,564 participants is one of the largest available studies of this research question. Additionally, this analysis followed LRA users and non-users from a common index date, using a time-updating treatment indicator to correctly assign time-at-risk to the exposed and unexposed groups. This approach has been less commonly used in the literature and allows for the mitigation of potential immortal time bias.

However, the findings of this analysis are subject to several limitations. There is a strong possibility of differential misclassification¹⁸⁴ of dementia-related conditions based on the exposure. As an illustrative example, those with memory complaints may be more likely to be classified as vascular dementia than Alzheimer's disease if their medical records contain prescriptions for lipid-regulating agents. Further, there is a potential for general non-differential misclassification of the outcome based on varying positive predictive value of electronic health record code lists to identify dementia cases.^{153,154}

Misclassification of outcomes is not the only issue introduced by the use of EHR codes to define outcomes. Comparing and contrasting between different studies is particularly difficult because of the impact that the use of different code list can have on the analysis. This is a particular challenge in comparing research across different time-periods and coding systems, as illustrated by the discrepancy

between the results when using the codes lists defined for this study and those used by Smeeth *et al.*

5.5.4 Enabling easy synthesis of this analysis

In light of my own experiences in attempting to extract information for papers assessing preventative treatments, as documented in Section 4.5.5, the outputs from this analysis are readily available.

All code, Read codelists and summary statistics (i.e. the tables presented in this chapter, plus summary tables of effect estimates) are readily available in machine readable formats (i.e. as comma separated values, or CSV, files) from the archived repository for this project (Zotero DOI: XXXX-XXXX).

The raw data supporting this analysis is not available, as access to the CPRD data is controlled by a data monitoring committee. In the context of access-controlled data, sharing the analysis code represents a way for readers to validate the findings.¹⁸⁵

Additionally, to help ensure accurate reporting of this analysis, the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines were used¹⁸⁶ (see Appendix ?? for the STROBE checklist).

This open approach will enable easy inclusion of this analysis in future evidence synthesis exercises, allowing new work to readily build on that presented here.

5.5.5 Conclusions

This chapter has provided new evidence on the potential repurposing of lipid-regulating agents for the prevention of all-cause dementia, Alzheimer's disease, vascular dementia, and other dementia. It made use of a large scale electronic health record database, the CPRD, and employed a time-varying Cox proportional hazards model to account for the potential immortal time bias.

5.5 - Discussion

It found no evidence for an effect of lipid-regulating agents on probable or possible Alzheimer's disease. However, lipid regulating agent use was associated with an increased risk of all-cause, vascular and other dementias. In all cases, the estimated associations were driven by those observed in the statin subgroup, which comprised the majority of exposed participants in this cohort.

This chapter attempted to account for important sources of bias and provide a comparison with other available literature, as identified in the systematic review presented in Chapter 4. However, there is a strong potential for uncontrolled confounding by indication and differential misclassification of the outcome on the basis of exposure, which raises questions about the findings, in particular the unexpected increase in risk of vascular dementia associated with statin use. This is supported by our findings for the negative and positive control outcomes used, which provide some evidence of uncontrolled vascular confounders that may both prompt LRA prescription and increase risk of vascular dementia. Future research using large scale electronic health records should aim to address these limitations, potentially by using a analytical design that more closely emulates a trial.¹⁸²

Regardless, this analysis has provided an additional source of evidence for the triangulation exercise presented in Chapter 7. In the following Chapter, the dataset described here is incorporated along with several other datasets as part of an IPD analysis to investigate the effect of blood lipid levels on dementia outcomes directly, rather than via the proxy of treatment.

5.6 References

— Hold — Hold

6

Individual participant data meta-analysis

Contents

6.1	Lay Summary	129
6.2	Introduction	129
6.3	Methods	130
6.4	Discussion	132

6.1 Lay Summary

6.2 Introduction

Is there a good way to visualise input of cohorts into an IPD analysis?

6.3 Methods

As part

As part of the IPD analysis, relevant cohort studies were identified from the style and approached. In the first instance, the corresponding author on the main man

Talk about how different covariates were managed across cohorts,

Talk about missing data was handled

6.3.1 Data sources

Several cohort studies were approach

As part of this Chapter, I will use individual patient data from a range of sources.

These sources are described here in detail for reference.

Should also include a list of reasons why specific additional cohorts were not included - some like the EHR might be too big to get

Each should have:

- Description (including observation period, numbers, numbers with outcome, etc)
- Whether they are a known genetically at-risk cohort
-

Epic Norfolk

The European Prospective Investigation of Cancer - Norfolk is a^{187,188}

BRACE

CaPS

ELSA

EPIC Norfolk

Generation Scotland

Memento

NICOLA

TRACK HD

Whitehall II

The Whitehall II study is a prospective cohort study of 10 308 participants (70% men), aged 35–55 years and recruited between 1985 and 1989 from 20 London-based Civil service departments (<https://www.ucl.ac.uk/whitehallII>). Clinical examinations have been performed in 1991-1994, 1997-1999, 2002-2004, 2007-2009, 2012-2013, 2015-2016 with the data from circulating metabolomic traits and cognitive testing for the present study obtained from the 1997-1999 clinic phase. - taken from EN ID: 2140

included in previous IPD meta-analysis (Endnote ID: 2140)

Different approaches to combining subgroups.¹⁸⁹

Several DPUK cohorts were excluded on the basis of a lack of exposure variables, including Cam-CAN (had cardiovascular category, but only contained blood pressure).

See

6.3.2 Risk of bias assessment

Risk of bias assessment was performed for each of the included cohorts using the relevant tool described in Section 4.3.7.

6.3.3 Data cleaning

6.4 Discussion

Letters sent to all cohorts identified through the systematic review - will need quite a bit here on the numbers approached and how they responded if at all. Will also need to comment on the additional tactics used.

This is likely due to my junior position as a early career research combined with the impact

Range of reasons why data is not readily available. Privacy concerns, concerns around scoping or “parasitic” behaviour, and a lack of trust (i.e. primary researchers do not trust secondary researchers).

Unfortunately,

As part of this, an analysis of

6.4.1 Reflections on data access

Describe your experience of trying to access the DPUK - while a great resources, frustrating at times. The portal would benefit from a centralised management system and

6.4.2 Study within a review

In future, it could have been worth running a “study within a review”, such as that used to examine whe

<https://bmjopen.bmj.com/content/9/1/e025273>

— Hold — Hold

7

Discussion

Contents

7.1	Triangulation across existing evidence and evidence produced in this thesis	135
7.2	Qualitative vs quantitative triangulation	136
7.3	Discussion of E-values	136
7.4	Use of consensus panel to bring everything together .	136
7.5	Additional topics for the discussion	136
7.6	Summary of findings (and implications for policy makers)	137
7.7	Strengths and Limitations	137
7.8	Reproducible research	137
7.9	Public involvement and engagement	138
7.10	Future work	138
7.11	Overall conclusions	138

7.1 Triangulation across existing evidence and evidence produced in this thesis

7.1.1 Overview of triangulation

Using the review of the literature as a guide to identify useful further studies that could be performed, I then performed .

Summary table of different studies and results

7.2 Qualitative vs quantitative triangulation

7.2.1 Researcher positionality statement

7.3 Discussion of E-values

How strong unmeasured confounding would have to be in order to explain the observed effect.

Heavily criticised in a range of papers.

7.4 Use of concensus panel to bring everything together

7.5 Additional topics for the discussion

There are several addition topics I would like to cover:

- The difference between the peer review of the software vs. the peer review of papers. A conversation until everything is fixed/clear reporting guidelines/checklists, with acceptance dependent on them being implemented /built-in version control, and cross linking discussion points to changes through the linking of commits and comments. Essentially, a more transparent and open

7.6 Summary of findings (and implications for policy makers)

7.7 Strengths and Limitations

There are several strengths and limitations to the work presented in this thesis. One particularly strength is the lengths gone to find all available published and unpublished evidence around the question, and to integrate this evidence in a coherent framework, taking into account the limitations of ach source and how these limitations may be used to provide

Need for large simple trials for common disease where small treatment effect can have large effect -¹⁹⁰

7.8 Reproducible research

Reproducible and science has been a key theme running through this thesis, as reflected by the development of an open source tool to help search medRxiv and bioRxiv preprint metadata. In line with this, an open source copy of the code used to produce this thesis is available on GitHub, as is the code used to perform the analysis contained within it.

Containerisation was used to ensure that the code is reproducible, in line iwht best practices

Commentary on the fact that the best you can do is replicate vs reproducible (due the closed nature of the data).

7.9 - Public involvement and engagement

One is the ability to recreate the results given the same data and code, the other is the ability to recreate the results given the same code but a different dataset. IN theory it is possible to gain access the dataset given the information presented in Chapter @ref{cprd-analysis-heading}. However, access is dependency on an ISAC application to the managing body of the CPRD.

7.9 Public involvement and engagement

Involving and engaging the public and patients has been a central theme to this thesis.

Public engagement activities included

Public involvement also steered the creation of the topic

P

7.10 Future work

7.11 Overall conclusions

Lasciate ogne speranza, voi ch'intrate. . .

8

Bibliography

1. Cerejeira, J., Lagarto, L. & Mukaetova-Ladinska, E. B. Behavioral and Psychological Symptoms of Dementia. *Frontiers in Neurology* **3**, (2012).
2. Kumar, C. S. & Kuriakose, J. R. End-of-life care issues in advanced dementia. *Mental Health in Family Medicine* **10**, 129–132 (2013).
3. Burns, A. & Iliffe, S. Dementia. *BMJ* **338**, b75 (2009).
4. Robinson, L., Tang, E. & Taylor, J.-P. Dementia: Timely diagnosis and early intervention. *BMJ* **350**, h3029 (2015).
5. Iadecola, C. The pathobiology of vascular dementia. *Neuron* **80**, (2013).
6. Venkat, P., Chopp, M. & Chen, J. Models and Mechanisms of Vascular Dementia. *Experimental neurology* **272**, 97–108 (2015).
7. Custodio, N. *et al.* Mixed dementia: A review of the evidence. *Dementia & Neuropsychologia* **11**, 364–370 (2017).
8. Sheehan, B. Assessment scales in dementia. *Therapeutic Advances in Neurological Disorders* **5**, 349–358 (2012).
9. Dubois, B. *et al.* Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDSADRDA criteria. *The Lancet Neurology* **6**, 734–746 (2007).
10. Prince, M. *et al.* Recent global trends in the prevalence and incidence of dementia, and survival with dementia. *Alzheimer's Research & Therapy* **8**, 23 (2016).
11. Flier, W. M. van der & Scheltens, P. Epidemiology and risk factors of dementia. *Journal of Neurology, Neurosurgery & Psychiatry* **76**, v2–v7 (2005).

12. Baker, C., Jarrett, T. & Powell, T. *Dementia: Policy, services and statistics overview*. (2019).
13. Wittenberg, R. *et al.* The costs of dementia in England. *International Journal of Geriatric Psychiatry* **34**, 1095–1103 (2019).
14. Cummings, J., Lee, G., Ritter, A., Sabbagh, M. & Zhong, K. Alzheimer's disease drug development pipeline: 2020. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* **6**, e12050 (2020).
15. Hampel, H. *et al.* The cholinergic system in the pathophysiology and treatment of Alzheimer's disease. *Brain* **141**, 1917–1933 (2018).
16. Pariente, A. *et al.* Prevalence of cholinesterase inhibitors in subjects with dementia in Europe. *Pharmacoepidemiology and Drug Safety* **17**, 655–660 (2008).
17. Marucci, G. *et al.* Efficacy of acetylcholinesterase inhibitors in Alzheimer's disease. *Neuropharmacology* **108** 352 (2020). doi:10.1016/j.neuropharm.2020.108352
18. Francis, P. T., Ramírez, M. J. & Lai, M. K. Neurochemical basis for symptomatic treatment of Alzheimer's disease. *Neuropharmacology* **59**, 221–229 (2010).
19. Feingold, K. R. Introduction to Lipids and Lipoproteins. in *Endotext* (eds. Feingold, K. R. et al.) (MDText.com, Inc., 2000).
20. Peters, R. *et al.* Combining modifiable risk factors and risk of dementia: A systematic review and meta-analysis. *BMJ Open* **9**, e022846 (2019).
21. Anstey, K. J., Ee, N., Eramudugolla, R., Jagger, C. & Peters, R. A Systematic Review of Meta-Analyses that Evaluate Risk Factors for Dementia to Evaluate the Quantity, Quality, and Global Representativeness of Evidence. *Journal of Alzheimer's Disease* **70**, S165–S186
22. Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K. & Brayne, C. Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. *The Lancet Neurology* **13**, 788–794 (2014).
23. Laufs, U., Parhofer, K. G., Ginsberg, H. N. & Hegele, R. A. Clinical review on triglycerides. *European Heart Journal* **41**, 99–109c (2020).
24. Zampelas, A. & Magriplis, E. New Insights into Cholesterol Functions: A Friend or an Enemy? *Nutrients* **11**, (2019).
25. Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, Without Use of the Preparative Ultracentrifuge. *Clinical Chemistry* **18**, 499–502 (1972).
26. Detection, N. C. E. P. (US). E. P. on. *Third report of the National Cholesterol Education Program (NCEP) Expert Panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III)*. (The Program, 2002).
27. Nelson, R. H. Hyperlipidemia as a Risk Factor for Cardiovascular Disease. *Primary care* **40**, 195–211 (2013).

28. Libby, P. *et al.* Atherosclerosis. *Nature Reviews Disease Primers* **5**, 1–18 (2019).
29. Collins, R. *et al.* Interpretation of the evidence for the efficacy and safety of statin therapy. *The Lancet* **388**, 2532–2561 (2016).
30. Collins, R. *et al.* Interpretation of the evidence for the efficacy and safety of statin therapy. *The Lancet* **388**, 2532–2561 (2016).
31. Law, M. R., Wald, N. J. & Rudnicka, A. R. Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: Systematic review and meta-analysis. *BMJ* **326**, 1423 (2003).
32. Schachter, M. Chemical, pharmacokinetic and pharmacodynamic properties of statins: An update. *Fundamental & Clinical Pharmacology* **19**, 117–125 (2005).
33. Sierra, S. *et al.* Statins as neuroprotectants: A comparative in vitro study of lipophilicity, blood-brain-barrier penetration, lowering of brain cholesterol, and decrease of neuron cell death. *Journal of Alzheimer's disease: JAD* **23**, 307–318 (2011).
34. Kosoglou, T. *et al.* Ezetimibe: A review of its metabolism, pharmacokinetics and drug interactions. *Clinical Pharmacokinetics* **44**, 467–494 (2005).
35. Genest, J. Combination of statin and ezetimibe for the treatment of dyslipidemias and the prevention of coronary artery disease. *The Canadian Journal of Cardiology* **22**, 863–867 (2006).
36. Chaudhary, R., Garg, J., Shah, N. & Sumner, A. PCSK9 inhibitors: A new era of lipid lowering therapy. *World Journal of Cardiology* **9**, 76–91 (2017).
37. McKenney, J. New perspectives on the use of niacin in the treatment of lipid disorders. *Archives of internal medicine* **164**, 697–705 (2004).
38. Skulas-Ray Ann C. *et al.* Omega-3 Fatty Acids for the Management of Hypertriglyceridemia: A Science Advisory From the American Heart Association. *Circulation* **140**, e673–e691 (2019).
39. Burns, M. P. *et al.* Co-localization of cholesterol, apolipoprotein E and fibrillar A β in amyloid plaques. *Molecular Brain Research* **110**, 119–125 (2003).
40. Mizuno, T. *et al.* Cholesterol-dependent Generation of a Seeding Amyloid β -Protein in Cell Culture *. *Journal of Biological Chemistry* **274**, 15110–15114 (1999).
41. Beecham, G. W. *et al.* Genome-Wide Association Meta-analysis of Neuropathologic Features of Alzheimer's Disease and Related Dementias. *PLoS Genetics* **10**, (2014).
42. Harold, D. *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease, and shows evidence for additional susceptibility genes. *Nature genetics* **41**, 1088–1093 (2009).
43. Meng, Y. *et al.* Association between SORL1 and Alzheimer disease in a genome-wide study. *Neuroreport* **18**, 1761–1764 (2007).

