

Determining the causal effects of lipid levels on risk of dementia: a triangulation of new and existing evidence

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.

Determining the causal effects of lipid levels on risk of dementia: a triangulation of new and existing evidence



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 Background, Theoretical framework, Aims & Objectives	3
1.1 Lay summary	3
1.2 Introduction	4
1.3 Dementia	5
1.3.1 Definition and underlying pathologies	5
1.3.2 Diagnostic criteria	6
1.3.3 Public health importance	8
1.3.4 Treatments	9
1.3.5 Risk factors	9
1.4 Serum lipids	10
1.4.1 Lipid fractions	10
1.4.2 Statins	12
1.4.3 Other lipid regulating agents (LRA)	13
1.5 Evidence for the association between blood lipids and dementia . .	15
1.5.1 Basic science	16
1.5.2 Observational studies	16
1.5.3 Randomised controlled trials	17

Contents

1.5.4	Mendelian randomisation	18
1.6	Theoretical framework: Evidence synthesis	19
1.6.1	Inclusion of preprints	20
1.6.2	Triangulating across study designs	21
1.6.3	Individual patient data meta-analysis	22
1.7	Thesis overview	23
1.7.1	Aims and objectives	23
1.7.2	Structure	24
1.8	Outputs from this thesis	25
1.8.1	Contributions to the scientific literature	25
1.8.2	Presentations/Talks	26
1.8.3	Software	27
1.9	Summary	28
1.10	References	29
2	medrxivr: an R package for systematically searching biomedical preprints	31
2.1	Lay summary	31
2.2	Introduction	32
2.3	Development	33
2.3.1	Success criteria	33
2.3.2	Alternative medRxiv/bioRxiv interfaces	34
2.3.3	Early versions	35
2.3.4	Package infrastructure	36
2.4	Usage	36
2.4.1	Installation	37
2.4.2	Importing preprint metadata	37
2.4.3	Performing a search	39
2.4.4	Refining a search	40
2.4.5	Exporting to a bibliography file	41

2.4.6	Downloading the PDFs of relevant records	42
2.5	Discussion	42
2.5.1	Reception and future plans	42
2.5.2	Use cases	43
2.5.3	Limitations of <code>medrxivr</code>	44
2.5.4	Role of open source tools in evidence synthesis	45
2.6	Summary	45
2.7	References	47
3	Systematic review of all available evidence on the association between blood lipids and dementia outcomes	49
3.1	Lay summary	49
3.2	Introduction	50
3.3	Methods	51
3.3.1	Protocol	51
3.3.2	Contributions	51
3.3.3	Search strategy	52
3.3.4	Study selection	54
3.3.5	Validation of screening process	57
3.3.6	Data extraction	58
3.3.7	Risk-of-bias assessment	60
3.3.8	Analysis methods	64
3.4	Results	68
3.4.1	Initial search and validation of search filters	68
3.4.2	Screening results	68
3.4.3	Validation of screening	70
3.4.4	Characteristics of included studies	71
3.4.5	Risk of bias	79
3.4.6	All-cause dementia	79
3.4.7	Alzheimer's disease	85

Contents

3.4.8	Vascular dementia	89
3.4.9	Dose response meta-analysis of lipid levels	92
3.4.10	Additional analyses	93
3.4.11	Missing evidence	93
3.4.12	Added evidential value of including preprints	94
3.5	Discussion	94
3.5.1	Summary of findings	95
3.5.2	Comparison with previous reviews	97
3.5.3	Inclusion of Mendelian randomisations studies	98
3.5.4	Inclusion of preprints	101
3.5.5	Open data sharing	103
3.5.6	Strengths and limitations	104
3.6	Conclusions	106
3.7	References	107
4	Systematic review of all available evidence on the association between blood lipids and dementia outcomes	109
5	Primary analysis of lipid-regulating agents and dementia outcomes	111
5.1	Lay summary	111
5.2	Introduction	112
5.3	Methods	113
5.3.1	Study protocol	113
5.3.2	Data source	113
5.3.3	Cohort definition	115
5.3.4	Exposures	116
5.3.5	Outcomes	116
5.3.6	Covariates	117
5.3.7	Missing data	120
5.3.8	Estimation methods	120

5.3.9	Immortal time bias and time-varying treatment indicators	121
5.3.10	Time axis	123
5.3.11	Sensitivity analyses	123
5.4	Results	127
5.4.1	Patient characteristics	127
5.4.2	Missing data	132
5.4.3	Primary analysis	133
5.4.4	Sensitivity analyses	135
5.5	Discussion	143
5.5.1	Summary of findings	143
5.5.2	Interpretation of results	143
5.5.3	Strengths and limitations	148
5.5.4	Enabling easy synthesis of this analysis	150
5.5.5	Conclusions	150
5.6	References	152
6	Individual participant data meta-analysis	153
6.1	Lay Summary	153
6.2	Introduction	154
6.3	Methods	155
6.3.1	Eligibility criteria	155
6.3.2	Applying for data access	155
6.3.3	Risk of bias assessment	156
6.3.4	Missing data	156
6.3.5	Variable definition	157
6.3.6	Statistical analysis	157
6.3.7	Comparison with published findings	159
6.3.8	Secondary analyses	159
6.4	Results	160
6.4.1	Data access	160

6.4.2	Included data sources	161
6.4.3	Excluded data sources	164
6.4.4	Analytical results	166
6.5	Discussion	166
6.5.1	Summary of findings	166
6.5.2	Limitations	167
6.5.3	Strengths	169
6.5.4	Reflections on the process	169
6.6	Conclusion	169
6.7	References	170
7	Aetiological triangulation across new and existing evidence sources	171
7.1	Lay summary	171
7.2	Introduction	172
7.2.1	Overview of triangulation	172
7.2.2	Terminology	172
7.3	Methods	173
7.3.1	Data sources	173
7.3.2	Qualitative triangulation	175
7.3.3	Quantitative triangulation	175
7.4	Results	183
7.4.1	Qualitative triangulation	183
7.4.2	Quantitative triangulation	185
7.5	Discussion	186
7.5.1	Summary of findings	186
7.5.2	Limitations and strengths	187
7.5.3	Alternative approaches	191
7.5.4	Need for new method	192
7.5.5	Filling in the gaps	193
7.6	Conclusion	193
7.7	References	194

8 Discussion	195
8.1 Lay summary	195
8.2 Chapter overview	195
8.3 Summary of principal findings	196
8.4 Overall strengths and limitations	196
8.5 Implications of this research	197
8.5.1 Clinical practice	197
8.5.2 Public health	197
8.6 Reproducible research	197
8.7 Public involvement and engagement	198
8.8 Future work	198
8.8.1 Inclusion of preprints	198
8.8.2 Evidence on vascular dementia	198
8.8.3 Reviewing of Mendelian randomisation studies	199
8.8.4 Empirical estimates of bias	199
8.8.5 Improved analytical designs	199
8.9 Overall conclusions	200
9 Bibliography	203
Appendices	
A By Chapter	229
A.1 Chapter 1	229
A.1.1 Publications beyond the scope of this thesis	229
A.1.2 Involvement of patients and the public	230
A.2 Chapter 2	230
A.2.1 Code for publication rate analysis	230
A.3 Chapter ??	231
A.3.1 Search strategy	231
A.3.2 Web of Science Databases Searched	234

Contents

A.3.3	Code to search preprints	235
A.3.4	Calculating Gwet's AC1	236
A.3.5	MR risk of bias tool	237
A.3.6	Forest plots for lipid fractions	237
A.4	Chapter 5	240
A.4.1	Amendments to protocol	240
A.4.2	Code lists	240
A.5	Chapter 6	240
A.5.1	Email and documents sent to potential collaborators	240
A.5.2	Section 2	240
A.6	Chapter 8	240
A.6.1	Catalogue of failures	240
B	Other Appendix	241
B.1	Software used to create this thesis	241
B.2	Producing risk-of-bias visualisations with <code>robvis</code>	241
B.2.1	Introduction	241
B.2.2	Development	243
B.2.3	Installation	243
B.2.4	Usage	244
B.2.5	Reception and Future Plans	246
B.3	Creating	247
B.4	Copies of papers arising from this thesis	248

List of Figures

1.1	Statin mechanism of action	12
2.1	Role of <code>medrxivr</code> in a systematic review workflow	33
2.2	Overview of <code>medrxivr</code> data sources	39
3.1	PRISMA flow diagram	69
3.2	Geographical distribution of study cohorts	78
3.3	Random-effects meta-analysis of statins on all-cause dementia . . .	80
3.4	Random-effects meta-analysis of statins on all-cause dementia . . .	81
3.5	Random-effects meta-analysis of statins on all-cause dementia . . .	82
3.6	Meta-analysis of hypercholesterolemia on all-cause dementia . . .	83
3.7	Random-effects meta-analysis of four lipid fractions on all-cause dementia	84
3.8	Random-effects meta-analysis of statins on Alzheimer's disease . . .	85
3.9	Random-effects meta-analysis of genetically lowered LDL-c via HMGCR inhibition on Alzheimer's disease	86
3.10	Meta-analysis of hypercholesterolemia on Alzheimer's disease . . .	87
3.11	Random-effects meta-analysis of four lipid fractions on Alzheimer's disease	88
3.12	Summary of duplication across two sample Mendelian randomisation studies	89
3.13	(ref:obsStatinDVaD-scap)	90
3.14	Meta-analysis of hypercholesterolemia on vascular dementia	91
3.15	Random-effects meta-analysis of four lipid fractions on Alzheimer's disease	91
3.16	Dose-response meta-analysis of total cholesterol	92

List of Figures

3.17	Study designs by year of publication	95
3.18	Overview of assumptions in Mendelian randomisation analyses	98
3.19	Growth of preprint repositories over time	102
5.1	Decision tree for assigning dementia subtypes	117
5.2	shortcap	122
5.3	Causal diagram for ideal negative outcome	125
5.4	Attrition of CPRD participants	128
5.5	Results from primary analyses of CPRD data	134
5.6	Complete case vs. imputed data analysis	136
5.7	Results of control outcome analysis	137
5.8	Comparison of different combinations of covariates	138
5.9	Sensitivity analysis: grouped year of entry	139
5.10	Comparison of pregnancy analysis	141
5.11	Analysis stratified by statin properties	142
5.12	Confounding by indication causal diagram	144
5.13	Comparison of code used in this analysis with those used in the Smeeth et al. 2010 paper	148
6.1	Flowchart	161
6.2	IPD analysis of total cholesterol and dementia	166
7.1	Bias direction plot, summarising internal biases related to the	185
B.1	Example risk of bias traffic light plot created using ‘robvis’	244
B.2	Example risk of bias summary plot created using ‘robvis’ and the example ROB2 dataset	245