44. Kivipelto, M. *et al.* Apolipoprotein E E4 Allele, Elevated Midlife Total Cholesterol Level, and High Midlife Systolic Blood Pressure Are Independent Risk Factors for Late-Life Alzheimer Disease. *Annals of Internal Medicine* **137**, 149–155 (2002).
45. Kivipelto, M. *et al.* Obesity and vascular risk factors at midlife and the risk of dementia and Alzheimer disease. *Archives of Neurology* **62**, 1556–1560 (2005).
46. Schilling, S. *et al.* Differential associations of plasma lipids with incident dementia and dementia subtypes in the 3C Study: A longitudinal, population-based prospective cohort study. *PLoS medicine* **14**, e1002265 (2017).
47. Solomon, A., Kivipelto, M., Wolozin, B., Zhou, J. & Whitmer, R. A. Midlife Serum Cholesterol and Increased Risk of Alzheimer’s and Vascular Dementia Three Decades Later. *Dementia and Geriatric Cognitive Disorders* **28**, 75–80 (2009).
48. Whitmer, R. A., Sidney, S., Selby, J., Johnston, S. C. & Yaffe, K. Midlife cardiovascular risk factors and risk of dementia in late life. *Neurology* **64**, 277–281 (2005).
49. Mainous, A. G., Eschenbach, S. L., Wells, B. J., Everett, C. J. & Gill, J. M. Cholesterol, transferrin saturation, and the development of dementia and Alzheimer’s disease: Results from an 18-year population-based cohort. *Family Medicine* **37**, 36–42 (2005).
50. Mielke, M. M. *et al.* The 32-year relationship between cholesterol and dementia from midlife to late life. *Neurology* **75**, 1888–1895 (2010).
51. Mielke, M. M. *et al.* High total cholesterol levels in late life associated with a reduced risk of dementia. *Neurology* **64**, 1689–1695 (2005).
52. Moroney, J. T. Low-Density Lipoprotein Cholesterol and the Risk of Dementia With Stroke. *JAMA* **282**, 254 (1999).
53. Anstey, K. J., Ashby-Mitchell, K. & Peters, R. Updating the Evidence on the Association between Serum Cholesterol and Risk of Late-Life Dementia: Review and Meta-Analysis. *Journal of Alzheimer’s Disease* **56**, 215–228
54. Ritchie, C. W., Terrera, G. M. & Quinn, T. J. Dementia trials and dementia tribulations: Methodological and analytical challenges in dementia research. *Alzheimer’s Research & Therapy* **7**, 31 (2015).
55. Trompet, S. *et al.* Pravastatin and cognitive function in the elderly. Results of the PROSPER study. *Journal of Neurology* **257**, 85–90 (2010).
56. Davey Smith, G. & Hemani, G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics* **23**, R89–R98 (2014).
57. Greenland, S. An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology* **29**, 722–729 (2000).

58. Davies, N. M., Holmes, M. V. & Smith, G. D. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
59. Larsson, S. C. *et al.* Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* **359**, j5375 (2017).
60. Østergaard, S. D. *et al.* Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. *PLOS Medicine* **12**, e1001841 (2015).
61. Kim, J., Basak, J. M. & Holtzman, D. M. The Role of Apolipoprotein E in Alzheimer's Disease. *Neuron* **63**, 287–303 (2009).
62. Benn, M., Nordestgaard, B. G., Frikke-Schmidt, R. & Tybjærg-Hansen, A. Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. *BMJ* **357**, j1648 (2017).
63. Donnelly, C. A. *et al.* Four principles to make evidence synthesis more useful for policy. *Nature* **558**, 361–364 (2018).
64. Chandler, J., Higgins, J., Deeks, J., Davenport, C. & Clarke, M. J. Chapter 1: Introduction. in *Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021)* **5**, (Cochrane, 2020).
65. Conn, V. S., Valentine, J. C., Cooper, H. M. & Rantz, M. J. Grey literature in meta-analyses. *Nursing Research* **52**, 256–261 (2003 Jul-Aug).
66. McAuley, L., Pham, B., Tugwell, P. & Moher, D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet (London, England)* **356**, 1228–1231 (2000).
67. Hopewell, S., McDonald, S., Clarke, M. & Egger, M. Grey literature in meta-analyses of randomized trials of health care interventions. *The Cochrane Database of Systematic Reviews* MR000010 (2007). doi:10.1002/14651858.MR000010.pub3
68. Lefebvre, C. *et al.* Section 4.4.4 -Searching for and selecting studies: Controlled vocabulary and text words. in *Cochrane Handbook for Systematic Reviews of Interventions* (eds. Higgins, J. *et al.*) (Cochrane, 2021).
69. Committee on Publication Ethics (COPE). *Discussion Document: Preprints*. (2018).
70. Vale, R. D. & Hyman, A. A. Priority of discovery in the life sciences. *eLife* **5**, e16931 (2016).
71. Rosenthal, R. The file drawer problem and tolerance for null results. *Psychological Bulletin* **86**, 638–641 (1979).
72. Maslove, D. M. Medical PreprintsA Debate Worth Having. *JAMA* **319**, 443–444 (2018).
73. Schalkwyk, M. C. I. van, Hird, T. R., Maani, N., Petticrew, M. & Gilmore, A. B. The perils of preprints. *BMJ* **370**, m3111 (2020).

74. Mahood, Q., Eerd, D. V. & Irvin, E. Searching for grey literature for systematic reviews: Challenges and benefits. *Research Synthesis Methods* **5**, 221–234 (2014).
75. Klein, M., Broadwell, P., Farb, S. E. & Grappone, T. Comparing published scientific journal articles to their pre-print versions. *International Journal on Digital Libraries* **20**, 335–350 (2019).
76. Nicholson, D. N. *et al.* Linguistic Analysis of the bioRxiv Preprint Landscape. *bioRxiv* 2021.03.04.433874 (2021). doi:10.1101/2021.03.04.433874
77. Munafò, M. R. & Smith, G. D. Robust research needs many lines of evidence. *Nature* **553**, 399–401 (2018).
78. Page, M. J. *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
79. McGuinness, L. A., Higgins, J. P. T. & Sterne, J. A. C. Assessing the Credibility of Findings From Nonrandomized Studies of Interventions. *JAMA Cardiology* **3**, 905–906 (2018).
80. Riley, R. D., Lambert, P. C. & Abo-Zaid, G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ* **340**, c221 (2010).
81. Stewart, L. A. & Parmar, M. K. B. Meta-analysis of the literature or of individual patient data: Is there a difference? *The Lancet* **341**, 418–422 (1993).
82. Arain, F. A., Kuniyoshi, F. H., Abdalrhim, A. D. & Miller, V. M. Sex/gender medicine. The biological basis for personalized care in cardiovascular medicine. *Circulation Journal: Official Journal of the Japanese Circulation Society* **73**, 1774–1782 (2009).
83. Clayton, J. A. Applying the new SABV (sex as a biological variable) policy to research and clinical care. *Physiology & Behavior* **187**, 2–5 (2018).
84. McCartney, M., Treadwell, J., Maskrey, N. & Lehman, R. Making evidence based medicine work for individual patients. *BMJ* **353**, i2452 (2016).
85. Letenneur, L. *et al.* Are sex and educational level independent predictors of dementia and Alzheimer's disease? Incidence data from the PAQUID project. *Journal of Neurology, Neurosurgery & Psychiatry* **66**, 177–183 (1999).
86. Stewart, L. A. & Tierney, J. F. To IPD or not to IPD?: Advantages and Disadvantages of Systematic Reviews Using Individual Patient Data. *Evaluation & the Health Professions* **25**, 76–97 (2002).
87. Tugwell, P. & Knottnerus, J. A. Advantages of individual patient data analysis in systematic reviews. *Journal of Clinical Epidemiology* **63**, 233–234 (2010).
88. Ventresca, M. *et al.* Obtaining and managing data sets for individual participant data meta-analysis: Scoping review and practical guide. *BMC Medical Research Methodology* **20**, 113 (2020).
89. Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H. & Ioannidis, J. P. A. Public Availability of Published Research Data in High-Impact Journals. *PLoS ONE* **6**, e24357 (2011).

90. Vines, T. H. *et al.* The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* **24**, 94–97 (2014).
91. Wartenberg, D. & Thompson, W. D. Privacy Versus Public Health: The Impact of Current Confidentiality Rules. *American Journal of Public Health* **100**, 407–412 (2010).
92. Bauermeister, S. *et al.* The Dementias Platform UK (DPUK) Data Portal. *European Journal of Epidemiology* **35**, 601–611 (2020).
93. Rawlinson, C. & Bloom, T. New preprint server for medical research. *BMJ* **365**, (2019).
94. Sever, R. *et al.* *bioRxiv: The preprint server for biology.* (Scientific Communication and Education, 2019). doi:10.1101/833400
95. Bealy, C. R Programming Humour. *Stack Exchange* (2013).

96. Bramer, W. M., De Jonge, G. B., Rethlefsen, M. L., Mast, F. & Kleijnen, J. A systematic approach to searching: An efficient and complete method to develop literature searches. *Journal of the Medical Library Association* **106**, (2018).
97. Gusenbauer, M. & Haddaway, N. R. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods* **11**, 181–217 (2020).
98. Wateridge, J. IT projects: A basis for success. *International Journal of Project Management* **13**, 169–172 (1995).
99. Abdill, R. J. & Blekhman, R. Rxivist.org: Sorting biology preprints using social media and readership metrics. *PLOS Biology* **17**, e3000269 (2019).
100. Iwema, C. L., LaDue, J., Zack, A. & Chattopadhyay, A. Search.bioPreprint: A discovery tool for cutting edge, preprint biomedical research articles. *F1000Research* **5**, 1396 (2016).
101. Shaw, M. "Self-healing": Softening precision to avoid brittleness: Position paper for WOSS '02: Workshop on self-healing systems. *Proceedings of the first workshop on Self-healing systems* 111–114 (2002). doi:10.1145/582128.582152
102. Laprie, J. C. Dependability: Basic Concepts and Terminology. in *Dependability: Basic Concepts and Terminology: In English, French, German, Italian and Japanese* (ed. Laprie, J. C.) 3–245 (Springer, 1992).
103. R Core Team. *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing, 2019).
104. Bramer, W. M., de Jonge, G. B., Rethlefsen, M. L., Mast, F. & Kleijnen, J. A systematic approach to searching: An efficient and complete method to develop literature searches. *Journal of the Medical Library Association : JMLA* **106**, 531–541 (2018).

105. Kodvanj, I., Homolak, J., Virag, D. & Trkulja, V. Publishing of COVID-19 Preprints in Peer-reviewed Journals, Preprinting Trends, Public Discussion and Quality Issues. *bioRxiv* 2020.11.23.394577 (2020). doi:10.1101/2020.11.23.394577
106. Noone, C. *et al.* Investigating and evaluating evidence of the behavioural determinants of adherence to social distancing measures A protocol for a scoping review of COVID-19 research. *HRB Open Research* **3**, 46 (2020).
107. Grassly, N. C. *et al.* Comparison of molecular testing strategies for COVID-19 control: A mathematical modelling study. *The Lancet Infectious Diseases* **20**, 1381–1389 (2020).
108. Boettiger, C., Chamberlain, S., Hart, E. & Ram, K. Building Software, Building Community: Lessons from the rOpenSci Project. *Journal of Open Research Software* **3**, e8 (2015).
109. Issue #380 - medrxivr: Accessing and searching medRxiv preprint data in R · · rOpenSci/software-review · GitHub. *GitHub · rOpenSci/software-review*
110. Bong, S. M., McKay, J. L., Factor, S. A. & Ting, L. H. *Perception of whole-body motion during balance perturbations is impaired in Parkinson's disease and is associated with balance impairment.* (*Neurology*, 2019). doi:10.1101/19000265
111. Song, F. *et al.* Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment* **14**, (2010).
112. Vuorre, M. & Crump, M. J. C. Sharing and organizing research products as R packages. *Behavior Research Methods* (2020). doi:10.3758/s13428-020-01436-x
113. Cochrane, A. L. 1931-1971: A critical review, with particular reference to the medical profession. *Medicines for the year 2000* (1979).
114. McGuinness, L., Higgins, J., Ben-Shlomo, Yoav, Coulthard, L. & Smith, G. The relationship between blood lipid levels and dementia: Protocol for a systematic review and triangulation of all available evidence. *OSF* (2020). doi:10.17605/OSF.IO/VTW5Y
115. Bramer, W. M., Giustini, D., de Jonge, G. B., Holland, L. & Bekhuis, T. De-duplication of database search results for systematic reviews in EndNote. *Journal of the Medical Library Association : JMLA* **104**, 240–243 (2016).
116. Hupe, M. EndNote X9. *Journal of Electronic Resources in Medical Libraries* **16**, 117–119 (2019).
117. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews* **5**, 210 (2016).
118. Petersen, R. C. *et al.* Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology* **56**, 303–308 (1999).
119. Gwet, K. L. Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology* **61**, 29–48 (2008).
120. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **20**, 37–46 (1960).

121. Wongpakaran, N., Wongpakaran, T., Wedding, D. & Gwet, K. L. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: A study conducted with personality disorder samples. *BMC Medical Research Methodology* **13**, 61 (2013).
122. Brennan, P. & Silman, A. Statistical methods for assessing observer variability in clinical measures. *BMJ : British Medical Journal* **304**, 1491–1494 (1992).
123. McHugh, M. L. Interrater reliability: The kappa statistic. *Biochemia Medica* **22**, 276–282 (2012).
124. Reynders, R. M., Ladu, L. & Girolamo, N. D. Contacting of authors modified crucial outcomes of systematic reviews but was poorly reported, not systematic, and produced conflicting results. *Journal of Clinical Epidemiology* **115**, 64–76 (2019).
125. Sterne, J. A. *et al.* ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355**, i4919 (2016).
126. Campbell, D. T. Factors relevant to the validity of experiments in social settings. *Psychological Bulletin* **54**, 297–312 (1957).
127. Jüni, P., Altman, D. G. & Egger, M. Assessing the quality of controlled clinical trials. *BMJ* **323**, 42–46 (2001).
128. Sterne, J. A. C. *et al.* RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, (2019).
129. Morgan R *et al.* A tool to assess Risk Of Bias In Non-randomised Studies - of Exposures (ROBINS-E). in *Advances in Evidence Synthesis: Special issue, Cochrane Database of Systematic Reviews* **9 Supp 1**, 320 (2020).
130. Wells, G. A. *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. (2000).
131. Mamluk, L. *et al.* Evidence of detrimental effects of prenatal alcohol exposure on offspring birthweight and neurodevelopment from a systematic review of quasi-experimental studies. *International Journal of Epidemiology* **49**, 1972–1995 (2020).
132. McGuinness, L. A. & Higgins, J. P. T. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods* **12**, 55–61 (2021).
133. Issue #102 - Feature request: Paired forest/risk-of-bias plot. *GitHub* · [mcguinlu/robvis](https://github.com/mcguinlu/robvis)
134. Feinstein, A. R. & Cicchetti, D. V. High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology* **43**, 543–549 (1990).
135. McGuinness, B., Craig, D., Bullock, R. & Passmore, P. Statins for the prevention of dementia. *The Cochrane Database of Systematic Reviews* CD003160 (2016). doi:10.1002/14651858.CD003160.pub3

136. Chu, C.-S. *et al.* Use of statins and the risk of dementia and mild cognitive impairment: A systematic review and meta-analysis. *Scientific Reports* **8**, 5804 (2018).
137. Yang, Z. *et al.* Association of blood lipids, atherosclerosis and statin use with dementia and cognitive impairment after stroke: A systematic review and meta-analysis. *Ageing Research Reviews* **57**, 100962 (2020).
138. Muangpaisan, W., Brayne, C. & the Alzheimer's Society Vascular Dementia Systematic Review Group. Systematic review of statins for the prevention of vascular dementia or dementia. *Geriatrics & Gerontology International* (2010). doi:10.1111/j.1447-0594.2009.00579.x
139. Poly, T. N. *et al.* Association between Use of Statin and Risk of Dementia: A Meta-Analysis of Observational Studies. *Neuroepidemiology* **54**, 214–226 (2020).
140. Kuźma, E. *et al.* Which Risk Factors Causally Influence Dementia? A Systematic Review of Mendelian Randomization Studies. *Journal of Alzheimer's disease: JAD* **64**, 181–193 (2018).
141. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *bioRxiv* (2017). doi:10.1101/168674
142. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nature Communications* **9**, 224 (2018).
143. Sohoni, Z. N. *et al.* Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (Q-Genie) tool. *BMC Genetics* **16**, 50 (2015).
144. Larsson, S. C. *et al.* Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* **359**, j5375 (2017).
145. Waffenschmidt, S. *et al.* Development and validation of study filters for identifying controlled non-randomized studies in PubMed and Ovid MEDLINE. *Research Synthesis Methods* **11**, 617–626 (2020).
146. Wagner, M., Rosumeck, S., Küffmeier, C., Döring, K. & Euler, U. A validation study revealed differences in design and performance of MEDLINE search filters for qualitative research. *Journal of Clinical Epidemiology* **120**, 17–24 (2020).
147. Booth, A. *et al.* An international registry of systematic-review protocols. *The Lancet* **377**, 108–109 (2011).
148. Danko, K. J., Dahabreh, I. J., Ivers, N. M., Moher, D. & Grimshaw, J. M. Contacting authors by telephone increased response proportions compared with emailing: Results of a randomized study. *Journal of Clinical Epidemiology* **115**, 150–159 (2019).
149. Meursinge Reynders, R., Ladu, L. & Di Girolamo, N. Contacting of authors modified crucial outcomes of systematic reviews but was poorly reported, not systematic, and produced conflicting results. *Journal of Clinical Epidemiology* **115**, 64–76 (2019).

150. Cooper, C., Bou, J. T. & Varley-Campbell, J. Evaluating the effectiveness, efficiency, cost and value of contacting study authors in a systematic review: A case study and worked example. *BMC Medical Research Methodology* **19**, 45 (2019).
151. Hennessy, E. A. *et al.* Ensuring Prevention Science Research is Synthesis-Ready for Immediate and Lasting Scientific Impact. *Prevention Science* (2021). doi:10.1007/s11121-021-01279-8
152. Hsieh, C.-Y. *et al.* Taiwan's National Health Insurance Research Database: Past and future. *Clinical Epidemiology* **11**, 349–358 (2019).
153. McGuinness, L. A., Warren-Gash, C., Moorhouse, L. R. & Thomas, S. L. The validity of dementia diagnoses in routinely collected electronic health records in the United Kingdom: A systematic review. *Pharmacoepidemiology and Drug Safety* **28**, 244–255 (2019).
154. Wilkinson, T. *et al.* Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimer's & Dementia* **14**, 1038–1051 (2018).
155. Dorn, H. F. Philosophy of Inferences from Retrospective Studies. *American Journal of Public Health and the Nations Health* **43**, 677–683 (1953).
156. Walker, V. M., Davies, N. M., Jones, T., Kehoe, P. G. & Martin, R. M. Can commonly prescribed drugs be repurposed for the prevention or treatment of Alzheimers and other neurodegenerative diseases? Protocol for an observational cohort study in the UK Clinical Practice Research Datalink. *BMJ Open* **6**, (2016).
157. von Elm, E. *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Journal of Clinical Epidemiology* **61**, 344–349 (2008).
158. Herrett, E. *et al.* Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology* **44**, 827–836 (2015).
159. Williams, T., van Staa, T., Puri, S. & Eaton, S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Therapeutic Advances in Drug Safety* **3**, 89–99 (2012).
160. Wood, L. & Coulson, R. Revitalizing the general practice research database: Plans, challenges, and opportunities. *Pharmacoepidemiology and drug safety* **10**, 379–383 (2001).
161. Mathur, R. *et al.* Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *Journal of Public Health (Oxford, England)* **36**, 684–692 (2014).
162. Booth, N. What are the Read Codes? *Health Libraries Review* **11**, 177–182 (1994).
163. Wolf, A. *et al.* Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology* **48**, 1740–1740g (2019).

164. Wishart, D. S. *et al.* DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research* **46**, D1074–D1082 (2017).
165. Charlson, M. E., Pompei, P., Ales, K. L. & MacKenzie, C. R. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of chronic diseases* **40**, 373–383 (1987).
166. Wells, B. J., Chagin, K. M., Nowacki, A. S. & Kattan, M. W. Strategies for handling missing data in electronic health record derived data. *EgEMS* **1**, (2013).
167. Sterne, J. A. C. *et al.* Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ* **338**, b2393 (2009).
168. Moons, K. G., Donders, R. A., Stijnen, T. & Harrell Jr, F. E. Using the outcome for imputation of missing predictor values was preferred. *Journal of clinical epidemiology* **59**, 1092–1101 (2006).
169. Pigott, T. D. A review of methods for missing data. *Educational research and evaluation* **7**, 353–383 (2001).
170. Hughes, R. A., Heron, J., Sterne, J. A. C. & Tilling, K. Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. *International Journal of Epidemiology* **48**, 1294–1304 (2019).
171. Lévesque, L. E., Hanley, J. A., Kezouh, A. & Suissa, S. Problem of immortal time bias in cohort studies: Example using statins for preventing progression of diabetes. *BMJ* **340**, b5087 (2010).
172. Lamarca, R., Alonso, J., Gomez, G. & Munoz, A. Left-truncated Data With Age as Time Scale: An Alternative for Survival Analysis in the Elderly Population. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **53A**, M337–M343 (1998).
173. Gail, M. H., Graubard, B., Williamson, D. F. & Flegal, K. M. Comments on “Choice of time scale and its effect on significance of predictors in longitudinal studies” by Michael J. Pencina, Martin G. Larson and Ralph B. D’Agostino, *Statistics in Medicine* 2007; **26** :1343-1359. *Statistics in Medicine* **28**, 1315–1317 (2009).
174. Pencina, M. J., Larson, M. G. & D’Agostino, R. B. Choice of time scale and its effect on significance of predictors in longitudinal studies. *Statistics in Medicine* **26**, 1343–1359 (2007).
175. Lipsitch, M., Tchetgen, E. T. & Cohen, T. Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies. *Epidemiology (Cambridge, Mass.)* **21**, 383–388 (2010).
176. Selva-O’Callaghan, A. *et al.* Statin-induced myalgia and myositis: An update on pathogenesis and clinical recommendations. *Expert review of clinical immunology* **14**, 215–224 (2018).
177. Herrett, E. *et al.* Statin treatment and muscle symptoms: Series of randomised, placebo controlled n-of-1 trials. *BMJ* **372**, n135 (2021).

178. Karalis, D. G., Hill, A. N., Clifton, S. & Wild, R. A. The risks of statin use in pregnancy: A systematic review. *Journal of Clinical Lipidology* **10**, 1081–1090 (2016 Sep-Oct).
179. Smeeth, L., Douglas, I., Hall, A. J., Hubbard, R. & Evans, S. Effect of statins on a wide range of health outcomes: A cohort study validated by comparison with randomized trials. *British Journal of Clinical Pharmacology* **67**, 99–109 (2009).
180. Salas, M., Hotman, A. & Stricker, B. H. Confounding by Indication: An Example of Variation in the Use of Epidemiologic Terminology. *American Journal of Epidemiology* **149**, 981–983 (1999).
181. Suttorp, M. M., Siegerink, B., Jager, K. J., Zoccali, C. & Dekker, F. W. Graphical presentation of confounding in directed acyclic graphs. *Nephrology Dialysis Transplantation* **30**, 1418–1423 (2015).
182. Danaei, G., García Rodríguez, L. A., Cantero, O. F., Logan, R. & Hernán, M. A. Observational data for comparative effectiveness research: An emulation of randomised trials to estimate the effect of statins on primary prevention of coronary heart disease. *Statistical methods in medical research* **22**, 70–96 (2013).
183. Korologou-Linden, R. *et al.* The causes and consequences of Alzheimer's disease: Phenome-wide evidence from Mendelian randomization. *medRxiv* 2019.12.18.19013847 (2020). doi:10.1101/2019.12.18.19013847
184. Porta, M. *A dictionary of epidemiology*. (Oxford University Press, 2014).
185. Goldacre, B., Morton, C. E. & DeVito, N. J. Why researchers should share their analytic code. *BMJ* **367**, l6365 (2019).
186. Vandebroucke, J. P. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLOS Medicine* **4**, e297 (2007).
187. Riboli, E. The EPIC Project: Rationale and study design. European Prospective Investigation into Cancer and Nutrition. *International Journal of Epidemiology* **26**, 6S–14 (1997).
188. Riboli, E. *et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): Study populations and data collection. *Public Health Nutrition* **5**, 1113–1124 (2002).
189. Fisher, D. J., Carpenter, J. R., Morris, T. P., Freeman, S. C. & Tierney, J. F. Meta-analytical methods to identify who benefits most from treatments: Daft, deluded, or deft approach? *BMJ* **356**, j573 (2017).
190. Yusuf, S., Collins, R. & Peto, R. Why do we need some large, simple randomized trials? *Statistics in Medicine* **3**, 409–420 (1984).
191. Wilson, G. *et al.* Best Practices for Scientific Computing. *PLOS Biology* **12**, e1001745 (2014).

192. Wilson, G. *et al.* Good enough practices in scientific computing. *PLOS Computational Biology* **13**, e1005510 (2017).
193. Simillis, C. *et al.* Postoperative chemotherapy improves survival in patients with resected high-risk stage II colorectal cancer: Results of a systematic review and meta-analysis. *Colorectal Disease Published online 30th January*, (2020).
194. Tanneru, K. *et al.* Meta-analysis and systematic review of intermediate-term follow-up of prostatic urethral lift for benign prostatic hyperplasia. *International Urology and Nephrology* (2020).

Appendices

A

By Chapter

A.1 Chapter 2

A.1.1 Publications beyond the scope of this thesis

Peer reviewed

- PRISMA main paper
- PRISMA E&E
- PRISMA 2020 software
- COVID Suicide Living Review
- Data extraction tools systematic review

Under review/Preprints

- MSc Paper on systematic reviews of this thesis topic
-

A.1.2 Involvement of patients and the public

Patients were involved at several stages of this research. When designing the PhD programme of work, a Patient and Public Advisory Group (PPAG) provided feedback on the relevance of the question.

Additionally,

Lay summaries appear at the beginning of each chapter, reviewed by the Patient and Public Involvement panel. They provide a plain language summary

A.2 Chapter 4

A.2.1 Search strategy

Table A.1: Overview of the full Medline search strategy.

#	Search term	Hits
1	dement*.ti,ab.	103404
2	alzheimer*.ti,ab.	132832
3	exp Dementia/	154234
4	exp Alzheimer Disease/	87346
5	Pick* disease.ti,ab	2794
6	globular glial tauopathy.ti,ab	24
7	primary progressive aphasia.ti,ab	1051
8	logopaenic aphasia.ti,ab	0
9	posterior cortical atrophy.ti,ab	381
10	(age-associated) adj2 (memory decline).ti,ab	11
11	((mild or slight) adj2 (cognitive or cognition) adj2 (disorder* or defect* or deficit* or disabilit* or dysfunction or impair*).ti,ab	14883
12	((cognit\$ or memory or cerebr\$ or mental\$) adj3 (declin\$ or impair\$ or los\$ or deteriorat\$ or degenerat\$ or complain\$ or disturb\$ or disorder\$)).ti,ab	182141
13	(MCI or aMCI or CIND or non-aMCI).ti,ab	16893
14	(cognitive impair*).ti,ab	56411
15	Cognition Disorders/	62602
16	Cognitive Dysfunction/	11999
17	Mild Cognitive Impairment/	11999
18	<i>or/1-17</i>	407352

Table A.1: Overview of the full Medline search strategy. (*continued*)

#	Search term	Hits
19	lipid*.ti,ab.	462968
20	lipoprotein*.ti,ab.	140438
21	cholesterol.ti,ab.	227679
22	hypercholesterol*.ti,ab.	33093
23	hypcholesterol*.ti,ab.	3347
24	triacylglycerol.ti,ab.	11077
25	lipemia*.ti,ab.	1836
26	dyslipid?emia.ti,ab.	29128
27	hyperlipid?emia*.ti,ab.	25134
28	hypolipid?emia.ti,ab.	271
29	HDL.ti,ab.	61231
30	LDL.ti,ab.	71176
31	VLDL.ti,ab.	12485
32	triglyceride*.ti,ab.	104904
33	exp Dyslipidemias/	76480
34	exp Cholesterol/	155339
35	exp Lipoproteins/	141558
36	or/19-35	777210
37	statin*.ti,ab.	39998
38	atorvastatin.ti,ab.	7994
39	cerivastatin.ti,ab	646
40	fluvastatin.ti,ab.	1795
41	pravastatin.ti,ab.	3940
42	rosuvastatin.ti,ab.	3175
43	simvastatin.ti,ab.	8933
44	pitavastatin.ti,ab	816
45	lovastatin.ti,ab.	3667
46	fibrat*.ti,ab.	3135
47	("fibric acid" adj3 derivat*).ti,ab.	341
48	bezafibrate.ti,ab	1523
49	fenofibrate.ti,ab	3109
50	gemfibrozil.ti,ab	1802
51	clofenapate.ti,ab	39
52	clofibrate.ti,ab	3035
53	ciprofibrate.ti,ab	481
54	(bile adj3 sequest*).ti,ab.	816
55	colestyramine.ti,ab	60
56	colestipol hydrochloride.ti,ab	52
57	colesevelam hydrochloride.ti,ab	71
58	nicotinic acid*.ti,ab.	5854
59	inositol nicotinate.ti,ab	30
60	niacin.ti,ab	4631

Table A.1: Overview of the full Medline search strategy. (*continued*)

#	Search term	Hits
61	ezetimibe.ti,ab.	2766
62	acipimox.ti,ab	292
63	evolocumab.ti,ab	394
64	alirocumab.ti,ab	350
65	lomitapide.ti,ab	150
66	(omega-3-acid adj2 ethyl ester*).ti,ab	85
67	meglutol.ti,ab	2
68	Meglutol/	134
69	exp Anticholesteremic Agents/	71609
70	exp Fibric Acids/	9523
71	exp Ezetimibe/	1954
72	exp Nicotinic Acids/	36409
73	or/37-72	138108
74	18 and 36	19659
75	18 and 73	2287
76	74 or 75	21029
77	Animals/ not (Animals/ and Humans/)	4552498
78	76 not 77	18226
79	epidemiologic studies/ or case-control studies/ or cross-sectional studies/ or cohort studies/ or follow-up studies/ or longitudinal studies/ or prospective studies/ or retrospective studies/	2299133
80	((epidemiologic or prospective or retrospective or cross-sectional or case control* or cohort or longitudinal or followup or follow-up) adj3 (study or studies)).ti,ab,kf.	1043484
81	(case control* or cross-sectional or cohort? or follow-up or followup or longitudinal or prospective or retrospective or observational or population).ti.	656500
82	(cohort? adj2 (analys* or compar* or data or study or studies)).ab.	184866
83	(population adj2 (based or data* or study or studies or register? or survey? or surveillance)).ab.	200506
84	or/79-83	2933516
85	controlled clinical trial.pt.	93095
86	randomized controlled trial.pt.	483099
87	clinical trials as topic/	187183
88	(randomi#ed or randomi#ation or randomi#ing).ti,ab,kf.	585795

Table A.1: Overview of the full Medline search strategy. (*continued*)

#	Search term	Hits
89	(RCT or "at random" or (random* adj3 (administ* or allocat* or assign* or class* or cluster or crossover or cross-over or control* or determine* or divide* or division or distribut* or expose* or fashion or number* or place* or pragmatic or quasi or recruit* or split or substitut* or treat*)).ti,ab,kf.	512675
90	placebo.ab,ti,kf.	203773
91	trial.ti.	199586
92	(control* adj3 group*).ab.	498141
93	(control* and (trial or study or group*) and (waitlist* or wait* list* or ((treatment or care) adj2 usual))).ti,ab,kf.	19035
94	((single or double or triple or treble) adj2 (blind* or mask* or dummy)).ti,ab,kf.	165010
95	double-blind method/ or random allocation/ or single-blind method/	266392
96	or/85-95	1616814
97	84 or 96	4175140
98	MENDELIAN RANDOMIZATION ANALYSIS/	736
99	Mendelian randomi*.ti,ab,kf.	1647
100	98 or 99	1738
101	RANDOMIZED CONTROLLED TRIALS AS TOPIC/	124147
102	(RCT? or (randomi#ed adj2 (control* or intervention* or experiment* or trial* or study or studies))).ti,ab,kf.	405207
103	((random* or comparative or intervention? or treatment?) adj3 (efficacy or effect*)).ti,ab,kf.	435773
104	(clinical adj (intervention? or trial?)).ti,ab,kf.	346211
105	CLINICAL TRIALS AS TOPIC/ or CONTROLLED CLINICAL TRIALS AS TOPIC/	192430
106	TREATMENT EFFECT/	904484
107	or/101-106	1894420
108	100 AND 107	313
109	instrument* variab*.ti,ab,kf.	2380
110	((causal* or causative) adj3 (associat* or infer* or implicat* or effect* or predict* or factor? or risk? or relat*)).ti,ab,kf.	54710
111	((gene* adj2 (associat* or risk? or varia* or determinant?)) or risk variant?).ti,ab,kf.	234808
112	(disease* adj3 (expos* or associat* or etiolog* or pathogenesis or risk?)).ti,ab,kf.	304605
113	risk factor?.mp.	1045594
114	exp CAUSALITY/	782487
115	"confounding factors (epidemiology)"/	9873
116	(confound* or nonconfound* or non-confound*).ti,ab,kf.	113902

Table A.1: Overview of the full Medline search strategy. (*continued*)

#	Search term	Hits
117	(statistics or epidemiolog* or ((genetic* or molecular) and medicine)).jw.	205082
118	or/109 -117	1768577
119	108 and 118	273
120	98 and 101	27
121	119 or 120	277
122	97 or 121	4175143
123	78 and 122	6045

A.2.2 Web of Science Databases Searched

Table A.2: Summary of Web of Science databases searched.