List of Tables

1.1	Overview of the DSM-5 criteria for dementia and vascular dementia.	7
1.2	Classification of blood lipid levels	11
1.3	Overview of common statins	13
1.4	Summary of available treatments for hyperlipidaemia.	15
3.1	searchOverview	53
3.2	Ranges for Gwet's AC1	57
3.3	Inter-rater agreement	70
3.4	Inter-rater agreement	70
3.5	Characteristics of included studies	72
5.1	Example of CPRD Read code hierarchy	114
5.2	Covariates adjusted for.	118
5.3	Crude rates, stratified by outcome and drug class of interest. . . .	130
5.4	Patient characteristics by drug class	131
5.5	Participants who stopped, switched or added treatments by initial treatment type	132
5.6	Frequency of diagnoses by grouped year of cohort entry	140
5.7	Summary of statin properties (lipophilicity vs hydrophilicity). . . .	141
6.1	Summary of included cohorts	162
6.2	Exclusion reasons for cohorts providing data	165
7.1	Summary of research designs included in this thesis	174
7.2	robLevelsMapping	177
7.3	Additive bias distributions	181

List of Tables

8.1 Covariates adjusted for.	201
A.1 searchHits	231
A.2 Summary of Web of Science databases searched.	235
A.3 Mendelian randomisation risk-of-bias assessment tool	238
B.1 Description of the arguments available in the two main ‘robvis‘ functions. ‘X’ indicates that the option is available for the respective function.	245

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

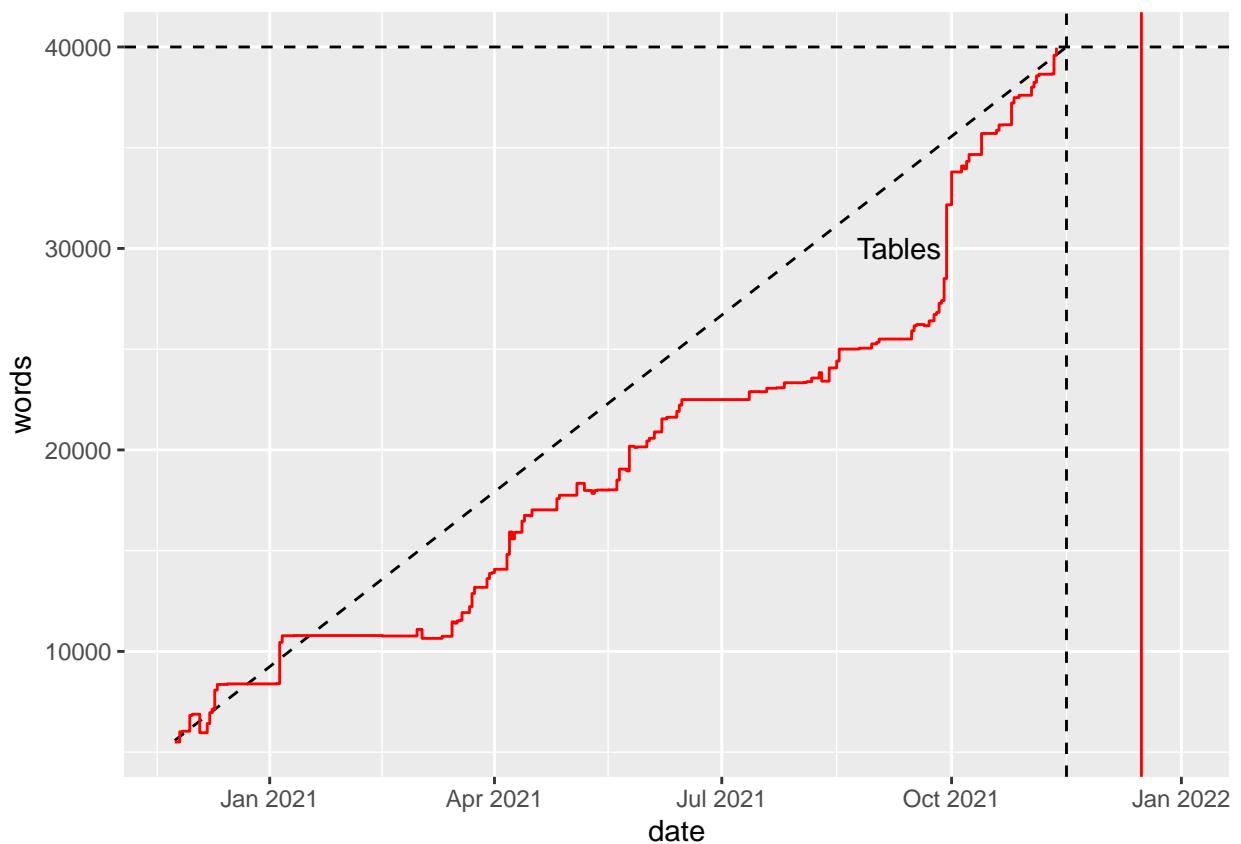
Word count: 39966

Table words: 3484

Days: 353

Words behind: 145

Words today: 380



1

Background, Theoretical framework, Aims & Objectives

Contents

1.1	Lay summary	3
1.2	Introduction	4
1.3	Dementia	5
1.4	Serum lipids	10
1.5	Evidence for the association between blood lipids and dementia	15
1.6	Theoretical framework: Evidence synthesis	19
1.7	Thesis overview	23
1.8	Outputs from this thesis	25
1.9	Summary	28
1.10	References	29

1.1 Lay summary

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.

1.2 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.

1.3 Dementia

1.3.1 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

1.3 - 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

1.3.2 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 1.1: 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 1.1).^{edition2013diagnostic?}

These criteria are outlined in Table 1.1, 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 1.1. 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 1.3.4). 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?}

1.3.3 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.

1.3.4 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

1.3.5 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 1.5 provides an overview of the existing evidence for an association between lipids and dementia outcomes.

1.4 Serum lipids

1.4.1 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

$$TC \approx LDLc + HDLc + kTG \quad (1.1)$$

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 1.2.

Table 1.2: 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%.

1.4.2 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 (HMGCR). 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.

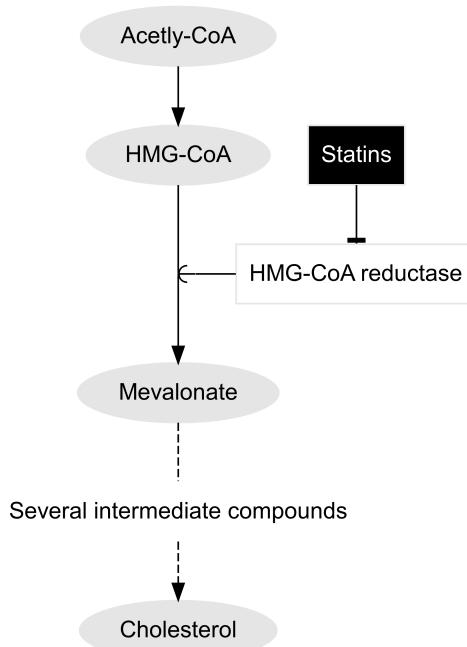


Figure 1.1: Overview of statins mechanism of action, inhibiting HMG-CoA reductase which controls the conversion of HMG-CoA to mevalonate, the rate-limiting step in cholesterol biosynthesis.

Several statin treatments have been widely available for some time (see Table 1.3). 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 1.3: 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	++

1.4.3 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.

1.4 - 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.

1.5 - Evidence for the association between blood lipids and dementia

Table 1.4: 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

1.5 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.

1.5 - Evidence for the association between blood lipids and dementia

1.5.1 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.

1.5.2 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,^{49–52} or a reduced susceptibility.^{53,54} With regards vascular dementia, decreased levels of HDL-c appear to be associated with increased risk,⁵⁴ while for LDL-c, studies have reported both positive and negative associations.^{54,55}

Several previous systematic review of observational studies examining the effect of lipids⁵⁶ and lipid-regulating agents^{57,poly2020c?} on dementia outcomes have been performed. However, these reviews have several limitations. Many did not consider grey literature sources (see Section 1.6.1). Additionally, many of the reviews of observational studies did not perform any risk-of-bias assessment^{chu2018b?} or used an outdated assessment tool.^{56,poly2020c?}

1.5.3 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 1.3.1), as they would require extremely long and costly follow-up.⁵⁸ 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 in 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,⁵⁹ reported not on dementia outcomes but on the change in cognitive scores over a mean of 3.2 years. As highlighted in Section 1.3.2, 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 1.3.1, 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.

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

1.5 - Evidence for the association between blood lipids and dementia

rather than dementia.^{59,2002?} 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?}

1.5.4 Mendelian randomisation

Newer methodological approaches, such as Mendelian randomisation (MR),⁶⁰ 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.⁶¹ 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.⁶⁰ The analytic method relies on several assumptions about the instrumental variable (IV),⁶² 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.^{63,64} 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.⁶⁵ 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 lipid levels. This was supported by further MR studies where *ApoE4* variants were intentionally excluded.⁶⁶

1.6 - Theoretical framework: Evidence synthesis

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.

1.6 Theoretical framework: Evidence synthesis

Evidence synthesis is the process of finding and integrating information from several sources to examine a research question.⁶⁷ A common type of evidence synthesis is a systematic review, either with or without a meta-analysis.⁶⁸

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.

1.6.1 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.^{69–71} Common, well-accepted forms of grey literature include conference abstracts and theses.⁷²

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’⁷³, preprints serve several purposes. They are used to establish primacy when submitting to a journal where the peer-review process may take several months,⁷⁴ to rapidly disseminate research findings, as occurred during the COVID-19 pandemic,⁷⁵ and to make available publications that may not have been accepted elsewhere in an attempt to combat publication bias or the “file-drawer” effect.⁷⁶

One of the major criticisms of using preprints as an evidence source is that they have not yet undergone formal peer review.^{77,78} 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.^{72,79} 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.^{80,81,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,⁷⁹ 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 Chapters 3, a new tool addressing these issues is presented in Chapter ??.

1.6.2 Triangulating across study designs

As illustrated in Section 1.5, 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.⁸²

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 Chapters 3/4. 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.^{83,84,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 ??.

1.6.3 Individual patient data meta-analysis

Individual patient data meta-analyses are commonly held to represent the gold standard in evidence synthesis methodology.^{85,86} 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.⁸⁵

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.⁸⁷⁻⁸⁹ 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.⁸⁵ In terms of this thesis, patient sex is considered to be of particular interest.^{87,90}

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.⁹¹ 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.⁹²

Despite their advantages, IPD meta-analysis are rarely performed.⁹³ 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.^{94,95} The data underlying primary studies are frequently not publicly available,^{96,federer2018a?} and the availability of data “available on request from authors” declines rapidly over time.⁹⁷ 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.⁹⁸

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),⁹⁹ 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 Chapters 3/4. 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.

1.7 Thesis overview

1.7.1 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

1.7.2 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 8.7 for a discussion of the group’s involvement and Appendix A.1.2 for more detail on the group).

- **Chapter 1:** 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 2:** This Chapter introduces a new tool, `medrxivr`, which was used to developed to allow for systematic searches of the health-related preprint repositories.
- **Chapters 3/4:** These Chapters describe, respectively, the methods and results of a comprehensive systematic review and meta-analysis. This review examined 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 8:** 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.

1.8 Outputs from this thesis

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

1.8.1 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 2. 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 2.2 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 Chapters 3/4 inspired a practical guide for researchers. This piece was co-written with Dr. Emily Hennessy (see Author Declarations in the front materials).