Database	Abbreviation	Years
Science Citation Index Expanded	SCI-EXPANDED	1900-present
Social Sciences Citation Index	SSCI	1956-present
Arts & Humanities Citation Index	A&HCI	1975-present
Conference Proceedings Citation Index - Science	CPCI-S	1990-present
Conference Proceedings Citation Index - Social Science & Humanities	CPCI-SSH	1990-present
Emerging Sources Citation Index	ESCI	2015-present

A.2.3 Code to search preprints

```
library(medrxivr)

mx_data <- mx_api_content(to_date = "2019-09-01")
```

```

bx_data <- mx_api_content(server = "biorxiv",
                           to_date = "2019-09-01")

topic1 <- c(mx_caps("statin"),
            mx_caps("ldl"),
            mx_caps("hdl"),
            mx_caps("TG"),
            mx_caps("triglycer"),
            paste0("\b",mx_caps("TC"),"\b"),
            mx_caps("ezetim"),
            mx_caps("fibrate"),
            mx_caps("bile acid"),
            mx_caps("lipoprotein"),
            mx_caps("lipid"),
            mx_caps("cholesterol"))

topic2 <- c(mx_caps("dementia"),
            mx_caps("alzheim"),
            mx_caps("MCI"),
            mx_caps("mild cognitive"))

query <- list(
  topic1,
  topic2
)

bx_results <- mx_search(bx_data, query)

```

```
mx_results <- mx_search(mx_data, query)
```

Note

A.2.4 MR risk of bias tool

Table A.3: Tool used to assess risk of bias in Mendelian randomisation studies, adapted from that developed by Mamluk et al.¹³¹

Bias domain	Question	High	Moderate	Low
Weak instrument bias	Strength of association between instrument and exposure F statistic < 10 in the same sample (< 10 indicating a weak instrument)	F<10 F= missing or F~10		
Genetic confounding bias	Reported test on association between confounders and IV (testing the assumption that the instrument is associated with your outcome only via your exposure)	Yes AND there is an obvious association	Not presented or Yes presented AND there is some degree of association	Presented and no obvious association
'Other' Confounding bias	Included confounders in the IV analysis	Yes		No

Table A.3: Tool used to assess risk of bias in Mendelian randomisation studies, adapted from that developed by Mamluk et al.¹³¹ (*continued*)

Bias domain	Question	High	Moderate	Low
Additional direct effects between IV and outcome (exclusion restriction assumption)	Presence of pleiotropy for genetic IVs	Genetic IVs with no knowledge of mechanism for G-lipid association (e.g. GWAS hit, could be acting through any pathway...)	Biologically plausible lipid-specific mechanism of association for G-lipid (e.g. lipid metabolising genetic variants)	Same as moderate AND checks that there is no other known effect of genetic variants on outcome or its risk-factors
Bias due to selection of participants	Homogenous population or similar ancestry If no: Stratified by ethnicity or adjusted for population stratification (yes/no)	Non-homogenous population (e.g. black and white together, etc.)	Population described as homogenous (e.g. whites only) BUT not corrected for ancestry informative markers like principal components derived from GWAS	Population described as homogenous (e.g. whites only) AND corrected for ancestry informative markers like principal components derived from GWAS

A.3 Chapter 5

A.3.1 Amendments to protocol

A.3.2 Code lists

A.4 Chapter 6

A.4.1 Section 1

A.4.2 Section 2

A.5 Chapter 7

A.5.1 Catalogue of failures

Short section detailing the things I tried to do but which did not work.

B

Other Appendix

B.1 Software used to create this thesis

This thesis was written in RMarkdown. Several R packages were used as part of this project. `R-base?`, `R-bookdown?`, `R-dagitty?`, `R-data.table?`, `R-dplyr?`, `R-flextable?`, `R-ggdag?`, `R-ggplot2?`, `R-kableExtra?`,

All projects in these thesis attempt to conform to minimal best practices for research computing.^{191,192}

B.2 Producing risk-of-bias visualisations with `robvis`

B.2.1 Introduction

Risk of bias assessment - evaluation of the internal validity of studies included in a systematic review - often forms a key part of the evidence synthesis process, particularly in the health sciences.^{cochranechpt7?} A well-developed family of tools is widely used, which have in common the characteristic that they evaluate specific domains of bias rather being constructed as a checklist or a quantitative score.^{cochranechpt7?} These tools include the RoB 2 tool for randomized trials,^{sterne2019rob?} the ROBINS-I tool for non-randomized studies of interventions,¹²⁵ the QUADAS 2 tool for test accuracy and the ROBIS tool for systematic reviews.^{whiting2011quadas?} Within each

*B.2 - Producing risk-of-bias visualisations with **robvis***

bias domains a judgement is reached about the strength of the study in that regard: for example, the first domain in the Cochrane RoB 2 tool deals with bias arising from the randomization process.^{sterne2019rob?} Accessible graphics summarizing the results of these domain-based risk-of-bias assessments are included in reports of systematic reviews. A convenient plot in many reviews is a “traffic light” plot, which tabulates the judgement for each study in each domain. For larger numbers of studies, when such a table become unmanageable, a popular alternative is a weighted bar plot, which show the proportion of information with each judgement for each domain.^{higgins2008assessing?}

Researchers can face a number of barriers in creating these plots. While some evidence synthesis platforms, such as Cochrane’s Review Manager,^{cochrane2014review?} are able to produce these visualizations, not all researchers use these systems to conduct their systematic reviews, and copying the risk-of-bias data into these systems simply to produce the plots is inefficient and error prone. Likewise, creating the figures by hand, through software such as MS PowerPoint or Adobe Illustrator, may lead to unintentional errors and require the plots to be redrawn during an update to the review. Additionally, while the field of evidence synthesis software has grown rapidly in recent years,^{marshall2015systematic?} this growth has not been equally distributed across the different aspects of the systematic review process. For example, a recent review found several software offerings aimed specifically at the abstract screening stage of the review process,^{harrison2020software?} but no similar time- and error-reducing tool has been proposed for visualizing the results of risk-of-bias assessments.

Fortunately, tools such as R, RStudio and Shiny (an R package for building interactive web apps) have made it easier than ever to produce such a tool.^{rref?,rstudioref?,shinyref?} Here, we present **robvis** (Risk Of Bias VISualiation),^{mccguinness2019a?} an R package and Shiny web-app that allows users to create publication-ready risk-of-bias plots quickly and easily. Originally created for use with the major risk-of-bias assessment tools used in health research, the tool allows users to visualize the results from any domain-based risk-of-bias assessment or quality appraisal tool.

The tool is open-source and available to use free of charge. Users can download a stable version of the R package from CRAN (<https://cran.r-project.org/package=robvis>); or access and contribute to the development version via GitHub (<https://github.com/mcguinlu/robvis>).

B.2.2 Development

Development of **robvis** began in April 2019 at the Evidence Synthesis Hackathon (ESH), an event which brings together interested researchers, practitioners and coders to discuss and develop new open-source evidence synthesis technologies. Test versions of both the R package and the web app were made available in early June 2019, with attendees of the ESH and members of the Bristol Appraisal and Review of Research (BARR) group at the University of Bristol being invited to test the tool and provide feedback. This feedback, along with other feature suggestions from the wider evidence synthesis community captured via GitHub issues, was incorporated and the first release version of the package was uploaded to CRAN in November 2019. The tool has been well received and is beginning to be cited in the evidence synthesis literature.^{193,194,gibb2019consistent?,habadi2019prevalence?,veloso2020effectiveness?}

B.2.3 Installation

A stable version of **robvis** is hosted on the Comprehensive R Archive Network (CRAN) and can be installed using:

```
install.packages("robvis")
```

As development of **robvis** is ongoing, new features are often available in the development version some time before they appear in the stable CRAN version. The most recent development version can be install from GitHub using:

```
devtools::install_github("mcguinlu/robvis")
```

B.2 - Producing risk-of-bias visualisations with *robvis*

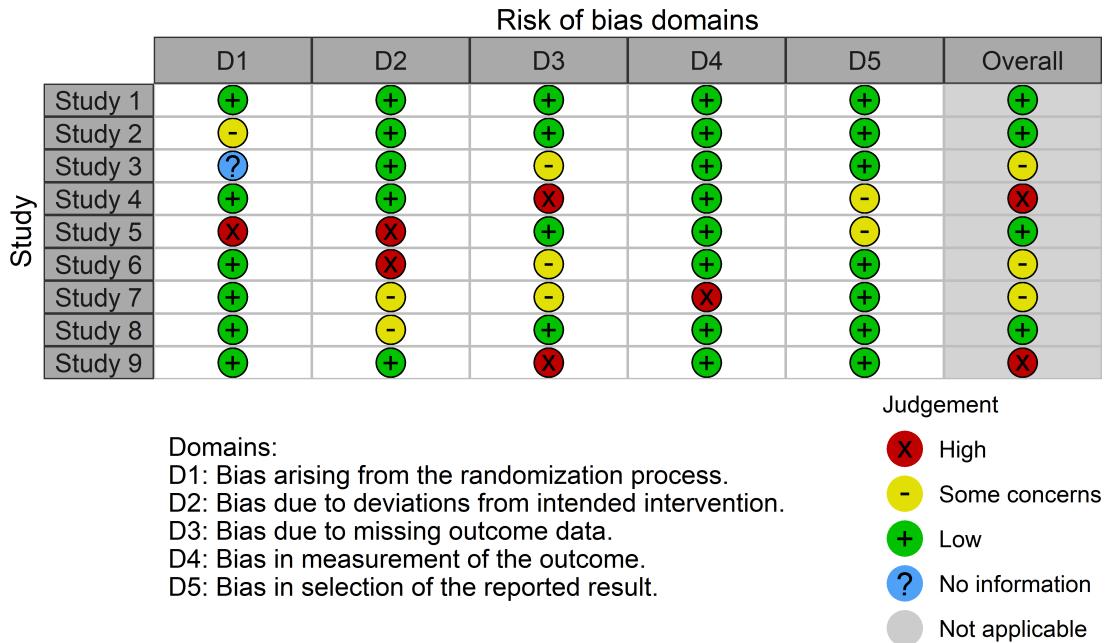


Figure B.1: Example risk of bias traffic light plot created using ‘robvis’

B.2.4 Usage

`robvis` contains two main functions. The first, `rob_traffic_light()`, creates a traffic light plot by tabulating each study by each domain, providing a more detailed view of the results of the risk-of-bias assessment. The second, `rob_summary()`, creates a weighted bar plot showing the proportion of information with each judgement for each domain in the assessment tool specified.

A worked example using these functions is outlined below, showing the ease with which risk-of-bias plots can be created using `robvis`. A detailed description of the additional options that can be used with each function is presented in Table B.1

Using the example data set (`data_rob2`) which is built into the package and is presented in Table ?? for reference, the traffic light plot shown in Figure B.1 is created using:

```
rob_traffic_light(data = data_rob2,
                  tool = "ROB2",
                  pszie = 15)
```

B.2 - Producing risk-of-bias visualisations with *robvis*

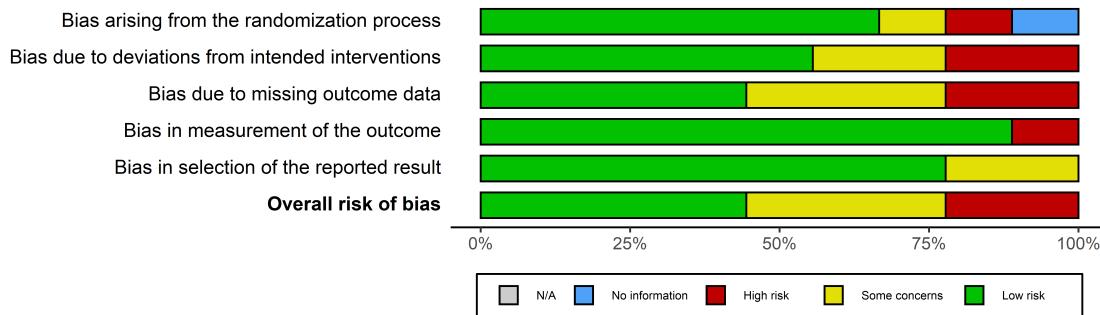


Figure B.2: Example risk of bias summary plot created using ‘robvis’ and the example ROB2 dataset

Similary, using the same data set, the summary barplot shown in Figure B.2 is created using:

```
rob_summary(data = data_rob2,
            tool = "ROB2",
            overall = TRUE)
```

A list of arguments available to the two functions in robvis are shown in Table B.1

Table B.1: Description of the arguments available in the two main ‘robvis’ functions. ‘X’ indicates that the option is available for the respective function.

Argument	rob_traffic_light()	rob_summary()	Description
data	X	X	Defines the dataframe containing the summary (domain) level risk-of-bias assessments. See the text and Table 1 for the format expected by ‘robvis’
tool	X	X	Defines the risk of bias assessment tool used. The RoB2 ('tool="ROB2"'), ROBINS-I ('tool="ROBINS-I"'), and QUADAS-2 ('tool="QUADAS-2"') assessments tools are currently supported. Other tools can be visualised using the generic template ('tool = "Generic"')
colour	X	X	Defines the colour scheme for the plot. The default is ‘colour = "cochrane” which uses the “Cochrane” (red, yellow, green) colours, while a preset option for a colour-blind friendly palette is also available ('colour = "colourblind"'). Alternatively, users can specify their own colour scheme e.g. ‘colour = c("#f442c8", "#bef441", "#000000")’

B.2 - Producing risk-of-bias visualisations with `robvis`

Table B.1: Description of the arguments available in the two main ‘`robvis`‘ functions. ‘X’ indicates that the option is available for the respective function. (*continued*)

Argument	<code>rob_traffic_light()</code>	<code>rob_summary()</code>	Description
overall		X	Defines whether to include an additional bar showing the distribution of overall risk of bias judgements in the summary barplot figure. Default is ‘overall = FALSE’.
weighted		X	Defines whether weights should be used to produce the summary barplot figure. Default is ‘weighted = TRUE’, in line with current Cochrane Collaboration guidance.
psize		X	Defines the size of the points in the traffic light plot. Default is ‘psize = 20’.

B.2.5 Reception and Future Plans

As of December 2021, `robvis` has been downloaded more than 13600 times. It has been well received by the systematic review community, and has been cited frequently in the published literature. A paper describing the tool was published in a special issue of Research Synthesis Methods focusing on data visualisation methods. A chapter on the tool has been incorporated into the “Doing Meta-Analysis in R” online textbook.[mathias_harrer_2019_2551803?](#)

While `robvis` is a stable package, a range of additional functionality could be added. At present, the number of tools with a specific template included in `robvis` is limited - adding additional templates is a priority. For example, a template for ROBIS, a tool for assessing risk of bias in systematic reviews, is in development.[whiting2016robis?](#) Additionally, the tool does not yet allow for the production of paired forest plots, where the risk-of-bias judgement is presented alongside each specific result included in the meta-analysis.[cochranechpt7?](#) This was initially considered to be beyond the scope of the tool, as it involves the visualization of something other than risk-of-bias assessments. However, following user-driven demand, this functionality is in development and will be available in the near future. Finally, we would like to add similar functionality to that provided by the `metafor::reporter()` function,

which generates a brief paragraph of text describing the results of a meta-analysis. The future `robvis::reporter()` function would provide a boilerplate description of the assessment tool used and the key domains at risk of bias.

B.3 Creating

B.4 Copies of papers arising from this thesis

medrxivr: Accessing and searching medRxiv and bioRxiv preprint data in R

Luke A McGuinness^{1, 2} and Lena Schmidt¹

¹ Department of Population Health Science, University of Bristol ² MRC Intergrative Epidemiology Unit, University of Bristol

DOI: [10.21105/joss.02651](https://doi.org/10.21105/joss.02651)

Software

- [Review ↗](#)
- [Repository ↗](#)
- [Archive ↗](#)

Editor: Daniel S. Katz ↗

Reviewers:

- [@danielskatz](#)

Submitted: 08 September 2020

Published: 09 October 2020

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

An increasingly important source of health-related bibliographic content are preprints: preliminary versions of research articles that have yet to undergo peer review. The two preprint repositories most relevant to health-related sciences are [medRxiv](#) and [bioRxiv](#), both of which are operated by the Cold Spring Harbor Laboratory, a not-for-profit research and educational institution (Rawlinson & Bloom, 2019).

The goal of the `medrxivr` R package is two-fold. In the first instance, it provides programmatic access to the Cold Spring Harbour Laboratory (CSHL) API, allowing users to download medRxiv and bioRxiv preprint metadata (e.g., title, abstract, author list.) This functionality will be of interest to anyone who wishes to import medRxiv and/or bioRxiv preprint metadata into R, for example to explore the distribution of preprints by subject area or by publication year. Examples of this type of usage have already been reported (e.g., by Brierley, 2020).

In the second instance, the package provides functions that allow users to search the downloaded preprint metadata for relevant preprints using complex search strings, including functionality such as search term truncation, Boolean operators (AND, OR, NOT), and term proximity. Helper functions are provided that allow users to export the results of their search to a .bib file for import into a reference manager (e.g., Zotero) and to download the full-text PDFs of preprints matching their search. This aspect of the package will be more relevant to systematic reviewers, health librarians and others performing literature searches, allowing them to perform and document transparent and reproducible searches in these important evidence sources.

Acknowledgements

We acknowledge funding from NIHR (LAM through NIHR Doctoral Research Fellowship (DRF-2018-11-ST2-048), and LS through NIHR Systematic Reviews Fellowship (RM-SR-2017-09-028)). LAM is a member of the MRC Integrative Epidemiology Unit at the University of Bristol. The views expressed in this article are those of the authors and do not necessarily represent those of the NHS, the NIHR, MRC, or the Department of Health and Social Care.

References

- Brierley, L. (2020). The role of research preprints in the academic response to the COVID-19 epidemic. *Authorea*. doi:[10.22541/au.158516578.89167184](https://doi.org/10.22541/au.158516578.89167184)
- Rawlinson, C., & Bloom, T. (2019). New preprint server for medical research. *BMJ*, 365. doi:[10.1136/bmj.l2301](https://doi.org/10.1136/bmj.l2301)

Risk-of-bias VISualization (**robvis**): An R package and Shiny web app for visualizing risk-of-bias assessments

Luke A. McGuinness^{1,2}  | Julian P. T. Higgins^{1,2}

¹MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

²Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Correspondence

Luke A. McGuinness, Bristol Medical School, University of Bristol, Canyng Hall, 39 Whatley Road, Bristol, BS8 2PS, United Kingdom.

Email: luke.mcguinness@bristol.ac.uk

Funding information

National Institute for Health Research, Grant/Award Number: NF-SI-0617-10145Research Trainees Coordinating Centre, Grant/Award Number: DRF-2018-11-ST2-048

Despite a major increase in the range and number of software offerings now available to help researchers produce evidence syntheses, there is currently no generic tool for producing figures to display and explore the risk-of-bias assessments that routinely take place as part of systematic review. However, tools such as the R programming environment and **Shiny** (an R package for building interactive web apps) have made it straightforward to produce new tools to help in producing evidence syntheses. We present a new tool, **robvis** (Risk-Of-Bias VISualization), available as an R package and web app, which facilitates rapid production of publication-quality risk-of-bias assessment figures. We present a timeline of the tool's development and its key functionality.

KEY WORDS

data visualization, evidence synthesis, R, risk of bias

1 | INTRODUCTION

Synthesis of evidence from the totality of relevant research is becoming more important than ever in informing policy across an increasingly wide range of fields.¹ Risk of bias assessment—evaluation of the internal validity of studies included in a systematic review—often forms a key part of the evidence synthesis process, particularly in the health sciences.² A well-developed family of tools is widely used, which have in common the characteristic that they evaluate specific domains of bias rather being constructed as a checklist or a quantitative score.² These tools include the RoB (Risk of Bias) 2 tool for randomized trials,³ the ROBINS-I (Risk Of Bias In Non-randomized Studies—of Interventions) tool,⁴ the QUADAS 2 (Quality and Applicability of Diagnostic Accuracy Studies) tool,⁵ and the ROBIS (Risk Of Bias in Systematic Reviews) tool.⁶ Within each bias domain a judgement is reached about the strength of the study in

that regard: for example, the first domain in the Cochrane RoB 2 tool deals with bias arising from the randomization process.³ Accessible graphics summarizing the results of these domain-based risk-of-bias assessments are included in reports of systematic reviews. A convenient plot in many reviews is a “traffic light” plot, which tabulates the judgement for each study in each domain. For larger numbers of studies, when such a table becomes unmanageable, a popular alternative is a (possibly weighted) bar plot, which shows the proportion of information with each judgement for each domain.⁷

Researchers can face a number of barriers in creating these plots. While some evidence synthesis platforms, such as Cochrane's Review Manager,⁸ can produce these visualizations, not all researchers use these systems to conduct their systematic reviews, and copying the risk-of-bias data into these systems solely to produce the plots is inefficient and error prone. On the other hand, creating the figures “by hand,” through software such as MS

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Research Synthesis Methods* published by John Wiley & Sons Ltd.

PowerPoint or Adobe Illustrator, may lead to unintentional errors and require the plots to be redrawn when a review is updated. Additionally, while the field of evidence synthesis software has grown rapidly in recent years,⁹ this growth has not been equally distributed across the different aspects of the systematic review process. For example, a recent review found several software offerings aimed specifically at the abstract screening stage of the review process,¹⁰ but no similar time- and error-reducing tool has been proposed for visualizing the results of risk-of-bias assessments.

It is now straightforward to produce such a tool, thanks to the availability of powerful computing offerings including R, RStudio, and **Shiny** (an R package for building interactive web apps).¹¹⁻¹³ Here, we present **robvis** (Risk Of Bias VISeualization),¹⁴ an R package and **Shiny** web-app that allows users to create publication-ready risk-of-bias plots quickly and easily. While primarily designed for use with the major risk-of-bias assessment tools used in health research (ROB2, ROBINS-I, and QUADAS-2), the tool allows users to visualize the results from any domain-based risk-of-bias assessment or quality appraisal tool.

The tool is open-source and available to use free of charge. Users can download a stable version of the R package from CRAN (<https://cran.r-project.org/package=robvis>); or access and contribute to the open-source code that powers the package via GitHub (<https://github.com/mcguinlu/robvis>). Extended guidance for the tool, including a step-by-step walk-through for those new to the R programming environment, is also available via the “Doing Meta-Analysis in R” online guide.¹⁵ Below, we discuss the tool’s development and key functionality.

2 | DEVELOPMENT

Development of **robvis** began in April 2019 at the Evidence Synthesis Hackathon (ESH), an event which brings together interested researchers, practitioners, and coders to discuss and develop new open-source evidence synthesis technologies. Test versions of both the R package and the web app were made available in early June 2019, with attendees of the ESH and members of the Bristol Appraisal and Review of Research (BARR) group at the University of Bristol being invited to test the tool and provide feedback. This feedback, along with other feature suggestions from the wider evidence synthesis community captured via GitHub issues, was incorporated and the first release version of the package was uploaded to CRAN in November 2019. The tool has been well received and is beginning to be cited in the evidence synthesis literature.¹⁶⁻²⁰

Highlights

- Risk-of-bias assessment is a key element of the systematic review workflow.
- No other tool exists explicitly for the purpose of visualizing risk-of-bias results.
- Here, we present **robvis**, an open-source R package and Shiny web app for creating publication-ready risk-of-bias assessment figures.
- **robvis** forms part of the metaverse, a collection of R packages designed to provide an evidence synthesis workflow in R.

3 | FUNCTIONALITY

3.1 | R package

3.1.1 | Tool templates and example data sets

At the time of writing, the tool includes templates for three major tools: the Cochrane RoB 2 tool for assessing randomized trials,³ the ROBINS-I tool for assessing non-randomized studies of interventions,⁴ and the QUADAS-2 tool for assessing diagnostic accuracy studies.⁵ These templates automatically apply the correct risk-of-bias domain names to the figures, and label the judgement levels appropriately (eg, “Low,” “Some Concerns,” and “High” in the case of the RoB 2 tool). In addition, **robvis** contains a general template that can be used to visualize the result of any domain-based assessment tool. The generic template has greater flexibility than the tool-specific templates, allowing a user-specified number of domains and custom domain titles to be used. It is suitable for use with the original version of the Cochrane risk-of-bias tool for randomized trials,²¹ in which flexibility in the specification of domains was permitted. Since users of more recent tools such as RoB 2 and ROBINS-I are not permitted to modify the domains, we strongly encourage use of the in-built templates for these.

In order to help users familiarize themselves with the package and its functionality, **robvis** contains built-in example data-sets for each template. We illustrate the example data for the RoB 2 tool for assessing risk of bias in randomized trials in Table 1.

3.1.2 | Data import

robvis expects the risk-of-bias data file to be arranged in a specific way (see Table 1 for an example). The first

TABLE 1 Example dataset for the ROB 2 tool contained within **robvis**. Data can be imported to the tool from with an Excel spreadsheet or a CSV file

Study	D1	D2	D3	D4	D5	Overall	Weight
Study 1	Low	Low	Low	Low	Low	Low	33.33
Study 2	Some concerns	Low	Low	Low	Low	Low	33.33
Study 3	Some concerns	Low	Some concerns	Low	Low	Some concerns	0.14
Study 4	Low	Low	High	Low	Some concerns	High	9.09
Study 5	High	High	Low	Low	Some concerns	Low	12.5
Study 6	Low	High	Some concerns	Low	Low	Some concerns	25
Study 7	Low	Some concerns	Some concerns	High	Low	Some concerns	200
Study 8	Low	Some concerns	Low	Low	Low	Low	11.11
Study 9	Low	Low	High	Low	Low	High	1.11

TABLE 2 Description of the arguments available in the two main functions of the **robvis** R package. “X” indicates that the option is available for the respective function

Argument	rob_traffic_light()	rob_summary()	Description
data	X	X	Defines the data frame containing the summary (domain) level risk-of-bias assessments. See the text and Table 1 for the format expected by robvis .
tool	X	X	Defines the risk of bias assessment tool used. The RoB2 (<code>tool = "ROB2"</code>), ROBINS-I (<code>tool = "ROBINS-I"</code>), and QUADAS-2 (<code>tool = "QUADAS-2"</code>) assessments tools are currently supported. Other tools can be visualized using the generic template (<code>tool = "Generic"</code>) [Note 1].
color	X	X	Defines the color scheme for the plot. The default is <code>colour = "cochrane"</code> which uses the “Cochrane” (red, yellow, green) colors, while a preset option for a color-blind friendly palette is also available (<code>colour = "colourblind"</code>). Alternatively, users can specify their own color scheme for example, <code>colour = c ("#f442c8," "#bef441," "#000000")</code> .
overall		X	Defines whether to include an additional bar showing the distribution of overall risk of bias judgments in the summary barplot figure. Default is <code>overall = FALSE</code> .
weighted		X	Defines whether weights should be used to produce the summary barplot figure. Default is <code>weighted = TRUE</code> , in line with current Cochrane Collaboration guidance.
psize	X		Defines the size of the points in the traffic light plot. Default is <code>psize = 20</code> .

Note: This option (`tool = "Generic"`) reflects the general template name used in the current development version of **robvis**, which will become the standard for all future iterations of the package. However, in the current CRAN version, the generic template is accessed using `tool = "ROB1"`.

column should contain the unique study/result identifier. The second-to-last column should contain the overall risk-of-bias judgments. The final column contains some measure of the result's precision (eg, the weight assigned to that result in a meta-analysis, or the sample size of the analysis that produced the result). This weight column is

used to create the summary bar plot, as current guidance recommends dividing the bars to show the proportion of information at each level of risk of bias, determined by the cumulative weight at that level, rather than simply showing the number of studies/results in each category.² If a measure of precision is not available, or to reproduce

“equally” weighted bar charts as have traditionally been presented in Cochrane Reviews to date, these weights may all be specified to be 1.

3.1.3 | Functions

robvis contains two main functions. The first, **rob_traffic_light()**, creates a *traffic light plot*. This displays every risk-of-bias judgement in a matrix, with domains along the horizontal and results/studies down the vertical, similar to the data set. The second function, **rob_summary()**, creates a *weighted bar plot*. This shows the proportion of information with each

risk-of-bias judgement separately for each domain in the assessment tool specified.

A worked example using these functions is outlined below, illustrating the simple steps involved in creating risk-of-bias plots using **robvis**. A detailed description of the additional options that can be used with each function is presented in Table 2. All examples produced in this article are created using the stable version available from CRAN.

To install and load the package in R, enter the following into the console:

```
install.packages("robvis")
library(robvis)
```

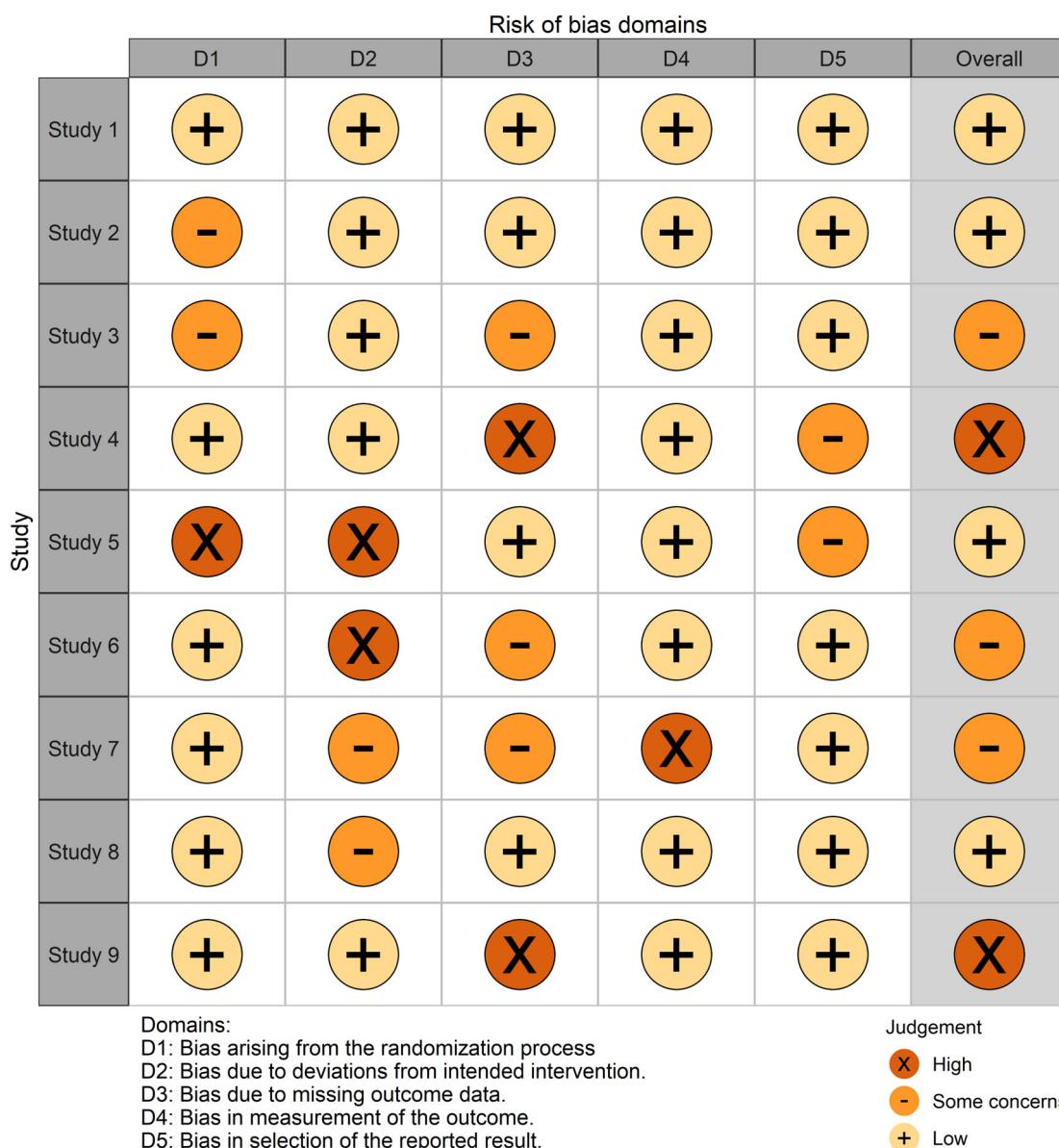


FIGURE 1 Example risk of bias traffic light plot of ROB2 assessments created using **robvis** and the colourblind palette [Colour figure can be viewed at wileyonlinelibrary.com]