1.8 - Outputs from this thesis

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 Chapters 3/4 has been published in Research Synthesis Methods. See Appendix B.2 for more details on this tool. Note that this publication does not describe the recently-developed functionality for producing bias direction plots, as described in Chapter 7.

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.

1.8.2 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.

1.8.3 Software

medrxvir

An R package that allows users to easily search and retrieve bibliographic data from the medRxiv¹⁰⁰ and bioRxiv¹⁰¹ preprint repositories. See Chapter 2 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")
```

1.9 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.

1.10 References

Why are open source statistical programming languages the best?

Because they R.

— Bealy, 2013¹⁰²

2

medrxivr: an R package for systematically searching biomedical preprints

Contents

2.1	Lay summary	31
2.2	Introduction	32
2.3	Development	33
2.4	Usage	36
2.5	Discussion	42
2.6	Summary	45
2.7	References	47

2.1 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.

2.2 Introduction

Preprints represent an increasingly important source of scientific information (see Section 1.6.1). 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 2013,¹⁰¹ and medRxiv, which launched in 2019 and was designed to replace the “Epidemiology” and “Clinical Trial” categories of bioRxiv.¹⁰⁰

Searching these preprints as part of the systematic review described in Chapter ?? was a necessity, as many of the existing reviews on the topic of lipids and dementia have not considered this important source of evidence. 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.^{103,104} 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 I created to facilitate the systematic 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 my own projects and by other researchers is discussed. As the majority of work on this aspect of my thesis is represented by lines of code or online documentation (available at <https://github.com/ropensci/medrxivr> and <https://docs.ropensci.org/medrxivr/> respectively), this chapter is an intentionally short, high-level summary of my work 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.¹⁰⁵

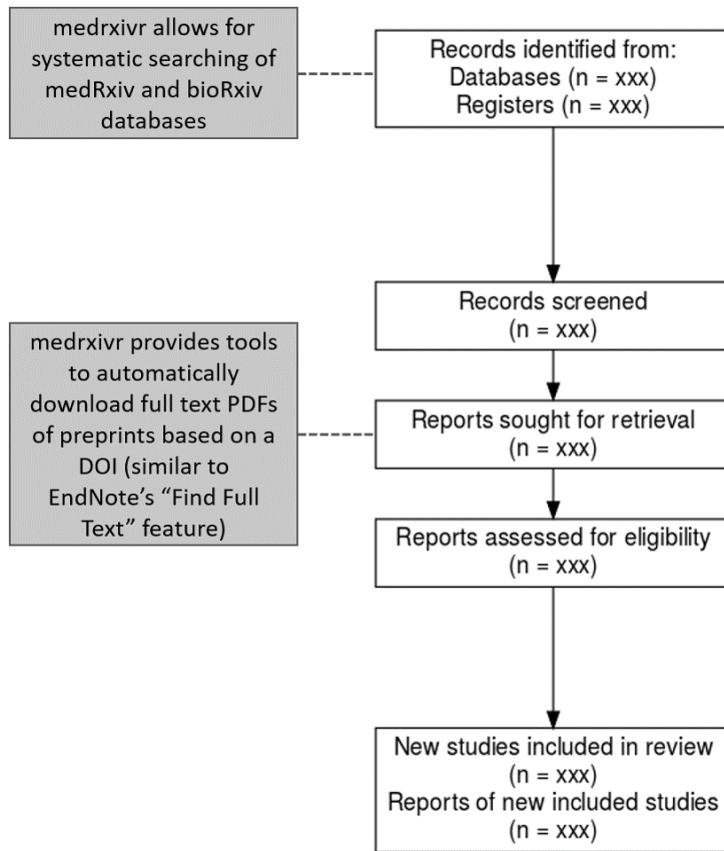


Figure 2.1: Role of medrxivr in a systematic review workflow - medrxivr allows for systematic searching of biomedical preprints as part of the initial literature searching. Following title and abstract screening, reviewers can then programmatically retrieve a copy of the PDF of included records to facilitate the full-text screening stage (similar to Endnote's "Find Full Text" feature).

2.3 Development

2.3.1 Success criteria

I developed the tool to meet three success criteria,¹⁰⁶ influenced both by the functionality required to perform systematic searches as part of the review in Chapter ??, 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.

2.3.2 Alternative medRxiv/bioRxiv interfaces

Prior to development of this tool, I conducted an audit of existing tools for accessing medRxiv and bioRxiv metadata. 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 Rxivist¹⁰⁷, 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 interface¹⁰⁸ 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.¹⁰⁹ 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.

2.3.3 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 ???. 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 I released the initial version of the `medrxivr` R package 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.^{110,111}. In the case of `medrxivr`, as the medRxiv/bioRxiv websites are regularly updated, ensuring the web-scraping performed as expected required me to regularly update or fix the script.

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.¹¹¹

2.3.4 Package infrastructure

I wrote the `medrxvir` package in R using RStudio,¹¹² and followed development best-practice, including development of 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.

2.4 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 users to interact with and search both medRxiv and bioRxiv metadata, as the process is identical for both, searching the medRxiv database is used as an illustrative example throughout this chapter.

2.4.1 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")
```

2.4.2 Importing preprint metadata

Prior to searching the metadata, it must first be imported in R. In `medrixvr`, I have provided two separate but related methods for users to import the data (Figure 2.2). The first of these methods, accessed 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()
```

This provides an up-to-the-minute reflection of the medRxiv preprint repository. However, this approach has two limitations. Firstly, as the API returns results as a series of pages limited to 100 records per page, downloading the entire database requires a time-intensive process of cycling through multiple pages. Secondly, the API can become unavailable, either during peak usage times or planned maintenance windows.

To address these limitations, I provide a second method of accessing medRxiv data, called via the `mx_snapshot()` function, which allows users to access a maintained static snapshot of the database.

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

This snapshot is created each morning at 6am using a process known as “git-scraping”,¹¹³ whereby the entire database is downloaded using the `mx_api_content()` function and saved as a comma separated value (CSV) file to an online server (Figure 2.2). Calling `mx_snapshot()` imports this CSV into R, and has the advantage of both faster loading of the data into R (as it is imported as a single file and does not require cycling through the output of the API) and an absence of any reliance on the API.

The one limitation of this approach is that the snapshot (by its nature) will not contain details of records added to the database since it was taken. However, given that the number of records added each day is relatively low, this should pose minor issues.

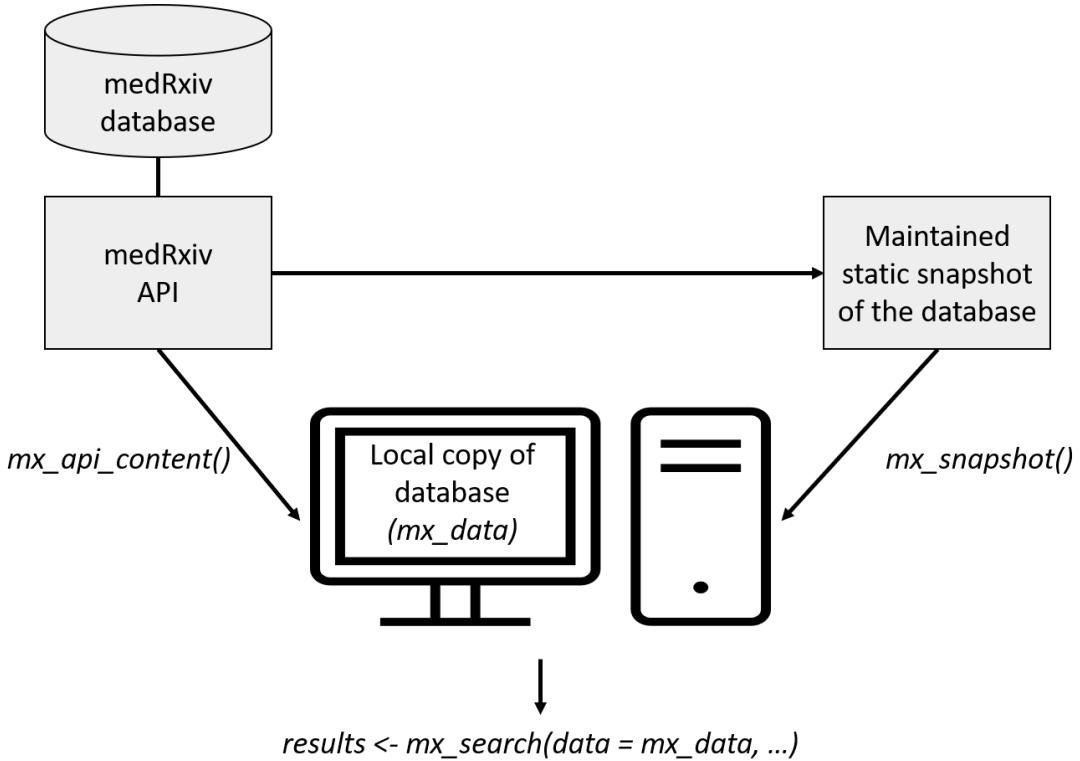


Figure 2.2: Overview of `medrxivr` data sources - Users can either access the API directly via `mx_api_content()`, or can import a maintained snapshot of the database, taken each morning at 6am, via the `mx_snapshot()` function. Note: due to the size of bioRxiv, only a maintained snapshot of the medRxiv repository is available via `mx_snapshot()`.

2.4.3 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)
```

2.4.4 Refining a 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.¹¹⁴

```

results <- mx_search(mx_data,
                      myquery,
                      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,¹¹⁵ 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.

2.4.5 Exporting to a bibliography file

In line with my second success criteria (Section 2.3.1), 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)
```