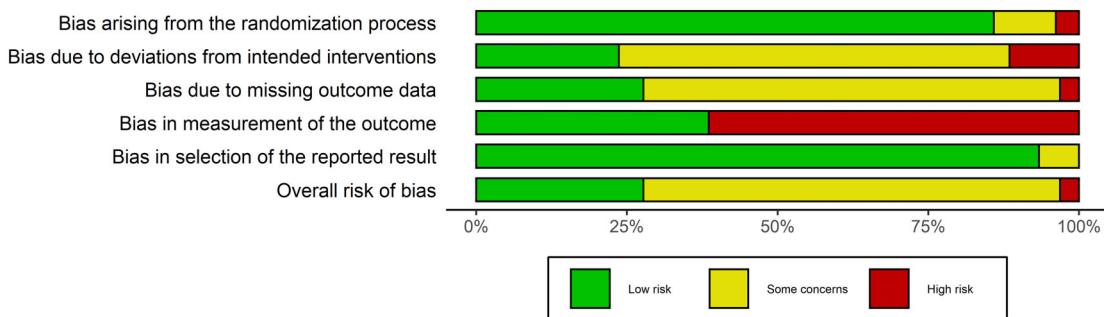


FIGURE 2 Example risk of bias summary plot of ROB2 assessments created using **robvis** and the standard Cochrane palette [Colour figure can be viewed at wileyonlinelibrary.com]

Using the example data set (`data_rob2`) that is built into the package and is presented in Table 1 for reference, the traffic light plot shown in Figure 1 is created using:

```
rob_traffic_light(data = data_rob2,
                  tool = "ROB2",
                  colour = "colourblind",
                  psize = 15)
```

Similarly, using the same data set, the summary barplot shown in Figure 2 is created using:

```
rob_summary(data = data_rob2,
            tool = "ROB2",
            overall = TRUE)
```

3.1.4 | Further customization

The `ggplot2` package in R, based on the “The Grammar of Graphics,” allows users to create detailed graphics and was used to create the templates found in **robvis**.^{22,23} As a result, both **robvis** functions return a `ggplot` object, meaning they can be easily customized further using the `ggplot2` framework. For example, to add a title to the summary bar plot:

```
library(ggplot2)

plot <- rob_summary(data_rob2,
                     tool = "ROB2")

plot +
  ggtitle("Summary Bar Plot")
```

For a full discussion of the range of post-production modifications that can be made to the plots via the `ggplot2` package, we refer the reader to the extensive guidance available.²³

3.2 | Shiny web app

Although **robvis** was originally designed for use in the command-line-based R programming environment, we developed a web app to make the tool accessible to those without knowledge of R. This is available via www.riskofbias.info.²⁴ The app was built using **Shiny**,¹³ an R package which makes it easy to produce interactive web-apps, and provides a graphical user interface (GUI) for the **robvis** package, allowing users to interact with the functions presented above without the need to download R or type any commands. Users can upload their data as either an Excel spreadsheet (recommended) or a comma-separate values (CSV) file, or can manually enter it directly into the app. Uploaded data are passed through a number of quality control checks that ensure the app will work correctly, and users are prevented from producing the plots until any issues identified have been addressed. Finally, users can customize their plots by defining a color scheme and other parameters, all through the online app.

4 | DISCUSSION

robvis facilitates the rapid production of two common risk-of-bias assessment figures at publication quality. By implementing **robvis** both as an R package and a **Shiny** web app, its functionality is available to evidence synthesists with varying levels of ability in R. **robvis** serves as an example of the advantages of “packaging” the R scripts that evidence synthesists often create for personal use.²⁵ It is likely that several other evidence synthesists have written scripts to produce similar risk-of-bias plots to those presented here—in fact, we personally know of at least one other research group that has done so. This duplication of time and effort is inefficient, and creating and sharing well-documented R packages represents one way to reduce this inefficiency. Taking this approach one step further, **Shiny** apps represent a straightforward way to provide a user-friendly GUI for a

newly created R package within a very short time-frame, expanding the potential pool of users of the package to anyone with an internet connection.

Creating a package using R has a number of particular advantages. R provides access to a range of powerful tools including the **ggplot** infrastructure as demonstrated above, and RMarkdown, which enables creation of documents that can be rendered in a range of formats such as PDF, HTML, or Word.²⁶ Furthermore, and focusing specifically on evidence synthesis, building new tools as packages in R allows for easy integration with the range of existing evidence synthesis packages. Recently, the **metaverse** project,²⁷ of which **robvis** is a part, has begun to curate a collection of R packages that cover different aspects of the systematic review and meta-analysis process which, when taken together, form a coherent end-to-end open-source alternative to commercial offerings such as Covidence or Review Manager. Key offerings in this suite of packages include **litsearcher**, which facilitates systematic search strategy development, **revtools**, a package for managing the review process and performing title and abstract screening, **metaDigitise**, a package for automatic extraction of data from figures in research papers, and **metafor**, a package for conducting meta-analyses in R.²⁸⁻³¹

While **robvis** is a stable package, a range of additional functionality could be added. At present, the number of tools with a specific template included in **robvis** is limited—adding additional templates is a priority. For example, a template for ROBIS, the risk-of-bias assessment tool for systematic reviews,⁶ is in development. Additionally, the **robvis** tool does not yet allow for the production of paired forest plots, where the risk-of-bias judgement is presented alongside each specific result included in the meta-analysis.² This was initially considered to be beyond the scope of the tool, as it involves the visualization of something other than risk-of-bias assessments. However, following user-driven demand, this functionality is in development and will be available in the near future. Finally, we would like to add similar functionality to that provided by the **metafor::reporter()** function, which generates a brief paragraph of text describing the results of a meta-analysis. The future **robvis::reporter()** function would provide a boilerplate description of the assessment tool used and the key domains at risk of bias.

ACKNOWLEDGMENTS

The following people provided valuable feedback that contributed to the development of the **robvis** tool: Emily Kothe, Eliza Grames, Matthew Page, Alexandra Bannach-Brown, Kyle Hamilton, Charles Gray, Vincent

Cheng, Wouter van Amsterdamm, Neal Haddaway, and Martin Westgate.

LAM is supported by an NIHR Doctoral Research Fellowship (DRF-2018-11-ST2-048). Attendance at the Evidence Synthesis Hackathon in April 2019 was supported by a grant from the Evidence Synthesis Hackathon itself and the University of Bristol Alumni Foundation. JPTH is a National Institute for Health Research (NIHR) Senior Investigator (NF-SI-0617-10145). JPTH is supported by NIHR Bristol Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. JPTH is supported by the NIHR Applied Research Collaboration West (ARC West) at University Hospitals Bristol NHS Foundation Trust. JPTH is a member of the MRC Integrative Epidemiology Unit at the University of Bristol. The views expressed in this article are those of the authors and do not necessarily represent those of the NHS, the NIHR, MRC, or the Department of Health and Social Care.

CONFLICT OF INTEREST

The author reported no conflict of interest.

DATA AVAILABILITY STATEMENT

The software and data presented in this paper are freely available on GitHub: <https://github.com/mcguinlu/robvis>

ORCID

Luke A. McGuinness  <https://orcid.org/0000-0001-8730-9761>

REFERENCES

1. Donnelly CA, Boyd I, Campbell P, et al. Four principles to make evidence synthesis more useful for policy. *Nature*. 2018; 558(7710):361-364. <https://doi.org/10.1038/d41586-018-05414-4>.
2. Cochrane Handbook for Systematic Reviews of Interventions Version 6.0 (Updated August 2019)Boutron I, Higgins JP, Altman DG, et al. Chapter 7: considering bias and conflicts of interest among the included studies. In: JPT H, Chandler J, Cumpston M, et al., eds. Cochrane; 2019. Retrieved from. www.training.cochrane.org/handbook.
3. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. <https://doi.org/10.1136/bmj.l4898>.
4. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. <https://doi.org/10.1136/bmj.i4919>.
5. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>.
6. Whiting P, Savović J, Higgins JPT, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin*

- Epidemiol.* 2016;69:225-234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>.
7. Higgins JPT, Altman DG. Chapter 8: Assessing risk of bias in included studies. In: JPT H, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Hoboken, New Jersey, United States: Wiley Online Library; 2008:187-241. <https://doi.org/10.1002/9780470712184.ch8>.
 8. Cochrane Collaboration. Review Manager (RevMan) [Computer Program]. 2014. <https://community.cochrane.org/help/tools-and-software/revman-5>.
 9. Marshall C, Brereton P. Systematic review toolbox: a catalogue of tools to support systematic reviews. Paper presented at: Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering. EASE'15. Association for Computing Machinery; 2015: Article 23; 1–6. <https://doi.org/10.1145/2745802.2745824>.
 10. Harrison H, Griffin SJ, Kuhn I, Usher-Smith JA. Software tools to support title and abstract screening for systematic reviews in healthcare: an evaluation. *BMC Med Res Methodol.* 2020;20(1):7. <https://doi.org/10.1186/s12874-020-0897-3>.
 11. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2019. <https://www.R-project.org/>.
 12. RStudio Team. *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.; 2015.
 13. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. Shiny: *Web Application Framework for R*. Boston, MA: RStudio; 2019. Retrieved from <https://CRAN.R-project.org/package=shiny>.
 14. McGuinness LA. Robvis: An R package for visualising risk-of-bias assessments. [Computer Software - v0.3.0]. November 2019. <https://doi.org/10.5281/zenodo.3552342>.
 15. Harrer M, Cuijpers P, Ebert D. *Doing Meta-Analysis in R: A Hands-on Guide*. Erlangen, Germany: Zenodo; 2019.
 16. Gibb K, Seeley A, Quinn T, et al. The consistent burden in published estimates of delirium occurrence in medical inpatients over four decades: a systematic review and meta-analysis study. *medRxiv*. 2019; 19005165. <https://doi.org/10.1101/19005165>.
 17. Habadi MI, Asiri AA, Aloufi SA, et al. Prevalence of panic disorders in the primary health care setting: a systematic review and meta-analysis. *EC Micro*. 2019;16(1):1-9.
 18. Veloso A, Vicente SG, Filipe MG. Effectiveness of cognitive training for school-aged children and adolescents with attention deficit/hyperactivity disorder: a systematic review. *Front Psych*. 2020;10:2983. <https://doi.org/10.3389/fpsyg.2019.02983>.
 19. Simillis C, Singh HKSI, Afentiotou T, et al. Postoperative chemotherapy improves survival in patients with resected high-risk stage II colorectal cancer: results of a systematic review and meta-analysis. *Colorectal Dis*. 2020; Published online 30th January. <https://doi.org/10.1111/codi.14994>.
 20. Tanneru K, Gautam S, Norez D, Kumar J., Alam M.U., Koocheckpour S., Balaji K.C., Joseph C. Meta-analysis and systematic review of intermediate-term follow-up of prostatic urethral lift for benign prostatic hyperplasia. *Int Urol Nephrol* February 2020. <https://doi.org/10.1007/s11255-020-02408-y>. Retrieved on 24 February 2020.
 21. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. <https://doi.org/10.1136/bmj.d5928>.
 22. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag; 2016.
 23. Wilkinson L. *The Grammar of Graphics*. 2nd ed. New York, NY: Springer-Verlag; 2012.
 24. McGuinness L. Robvis web application. 2019. Retrieved on 20 January 2020 from <https://www.riskofbias.info/welcome/robvis-visualization-tool>
 25. Wickham H. R Packages: Organize, Test, Document, and Share Your Code. First. O'Reilly Media, Inc.; 2015. Retrieved on 20 February 2020 from <http://r-pkgs.had.co.nz/>
 26. Xie Y, Allaire JJ, Grolemund G. R Markdown: the Definitive Guide. Chapman and Hall/CRC; 2018. Retrieved on 1 February 2020 from <https://bookdown.org/yihui/rmarkdown/>
 27. Various Authors. Metaverse: an R Ecosystem for Meta-Research. 2020. Retrieved on 24 February 2020 from <https://rmetaverse.github.io/>
 28. Grames EM, Stillman AN, Tingley MW, Elphick CS. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods Ecol Evol*. 2019; 10:1645-1654. <https://doi.org/10.1111/2041-210X.13268>.
 29. Viechtbauer W. Conducting meta-analyses in R with the metanfor package. *J Stat Software*. 2010;36(3):1-48. <https://doi.org/10.18637/jss.v036.i03>.
 30. Pick JL, Nakagawa S, Noble DWA. Reproducible, flexible and high-throughput data extraction from primary literature: the metaDigitise R package. *Methods Ecol Evol*. 2018;10:426-431. <https://doi.org/10.1101/247775>.
 31. Westgate MJ. Revtools: an R package to support article screening for evidence synthesis. *Res Syn Meth*. 2019;10:606-614. <https://doi.org/10.1002/jrsm.1374>.

How to cite this article: McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Syn Meth*. 2021;12: 55–61. <https://doi.org/10.1002/jrsm.1411>

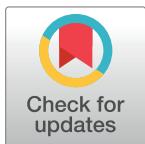
RESEARCH ARTICLE

A descriptive analysis of the data availability statements accompanying medRxiv preprints and a comparison with their published counterparts

Luke A. McGuinness^{1,2*}, Athena L. Sheppard^{1,3}

1 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom, **2** MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom, **3** Department of Health Sciences, University of Leicester, Leicester, United Kingdom

* luke.mcguinness@bristol.ac.uk



Abstract

Objective

To determine whether medRxiv data availability statements describe open or closed data—that is, whether the data used in the study is openly available without restriction—and to examine if this changes on publication based on journal data-sharing policy. Additionally, to examine whether data availability statements are sufficient to capture code availability declarations.

Design

Observational study, following a pre-registered protocol, of preprints posted on the medRxiv repository between 25th June 2019 and 1st May 2020 and their published counterparts.

Main outcome measures

Distribution of preprinted data availability statements across nine categories, determined by a prespecified classification system. Change in the percentage of data availability statements describing open data between the preprinted and published versions of the same record, stratified by journal sharing policy. Number of code availability declarations reported in the full-text preprint which were not captured in the corresponding data availability statement.

Results

3938 medRxiv preprints with an applicable data availability statement were included in our sample, of which 911 (23.1%) were categorized as describing open data. 379 (9.6%) preprints were subsequently published, and of these published articles, only 155 contained an applicable data availability statement. Similar to the preprint stage, a minority (59 (38.1%)) of these published data availability statements described open data. Of the 151 records eligible for the comparison between preprinted and published stages, 57 (37.7%) were

OPEN ACCESS

Citation: McGuinness LA, Sheppard AL (2021) A descriptive analysis of the data availability statements accompanying medRxiv preprints and a comparison with their published counterparts. PLoS ONE 16(5): e0250887. <https://doi.org/10.1371/journal.pone.0250887>

Editor: Rafael Sarkis-Onofre, Meridional Faculty IMED, BRAZIL

Received: September 21, 2020

Accepted: April 15, 2021

Published: May 13, 2021

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.0250887>

Copyright: © 2021 McGuinness, Sheppard. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All materials (data, code and [supporting information](#)) are available from GitHub (<https://github.com/mcguinlu/data->

[availability-impact](#)), archived at time of submission on Zenodo (DOI: [10.5281/zenodo.3968301](https://doi.org/10.5281/zenodo.3968301)).