2.4.6 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 2.1). `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
```

2.5 Discussion

2.5.1 Reception and future plans

The tool has been well received by the community (as of December 2021, `medrxivr` has been downloaded more than 5900 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,¹¹⁶ perform searches of preprints as part of a systematic review,^{117,118} and examine how data-sharing behaviour is affected by journal policies (see 1.8).^{mcguinness2020c?}

The package has been accepted 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”.¹¹⁹ As part of this process, following rigorous peer-review, an associated article introducing the tool was published by the Journal of Open Source Software.^{maguiness2020a?} The entire review discussion is publicly available and can be viewed online.¹²⁰ 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.

2.5.2 Use cases

In addition to being used to search systematically search health-related preprint servers, as illustrated in the systematic review presented in Chapter ??, `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`. By comparing the preprinted and published versions of the data availability statement, I could examine the same manuscript (same content, authors and funders) under two different publication policies, and examine whether stricter policies which require data sharing as a condition of publication actually result in increased data availability. 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.

Secondly, using `medrxivr`, an analysis of the publication rate for medRxiv preprints was performed (see Appendix A.2). Eighty-seven (67.4%) of the 129 records posted on medRxiv in July 2019 were published by 30th July 2021 (i.e. allowing for a two-year lag between preprint posting and publication). This finding agrees with previous work demonstrating that two-thirds of bioRxiv preprints are published in a peer-reviewed journal within two years of posting,¹²¹ indicating that a non-insignificant number of preprints are never formally published but remain accessible as preprints.

In short, these use cases illustrate that easy access to medRxiv/bioRxiv metadata has applications beyond systematic searching of preprints by enabling meta-research/methodological analyses.

2.5.3 Limitations of `medrxivr`

While searching of the medRxiv and bioRxiv databases was crucial for the systematic review element of my thesis presented in Chapter ??, 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.¹²²

There is also the potential that the cross-section of literature posted on medrxiv/bioRxiv is substantially different those suffering from publication bias (studies or analyses that are not published for a range of reasons including results that are not deemed

“novel” or are not statistically significant).¹²³ This is because simply lowering the barriers to publication may well encourage authors to published “null” results, but due to the effort involved in writing up a distributable manuscript, it is unlikely to completely address the “file drawer” effect.⁷⁶

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.

2.5.4 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 hamper the efforts of 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.¹²⁴

2.6 Summary

- In this Chapter, I have introduced a new tool, `medrxivr`, for performing complex systematic searches of the medRxiv and bioRxiv preprint repositories.

2.6 - Summary

- 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 ??.
- 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.

2.7 References

3

Systematic review of all available evidence on the association between blood lipids and dementia outcomes

Contents

3.1	Lay summary	49
3.2	Introduction	50
3.3	Methods	51
3.4	Results	68
3.5	Discussion	94
3.6	Conclusions	106
3.7	References	107

3.1 Lay summary

Systematic reviews are a type of research study that aim to collect and combine all existing evidence to provide the best possible answer to an important research question. Well-performed reviews involve multiple steps including: searching of existing studies; assessment of the studies against predefined inclusion criteria; collection of data from each study; and 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 such as statins, and dementia outcomes.

My review included 81 primary studies that contained information on this relationship. I found that statins appear to reduce the risk of Alzheimer's disease, but had no effect of vascular dementia. Lipid levels 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 added value of including preprinted study reports, made possible using the tool described in the previous chapter, along with the use of the results of this review in subsequent chapters, is discussed.

3.2 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 1.5), several diverse forms of evidence on the relationship of lipids and dementia exist. These include randomised controlled trials, observational studies of different analytical design, 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 describing prospective analyses, 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 synthesists but report relevant unpublished analyses. As a sensitivity analysis to this review presented in this chapter, I sought to quantify the additional evidential value of including preprints, making use of the preprint search tool presented in Chapter 2.

The results of this review are used to guide the primary analysis presented in Chapter 5 and 6, in addition to forming a key evidence source used in the triangulation exercise presented in Chapter 7.

3.3 Methods

3.3.1 Protocol

A pre-specified protocol for this analysis was registered on the Open Science Framework platform and is available for inspection.¹²⁶ Deviations from this protocol are detailed in the relevant sections.

3.3.2 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 review was performed in conjunction with a team of secondary reviewers (see Acknowledgements and Author Declaration in the front matter).

3.3.3 Search strategy

I systematically searched several electronic bibliographic databases to identify potentially relevant entries (hereafter referred to as “records”). 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,¹²⁷ the specific databases and date ranges for each database searched via this platform are listed in Appendix A.3.2. The search strategy used in each database was developed in an iterative manner using a combination of free text and controlled vocabulary (MeSH/EMTREE)⁷² terms to identify studies which have examined the relationship between blood lipids levels and dementia, incorporating input from an information specialist. The strategy included terms related to lipids, lipid modifying treatments, and dementia, and was designed for MEDLINE before being adapted for use in the other bibliography databases listed. A high-level outline of the strategy is presented in the Table 3.1 below and the full search strategies for each database are presented in Appendix A.3.1.

Table 3.1: 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.3.1.

No.	Concept
1	Dementia
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 were 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 3.1) was screened.

I also searched 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 2 to identify potentially relevant preprinted studies (see Appendix A.3.3 for the code used to search these preprint repositories).

Grey literature was 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”).

3.3.4 Study selection

Records were imported into Endnote and de-duplicated using the method outlined in Bramer et al. (2016).¹²⁸ 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 identified as duplicates based on the “Title” field alone).

Following de-duplication of records, screening (both title/abstract and full-text) was performed using a combination of Endnote, a citation management tool,¹²⁹ and Rayyan, a web-based screening application.¹³⁰ Title and abstract screening to remove obviously irrelevant records was performed primarily by me, with a random ~10% sample of excluded records being screened in duplicate to ensure consistency with the inclusion criteria. Additionally, I re-screened the same ~10% sample with 3 month lag to assess intra-rater consistency.

Similarly, I completed all full-text screening, with a random ~20% being screened in duplicate by a second reviewer, in addition to any records 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.⁸³

The criteria against which records were assessed for eligibility are presented in the subsequent sections.

Inclusion criteria