Funding: LAM is supported by an National Institute for Health Research (NIHR; <https://www.nihr.ac.uk/>) Doctoral Research Fellowship (DRF-2018-11-ST2-048). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The views expressed in this article are those of the authors and do not necessarily represent those of the NHS, the NIHR, MRC, or the Department of Health and Social Care.

Competing interests: The authors have declared that no competing interests exist.

published in journals which mandated open data sharing. Data availability statements more frequently described open data on publication when the journal mandated data sharing (open at preprint: 33.3%, open at publication: 61.4%) compared to when the journal did not mandate data sharing (open at preprint: 20.2%, open at publication: 22.3%).

Conclusion

Requiring that authors submit a data availability statement is a good first step, but is insufficient to ensure data availability. Strict editorial policies that mandate data sharing (where appropriate) as a condition of publication appear to be effective in making research data available. We would strongly encourage all journal editors to examine whether their data availability policies are sufficiently stringent and consistently enforced.

1 Introduction

The sharing of data generated by a study is becoming an increasingly important aspect of scientific research [1, 2]. Without access to the data, it is harder for other researchers to examine, verify and build on the results of that study [3]. As a result, many journals now mandate data availability statements. These are dedicated sections of research articles, which are intended to provide readers with important information about whether the data described by the study are available and if so, where they can be obtained [4].

While requiring data availability statements is an admirable first step for journals to take, and as such is viewed favorably by journal evaluation rubrics such as the Transparency and Openness Promotion [TOP] Guidelines [5], a lack of review of the contents of these statements often leads to issues. Many authors claim that their data can be made “available on request”, despite previous work establishing that these statements are demonstrably untrue in the majority of cases—that when data is requested, it is not actually made available [6–8]. Additionally, previous work found that the availability of data “available on request” declines with article age, indicating that this approach is not a valid long term option for data sharing [9]. This suggests that requiring data availability statements without a corresponding editorial or peer review of their contents, in line with a strictly enforced data-sharing policy, does not achieve the intended aim of making research data more openly available. However, few journals actually mandate data sharing as a condition of publication. Of a sample of 318 biomedical journals, only ~20% had a data-sharing policy that mandated data sharing [10].

Several previous studies have examined the data availability statements of published articles [4, 11–13], but to date, none have examined the statements accompanying preprinted manuscripts, including those hosted on medRxiv, the preprint repository for manuscripts in the medical, clinical, and related health sciences [14]. Given that preprints, particularly those on medRxiv, have impacted the academic discourse around the recent (and ongoing) COVID-19 pandemic to a similar, if not greater, extent than published manuscripts [15], assessing whether these studies make their underlying data available without restriction (i.e. “open”), and adequately describe how to access it in their data availability statements, is worthwhile. In addition, by comparing the preprint and published versions of the data availability statements for the same paper, the potential impact of different journal data-sharing policies on data availability can be examined. This study aimed to explore the distribution of data availability statements’ description of the underlying data across a number of categories of “openness” and to

assess the change between preprint and journal-published data availability statements, stratified by journal data-sharing policy. We also intended to examine whether authors planning to make the data available upon publication actually do so, and whether data availability statements are sufficient to capture code availability declarations.

2 Methods

2.1 Protocol and ethics

A protocol for this analysis was registered in advance and followed at all stages of the study [16]. Any deviations from the protocol are described. Ethical approval was not required for this study.

2.2 Data extraction

The data availability statements of preprints posted on the medRxiv preprint repository between 25th June 2019 (the date of first publication of a preprint on medRxiv) and 1st May 2020 were extracted using the medrxivr and rvest R packages [17, 18]. Completing a data availability statement is required as part of the medRxiv submission process, and so a statement was available for all eligible preprints. Information on the journal in which preprints were subsequently published was extracted using the published DOI provided by medRxiv and rcross-ref [19]. Several other R packages were used for data cleaning and analysis [20–33].

To extract the data availability statements for published articles and the journals data-sharing policies, we browsed to the article or publication website and manually copied the relevant material (where available) into an Excel file. The extracted data are available for inspection (see Material availability section).

2.3 Categorization

A pre-specified classification system was developed to categorize each data availability statement as describing either open or closed data, with additional ordered sub-categories indicating the degree of openness (see Table 1). The system was based on the “Findability” and “Accessibility” elements of the FAIR framework [34], the categories used by previous effort to categorize published data availability statements [4, 11], our own experience of medRxiv data availability statements, and discussion with colleagues. Illustrative examples of each category were taken from preprints included in our sample [35–43].

The data availability statement for each preprinted record were categorized by two independent researchers, using the groups presented in Table 1, while the statements for published articles were categorized using all groups barring Category 3 and 4 (“Available in the future”). Records for which the data availability statement was categorized as “Not applicable” (Category 1 from Table 1) at either the preprint or published stage were excluded from further analyses. Researchers were provided only with the data availability statement, and as a result, were blind to the associated preprint metadata (e.g. title, authors, corresponding author institution) in case this could affect their assessments. Any disagreements were resolved through discussion.

Due to our large sample, if authors claimed that all data were available in the manuscript or as a S1 File, or that their study did not make use of any data, we took them at their word. Where a data availability statement met multiple categories or contained multiple data sources with varying levels of openness, we took a conservative approach and categorized it on the basis of the most restrictive aspect (see S1 File for some illustrative examples). We plotted the

Table 1. Categories used to classify the data availability statements.

Key	Main category	Sub-category	Example
0	Not applicable (protocol for a review, commentary, etc)		"Data sharing not applicable to this article as no datasets were generated or analysed during the current study." [35]
1	"Closed"	Data not made available	"Not available for public" [36]
2	"Closed"	Data available on request to authors	"Data can be available upon reasonable request to the corresponding author." [37]
3	"Closed"	Data will be made available in the future (link provided)	"The protocol and full dataset will be available at Open Science Framework upon peer review publication (https://osf.io/rvbuyl/)." [38]
4	"Closed"	Data will be made available in the future (no link provided)	"Data will be deposited in Dryad upon publication" [39]
5	"Closed"	Data available from central repository (access-controlled or open access), but insufficient detail available to find specific dataset	"Data were obtained from the international MSBase cohort study. Information regarding data availability can be obtained at https://www.msbase.org/ ." OR Daily diagnosis number of countries outside China is download from WHO situation reports (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports [40]
6	"Closed"	Data available from central access-controlled repository, and sufficient details included to identify specific dataset e.g. via extract or accession ID or date stamp	"This research has been conducted using the UK Biobank Resource under application number 24494. All bona fide researchers can apply to use the UK Biobank resource for health related research that is in the public interest." [41]
7	"Open"	Data available in the manuscript/ S1 File	"All data related to this study are present in the paper or the S1 File ." [42]
8	"Open"	Data available via a online repository that is not access-controlled e.g. Dryad, Zenodo	"Extracted data used in this meta-analysis and analysis code are available at www.doi.org/10.5281/zenodo.3149365 ." [43]

Illustrative examples of each category were taken from preprints included in our sample (see "Data extraction").

<https://doi.org/10.1371/journal.pone.0250887.t001>

distribution of preprint and published data availability statements across the nine categories presented in [Table 1](#).

Similarly, the extracted data-sharing policies were classified by two independent reviewers according to whether the journal mandated data sharing (1) or not (0). Where the journal had no obvious data sharing policy, these were classified as not mandating data sharing.

2.4 Changes between preprinted and published statements

To assess if data availability statements change between preprint and published articles, we examined whether a discrepancy existed between the categories assigned to the preprinted and published statements, and the direction of the discrepancy ("more closed" or "more open"). Records were deemed to become "more open" if their data availability statement was categorized as "closed" at the preprint stage and "open" at the published stage. Conversely, records described as "more closed" were those moving from "open" at preprint to "closed" on publication.

We declare a minor deviation from our protocol for this analysis [16]. Rather than investigating the data-sharing policy only for journals with the largest change in openness as intended, which involved setting an arbitrary cut-off when defining "largest change", we systematically extracted and categorized the data-sharing policies for all journals in which preprints had subsequently been published using two categories (1: "requiring/mandating data sharing" and, 2: "not requiring/mandating data sharing"), and compared the change in openness between these two categories. Note that Category 2 includes journals that encourage data sharing, but do not make it a condition of publication.

To assess claims that data will be provided on publication, the data availability statements accompanying the published articles for all records in Category 3 (“Data available on publication (link provided)”) or Category 4 (“Data available on publication (no link provided)”) from [Table 1](#) were assessed, and any difference between the two categories examined.

2.5 Code availability

Finally, to assess whether data availability statements also capture the availability of programming code, such as STATA do files or R scripts, the data availability statement and full text PDF for a random sample of 400 preprinted records were assessed for code availability (1: “code availability described” and 2: “code availability not described”).

3 Results

The data availability statements accompanying 4101 preprints registered between 25th June 2019 and 1st May 2020 were extracted from the medRxiv preprint repository on the 26th May 2020 and were coded by two independent researchers according to the categories in [Table 1](#). During this process, agreement between the raters was high (Cohen’s Kappa = 0.98; “almost perfect agreement”) [44].

Of the 4101 preprints, 163 (4.0%) in Category 0 (“Not applicable”) were excluded following coding, leaving 3938 remaining records. Of these, 911 (23.1%) had made their data open as per the criteria in [Table 1](#). The distribution of data availability statements across the categories can be seen in [Fig 1](#). A total of 379 (9.6%) preprints had been subsequently published, and of these, only 159 (42.0%) had data availability statements that we could categorize. 4 (2.5%) records in

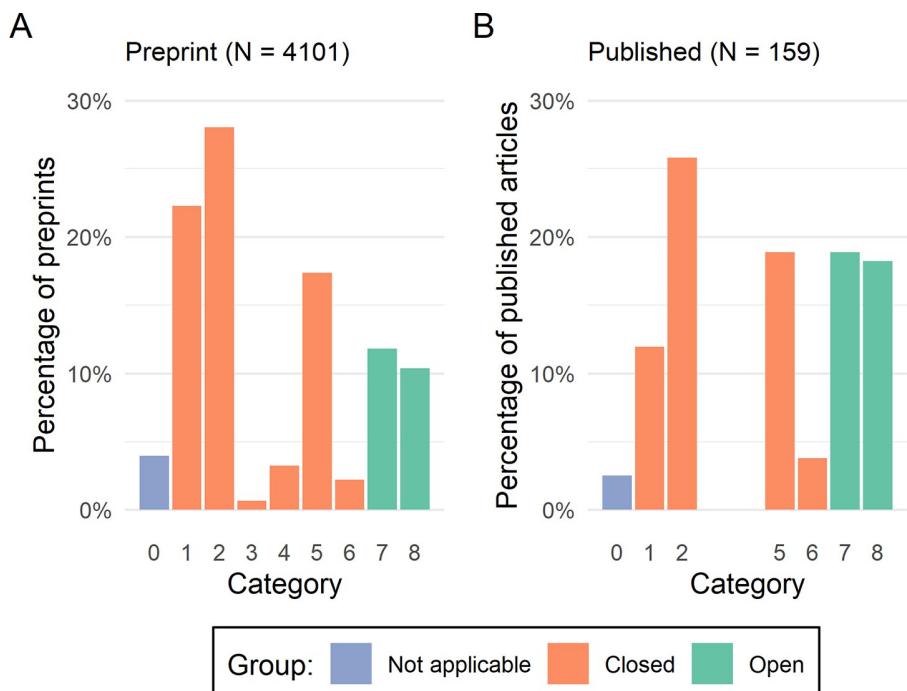


Fig 1. Distribution of the data availability statements of preprinted (Panel A) and published (Panel B) records by category from [Table 1](#).

<https://doi.org/10.1371/journal.pone.0250887.g001>

Table 2. Change in openness of data availability statements from preprint to published article, grouped by journal data-sharing policy.

Journal data sharing policy	Preprinted records subsequently published (N)	Open DAS in preprinted version % (N)	Open DAS in published version % (N)	Change in DAS from preprint to publication		
				More open (N)	More closed (N)	No change (N)
Does not mandate open data	94	20.2% (19)	22.3% (21)	10	8	76
Mandates open data	57	33.3% (19)	61.4% (35)	16	0	41

<https://doi.org/10.1371/journal.pone.0250887.t002>

Category 0 (“Not applicable”) were excluded, and of the 155 remaining, 59 (38.1%) had made their data open as per our criteria.

For the comparison of preprinted data availability statements with their published counterparts, we excluded records that were not published, that did not have a published data availability statement or that were labeled as “Not applicable” at either the preprint or published stage, leaving 151 records (3.7% of the total sample of 4101 records) records.

Data availability statements more frequently described open data on publication compared to the preprinted record when the journal mandated data sharing (Table 2). Moreover, the data availability statements for 8 articles published in journals that did not mandate open data sharing became less open on publication. The change in openness for preprints grouped by category and stratified by journal policy is shown in S1 Table in [S1 File](#), while the change for each individual journal included in our analysis is shown in S2 Table in [S1 File](#).

Interestingly, 22 records published in a journal mandating open data sharing did not have an open data availability statement. The majority of these records described data that was available from a central access-controlled repository (Category 5 or 6), while in others, legal restrictions were cited as the reason for lack of data sharing. However, in some cases, data was either insufficiently described or was only available on request (S3 Table in [S1 File](#)), indicating that journal policies which mandate data sharing may not always be consistently applied allowing some records to slip through the gaps.

161 (4.1%) preprints stated that data would be available on publication, but only 10 of these had subsequently been published (Table 3) and the number describing open data on publication did not seem to vary based on whether the preprinted data availability statements include a link to an embargoed repository or not, though the sample size is small.

Of the 400 records for which code availability was assessed, 75 mentioned code availability in the preprinted full-text manuscript. However, only 22 (29.3%) of these also described code availability in the corresponding data availability statement (S4 Table in [S1 File](#)).

Table 3. Assessment of whether researchers promising to make data available on publication actually do so, and whether this differs if researchers included a link to an embargoed repository or not.

Preprint Category	Number of preprints	Published Category	Number of published studies
Data available in the future, with a link to an embargoed repository provided	3	1. Data not made available 5. Data available from central repository (access-controlled or open access), but insufficient detail available to find specific dataset 8. Data available via a online repository that is not access-controlled e.g. Dryad, Zenodo	1 (33.3%) 1 (33.3%) 1 (33.3%)
Data available in the future, with no details of embargoed repository given	7	1. Data not made available 2. Data available on request to authors 7. Data available in the manuscript/ S1 File 8. Data available via a online repository that is not access-controlled e.g. Dryad, Zenodo	1 (14.3%) 1 (14.3%) 1 (14.3%) 4 (57.1%)

<https://doi.org/10.1371/journal.pone.0250887.t003>

4 Discussion

4.1 Principal findings and comparison with other studies

We have reviewed 4101 preprinted and 159 published data availability statements, coding them as “open” or “closed” according to a predefined classification system. During this labor-intensive process, we appreciated statements that reflected the authors’ enthusiasm for data sharing (“YES”) [45], their bluntness (“Data is not available on request.”) [46], and their efforts to endear themselves to the reader (“I promise all data referred to in the manuscript are available.”) [47]. Of the preprinted statements, almost three-quarters were categorized as “closed”, with the largest individual category being “available on request”. In light of the substantial impact that studies published as preprints on medRxiv have had on real-time decision making during the current COVID-19 pandemic [15], it is concerning that data for these preprints is so infrequently readily available for inspection.

A minority of published records we examined contained a data availability statement ($n = 159$ (42.0%)). This lack of availability statement at publication results in a loss of useful information. For at least one published article, we identified relevant information in the pre-printed statement that did not appear anywhere in the published article, due to it not containing a data availability statement [48, 49].

We provide initial descriptive evidence that strict data-sharing policies, which mandate that data be made openly available (where appropriate) as a condition of publication, appear to succeed in making research data more open than those that do not. Our findings, though based on a relatively small number of observations, agree with other studies on the effect of journal policies on author behavior. Recent work has shown that “requiring” a data availability statement was effective in ensuring that this element was completed [4], while “encouraging” authors to follow a reporting checklist (the ARRIVE checklist) had no effect on compliance [50, 51].

Finally, we also provide evidence that data availability statements alone are insufficient to capture code availability declarations. Even when researchers wish to share their code, as evidenced by a description of code availability in the main paper, they frequently do not include this information in the data availability statement. Code sharing has been advocated strongly elsewhere [52–54], as it provides an insight into the analytic decisions made by the research team, and there are few, if any, circumstances in which it is not possible to share the analytic code underpinning an analysis. Similar to data availability statements, a dedicated code availability statement which is critically assessed against a clear code-sharing policy as part of the editorial and peer review processes will help researchers to appraise published results.

4.2 Strengths and limitations

A particular strength of this analysis is that the design allows us to compare what is essentially the same paper (same design, findings and authorship team) under two different data-sharing policies, and assess the change in the openness of the statement between them. To our knowledge this is the first study to use this approach to examine the potential impact of journal editorial policies. This approach also allows us to address the issue of self-selection. When looking at published articles alone, it is not possible to tell whether authors always intended to make their data available and chose a given journal due to its reputation for data sharing. In addition, we have examined all available preprints within our study period and all corresponding published articles, rather than taking a sub-sample. Finally, categorization of the statements was carried out by two independent researchers using predefined categories, reducing the risk of misclassification.

However, our analysis is subject to a number of potential limitations. The primary one is that manuscripts (at both the preprint and published stages) may have included links to the data, or more information that uniquely identifies the dataset from a data portal, within the text (for example, in the Methods section). While this might be the case, if readers are expected to piece together the relevant information from different locations in the manuscript, it throws into question what having a dedicated data availability statement adds. A second limitation is that we do not assess the veracity of any data availability statements, which may introduce some misclassification bias into our categorization. For example, we do not check whether all relevant data can actually be found in the manuscript/S1 File (Category 7) or the linked repository (Category 8), meaning our results provide a conservative estimate of the scale of the issue, as previous work has suggested that this is unlikely to be the case [12]. A further consideration is that for Categories 1 (“No data available”) and 2 (“Available on request”), there will be situations where making research data available is not feasible, for example, due to cost or concerns about patient re-identifiability [55, 56]. This situation is perfectly reasonable, as long as statements are explicit in justifying the lack of open data.

4.3 Implications for policy

Data availability statements are an important tool in the fight to make studies more reproducible. However, without critical review of these statements in line with strict data-sharing policies, authors default to not sharing their data or making it “available on request”. Based on our analysis, there is a greater change towards describing open data between preprinted and published data availability statements in journals that mandate data sharing as a condition of publication. This would suggest that data sharing could be immediately improved by journals becoming more stringent in their data availability policies. Similarly, introduction of a related code availability section (or composite “material” availability section) will aid in reproducibility by capturing whether analytic code is available in a standardized manuscript section.

It would be unfair to expect all editors and reviewers to be able to effectively review the code and data provided with a submission. As proposed elsewhere [57], a possible solution is to assign an editor or reviewer whose sole responsibility in the review process is to examine the data and code provided. They would also be responsible for judging, when data and code are absent, whether the argument presented by the authors for not sharing these materials is valid.

However, while this study focuses primarily on the role of journals, some responsibility for enacting change rests with the research community at large. If researchers regularly shared our data, strict journal data-sharing policies would not be needed. As such, we would encourage authors to consider sharing the data underlying future publications, regardless of whether the journal actually mandates it.

5 Conclusion

Requiring that authors submit a data availability statement is a good first step, but is insufficient to ensure data availability, as our work shows that authors most commonly use them to state that data is only available on request. However, strict editorial policies that mandate data sharing (where appropriate) as a condition of publication appear to be effective in making research data available. In addition to the introduction of a dedicated code availability statement, a move towards mandated data sharing will help to ensure that future research is readily reproducible. We would strongly encourage all journal editors to examine whether their data availability policies are sufficiently stringent and consistently enforced.

Supporting information

S1 File.

(DOCX)

Acknowledgments

We must acknowledge the input of several people, without whom the quality of this work would have been diminished: Matthew Grainger and Neal Haddaway for their insightful comments on the subject of data availability statements; Phil Gooch and Sarah Nevitt for their skill in identifying missing published papers based on the vaguest of descriptions; Antica Culina, Julian Higgins and Alfredo Sánchez-Tójar for their comments on the preprinted version of this article; and Ciara Gardiner, for proof-reading this manuscript.

Author Contributions

Conceptualization: Luke A. McGuinness.

Data curation: Luke A. McGuinness.

Formal analysis: Luke A. McGuinness, Athena L. Sheppard.

Investigation: Luke A. McGuinness, Athena L. Sheppard.

Methodology: Luke A. McGuinness, Athena L. Sheppard.

Project administration: Luke A. McGuinness.

Software: Luke A. McGuinness.

Supervision: Luke A. McGuinness.

Validation: Luke A. McGuinness, Athena L. Sheppard.

Visualization: Luke A. McGuinness.

Writing – original draft: Luke A. McGuinness.

Writing – review & editing: Luke A. McGuinness, Athena L. Sheppard.

References

1. Packer M. Data sharing in medical research. *BMJ* 2018; k510. <https://doi.org/10.1136/bmj.k510> PMID: 29444885
2. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data. *BMJ* 2016; i255. <https://doi.org/10.1136/bmj.i255> PMID: 26790902
3. Krumholz HM. Why data sharing should be the expected norm. *BMJ (Clinical research ed)* 2015; 350: h599. <https://doi.org/10.1136/bmj.h599> PMID: 25655137
4. Federer LM, Belter CW, Joubert DJ, et al. Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLOS ONE* 2018; 13:e0194768. <https://doi.org/10.1371/journal.pone.0194768> PMID: 29719004
5. Nosek BA, Alter G, Banks GC, et al. Promoting an open research culture. *Science* 2015; 348:1422–5. <https://doi.org/10.1126/science.aab2374> PMID: 26113702
6. Naudet F, Sakarovitch C, Janiaud P, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: Survey of studies published in *The BMJ* and *PLOS Medicine*. *BMJ* 2018; k400. <https://doi.org/10.1136/bmj.k400> PMID: 29440066
7. Miyakawa T. No raw data, no science: Another possible source of the reproducibility crisis. *Molecular Brain* 2020; 13:24. <https://doi.org/10.1186/s13041-020-0552-2> PMID: 32079532

8. Krawczyk M, Reuben E. (Un)Available upon Request: Field Experiment on Researchers' Willingness to Share Supplementary Materials. *Accountability in Research* 2012; 19:175–86. <https://doi.org/10.1080/08989621.2012.678688> PMID: 22686633
9. Vines TH, Albert AYK, Andrew RL, et al. The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* 2014; 24:94–7. <https://doi.org/10.1016/j.cub.2013.11.014> PMID: 24361065
10. Vasilevsky NA, Minnier J, Haendel MA, et al. Reproducible and reusable research: Are journal data sharing policies meeting the mark? *PeerJ* 2017; 5. <https://doi.org/10.7717/peerj.3208> PMID: 28462024
11. Colavizza G, Hrynaszkiewicz I, Staden I, et al. The citation advantage of linking publications to research data. *PLOS ONE* 2020; 15:e0230416. <https://doi.org/10.1371/journal.pone.0230416> PMID: 32320428
12. Roche DG, Kruuk LEB, Lanfear R, et al. Public Data Archiving in Ecology and Evolution: How Well Are We Doing? *PLOS Biology* 2015; 13:e1002295. <https://doi.org/10.1371/journal.pbio.1002295> PMID: 26556502
13. Tan SC, Flanagan D, Morris E, et al. Research data repositories chosen by researchers across broad range of disciplines, from an analysis of 145,000 data availability statements. *Authorea Published Online First*: July 2020. <https://doi.org/10.22541/au.159422974.49069472>
14. Rawlinson C, Bloom T. New preprint server for medical research. *BMJ* 2019; 365. <https://doi.org/10.1136/bmj.l2301> PMID: 31167753
15. Fraser N, Brierley L, Dey G, et al. Preprinting a pandemic: The role of preprints in the COVID-19 pandemic. *bioRxiv* 2020;2020.05.22.111294. <https://doi.org/10.1101/2020.05.22.111294>
16. McGuinness LA, Sheppard AL. Protocol for a descriptive analysis of the data availability statements accompanying medRxiv preprints. 2020.
17. McGuinness L, Schmidt L. Medrxivr: Accessing and searching medRxiv and bioRxiv preprint data in R. *Journal of Open Source Software* 2020; 5:2651. <https://doi.org/10.21105/joss.02651>
18. Wickham H. *Rvest: Easily harvest (scrape) web pages*. 2019. <https://CRAN.R-project.org/package=rvest>
19. Chamberlain S, Zhu H, Jahn N, et al. *Rcrossref: Client for various 'CrossRef' APIs*. 2020. <https://CRAN.R-project.org/package=rcrossref>
20. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria:: R Foundation for Statistical Computing 2019. <https://www.R-project.org/>
21. Wickham H, Hester J, Chang W. *Devtools: Tools to make developing r packages easier*. 2019. <https://CRAN.R-project.org/package=devtools>
22. Wickham H, François R, Henry L, et al. *Dplyr: A grammar of data manipulation*. 2019. <https://CRAN.R-project.org/package=dplyr>
23. Gohel D. *Flextable: Functions for tabular reporting*. 2020. <https://CRAN.R-project.org/package=flextable>
24. Wickham H. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York 2016. <https://ggplot2.tidyverse.org>
25. Rodriguez-Sanchez F. *Grateful: Facilitate citation of r packages*. 2018. <https://github.com/Pakillo/grateful>
26. Müller K. *Here: A simpler way to find your files*. 2017. <https://CRAN.R-project.org/package=here>
27. Gamer M, Lemon J, Fellows I, et al. *Irr: Various coefficients of interrater reliability and agreement*. 2019. <https://CRAN.R-project.org/package=irr>
28. Gohel D. *Officer: Manipulation of microsoft word and PowerPoint documents*. 2020. <https://CRAN.R-project.org/package=officer>
29. Pedersen TL. *Patchwork: The composer of plots*. 2019. <https://CRAN.R-project.org/package=patchwork>
30. Neuwirth E. *RColorBrewer: ColorBrewer palettes*. 2014. <https://CRAN.R-project.org/package=RColorBrewer>
31. Chan C, Chan GC, Leeper TJ, et al. *Rio: A swiss-army knife for data file i/o*. 2018.
32. Wickham H. *Stringr: Simple, consistent wrappers for common string operations*. 2019. <https://CRAN.R-project.org/package=stringr>
33. Müller K, Wickham H. *Tibble: Simple data frames*. 2019. <https://CRAN.R-project.org/package=tibble>
34. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 2016; 3:160018. <https://doi.org/10.1038/sdata.2016.18> PMID: 26978244

35. Ehrlich OG, Testaverde J, Heller C, et al. Crohns disease and ulcerative colitis patient perspectives on clinical trials and participation. *medRxiv* 2019;19000273. <https://doi.org/10.1101/19000273>
36. Septiandi AA, Aditiawarman A, Tjiong R, et al. Cost-Sensitive Machine Learning Classification for Mass Tuberculosis Screening. *medRxiv* 2019;19000190. <https://doi.org/10.1101/19000190>
37. Solis JCA, Storvoll I, Vanbelle S, et al. Impact of spectrograms on the classification of wheezes and crackles in an educational setting. An interrater study. *medRxiv* 2019;19005504. <https://doi.org/10.1101/19005504>
38. Ebbeling CB, Bielak L, Lakin PR, et al. Higher energy requirement during weight-loss maintenance on a low- versus high-carbohydrate diet: Secondary analyses from a randomized controlled feeding study. *medRxiv* Published Online First: July 2019. <https://doi.org/10.1101/19001248>
39. Barry A, Bradley J, Stone W, et al. Increased gametocyte production and mosquito infectivity in chronic versus incident Plasmodium falciparum infections. *medRxiv* 2020;2020.04.08.20057927. <https://doi.org/10.1101/2020.04.08.20057927>
40. Malpas CB, Ali Manouchehrinia A, Sharmin S, et al. Early clinical markers of aggressive multiple sclerosis. *medRxiv* Published Online First: July 2019. <https://doi.org/10.1101/19002063>
41. Knuppel A, Papier K, Fensom GK, et al. Meat intake and cancer risk: Prospective analyses in UK Biobank. *medRxiv* 2019;19003822. <https://doi.org/10.1101/19003822>
42. Thompson ER, Bates L, Ibrahim IK, et al. Novel delivery of cellular therapy to reduce ischaemia reperfusion injury in kidney transplantation. *medRxiv* 2019;19005546. <https://doi.org/10.1101/19005546>
43. Moriarty F, Ebell MH. A comparison of contemporary versus older studies of aspirin for primary prevention. *medRxiv* 2019;19004267. <https://doi.org/10.1101/19004267>
44. McHugh ML. Interrater reliability: The kappa statistic. *Biochemia Medica* 2012; 22:276–82. PMID: [23092060](#)
45. Chen L, Du X, Liu Y, et al. Comparison of the Clinical Implications among Two Different Nutritional Indic平es in Hospitalized Patients with COVID-19. *medRxiv* Published Online First: May 2020. <https://doi.org/10.1101/2020.04.28.20082644>
46. Hashmi M, Taqi A, Memon MI, et al. A national landscaping survey of critical care services in hospitals accredited for training in a lower-middle income country: Pakistan. *medRxiv* Published Online First: April 2020. <https://doi.org/10.1101/2020.04.22.20071555>
47. Peng L, Liu J, Xu W, et al. 2019 Novel Coronavirus can be detected in urine, blood, anal swabs and oropharyngeal swabs samples. *medRxiv* Published Online First: February 2020. <https://doi.org/10.1101/2020.02.21.20026179>
48. Martin J, Hosking G, Wadon M, et al. A brief report: De novo copy number variants in children with attention deficit hyperactivity disorder. *medRxiv* Published Online First: December 2019. <https://doi.org/10.1101/12.12.19014555>
49. Martin J, Hosking G, Wadon M, et al. A brief report: De novo copy number variants in children with attention deficit hyperactivity disorder. *Translational Psychiatry* 2020; 10:135. <https://doi.org/10.1038/s41398-020-0821-y> PMID: [32398668](#)
50. Hair K, Macleod MR, Sena ES, et al. A randomised controlled trial of an Intervention to Improve Compliance with the ARRIVE guidelines (IICARus). *Research Integrity and Peer Review* 2019; 4:12. <https://doi.org/10.1186/s41073-019-0069-3> PMID: [31205756](#)
51. Kilkenny C, Browne WJ, Cuthill IC, et al. Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLOS Biology* 2010; 8:e1000412. <https://doi.org/10.1371/journal.pbio.1000412> PMID: [20613859](#)
52. Goldacre B, Morton CE, DeVito NJ. Why researchers should share their analytic code. *BMJ* 2019; i6365. <https://doi.org/10.1136/bmj.i6365> PMID: [31753846](#)
53. Eglen SJ, Marwick B, Halchenko YO, et al. Towards standard practices for sharing computer code and programs in neuroscience. *Nature neuroscience* 2017; 20:770–3. <https://doi.org/10.1038/nn.4550> PMID: [28542156](#)
54. Culina A, van den Berg I, Evans S, et al. Low availability of code in ecology: A call for urgent action. *PLOS Biology* 2020; 18:e3000763. <https://doi.org/10.1371/journal.pbio.3000763> PMID: [32722681](#)
55. Goodhill GJ. Practical costs of data sharing. *Nature* 2014; 509:33–3. <https://doi.org/10.1038/509033b> PMID: [24784208](#)
56. Courbier S, Dimond R, Bros-Facer V. Share and protect our health data: An evidence based approach to rare disease patients' perspectives on data sharing and data protection—quantitative survey and recommendations. *Orphanet Journal of Rare Diseases* 2019; 14:175. <https://doi.org/10.1186/s13023-019-1123-4> PMID: [31300010](#)
57. Sánchez-Tójar A, Lagisz M, Moran NP, et al. The jury is still out regarding the generality of adaptive “transgenerational” effects. *Ecology Letters*;n/a. <https://doi.org/10.1111/ele.13479> PMID: [32844521](#)