I sought to include studies that examined blood lipid levels as a risk factor for dementia outcomes, defined either as binary hypercholesterolemia variable or by category/1-standard-deviation increase of a specific lipid fraction (total cholesterol, high- and low-density lipoprotein cholesterol and triglycerides). I also aimed to include studies examining the effect of treatments that modify lipids levels as a source of indirect evidence. Eligible study designs included 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.

Eligible studies screened participants for dementia at baseline and excluded any prevalent cases. Alternatively, where no baseline screening was employed, participants were assumed to be dementia free if less than 50 years of age 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 outcomes according to recognised criteria, for example the International Classification of Diseases (ICD),¹³¹ National Institute of Neurological Disorders and Stroke Association-Internationale pour la Recherche en l'Enseignement en Neurosciences (NINDS-AIREN),¹³² or Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria.¹³³ Studies utilising electronic health records were the exception to this, as it was assumed that a valid criteria was employed when entering used when entering the outcome into the EHR.

3.3 - Methods

Conference abstracts with no corresponding full-text publication were eligible, and where required, I contacted authors to obtain information on the study's status. No limitations were imposed on publication status, date, venue or language.

Exclusion criteria

Due to the significant impact of a memory-related outcome such as dementia on exposure recall, case-control studies were excluded, though nested case-control studies, where historical records are used to determine the exposure status, were eligible for inclusion. Cross-sectional studies, qualitative studies, case reports/series and narrative reviews were also excluded, as were 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/dementia). 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 a 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. Note, this is distinct from studies which delivered a simple tablet-based omega-3 intervention, which would have been eligible for inclusion.

3.3.5 Validation of screening process

Inter- and intra-rater reliability during the screening stages were assessed for a 10% sub-sample of records. 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).¹³⁴ This measure was chosen over other methods 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.¹³⁴ Given the small number of included studies in this review as a proportion of the total number screened, this is a useful characteristic.

How to interpret agreement co-efficients is widely debated, and while arbitrary cut-off values may mislead readers,¹³⁷ 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,¹³⁸ presented in Table 3.2.

Table 3.2: 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

Intra- and inter-rater reliability was assessed against these cut-offs. If this assessment demonstrated issues with the screening process (defined as an AC1 of less than .9), a larger proportion of records would have been dual-screened.

3.3.6 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 (study location, data source, 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, p-value). I extracted all data in the first instance, which was subsequently checked for accuracy by a second member of the review team.

Grouping multiple reports into studies

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 was common in cases where multiple papers report results on the same cohort but at different time points. This process builds out the most comprehensive account of a given studies possible by incorporating information from all available records. This was particularly relevant to preprints and published papers reporting the same study, which were not considered to be duplicate records but instead 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.

Combining across groups

In line with best practice, where summary data was presented across two groups (e.g. age at baseline stratified by hypercholesterolemia status), the following approach was used to combine the groups:¹³⁹

$$N = N_1 + N_2 \quad (3.1)$$

$$Mean = \frac{(N_1 M_1 + N_2 N_2)}{(N_1 + N_2)} \quad (3.2)$$

$$SD = \sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2 + \frac{N_1 N_2}{N_1 + N_2}(M_1^2 + M_2^2 - 2M_1 M_2)}{N_1 + N_2 - 1}} \quad (3.3)$$

This was implemented in a systematic manner, with the raw group data being extracted and a cleaning script employed to combine the groups for analysis.

Harmonisation of cholesterol measures

Where necessary, lipid levels reported in *mmol/L* were converted in *mg/dL* using the following formula:

$$mg/dL = mmol/L \times Z \quad (3.4)$$

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 1.2 in Section 1.4.1.

Following up with authors

Where additional data points not included in the report of an analysis were required either for the analysis or risk-of-bias assessment, the corresponding author of the study was contacted. This approach was taken due to the potentially large impact of following up with authors on the results of the review.¹⁴⁰

Analysis of varying effect 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. As such, different effect estimates can be one potential problem that precludes a meta-analysis of all studies.¹⁴¹ If the outcome is rare, as is the case for dementia outcomes, the estimated prevalence of odds and risk ratios will approximate each other. However, hazard ratios provide a very different interpretation, taking into account person-time-at-risk 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 estimate. However, in this review, the small subset of studies reporting odds/risk ratios are synthesised separately to those reporting hazard ratios.

3.3.7 Risk-of-bias assessment

A key aim 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 7.2.1 for an overview of triangulation and Chapter 7 for the results of this analysis). To enable this triangulation 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 away from examining *methodological quality* to assessing *risk of bias*,^{84,142} 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.^{143,144}

This move was prompted by a unclear definition of “methodological quality” which could include facets such as unclear reporting, in addition to challenges in the comparison of results from different tools. As part of this shift, the focus shifted from checklist or score based tools towards domain-based methods, in which different potential sources of bias in a study are assessed in order. Finally, the new tools move from assessing bias at the study level to considering separately each individual numerical result reported. For example, a study may report on the efficacy of an intervention at six months and two years follow-up. In this case, missing outcome data that is not an issue at six months may introduce bias after two years of follow-up, and assigning a single risk-of-bias judgement to the study as a whole masks the different biases applicable to each unique 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.

Randomised controlled trials

Randomized controlled trials were assessed using the RoB 2 tool.¹⁴⁵ 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 for each domain include: “low risk”, “some concerns”, “high risk”. Each of the five domains contains a series of signalling questions or prompts which guide the user through the tool. Once a domain-level judgement for each

3.3 - Methods

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.¹⁴² 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”. In the context of the tool, observational studies are assessed in reference to an idealised randomised controlled trial. Under this approach, the (rare) overall judgement of “Low” indicates that the results should be considered equivalent to produced by a randomised controlled trial.

While a risk-of-bias tool for non-randomised studies of exposures (NRSE) is currently under development,¹⁴⁶ but was insufficiently developed at the time the risk-of-bias assessments for this review were performed. Instead, I used a version of the ROBINS-I tool informed by the preliminary ROBINS-E tool (Risk of Bias In Non-randomised Studies – of Exposure), which I had previously applied in a published review.¹⁴⁷ The version had no signalling questions and so judgements, using the same four levels of bias as ROBINS-I, were made at the domain level. The motivation using this tool above other established tools such as the Newcastle-Ottawa scale (NOS)¹⁴⁸ was two-fold. In the first instance, as mentioned in the introduction to this section, using a domain-based tool has distinct advantages over better-developed checklist-type tools including the NOS. Additionally, 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 as part of this review.

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,¹⁴⁹ as identified by a review of risk-of-bias assessments in systematic reviews of Mendelian randomisation studies (advance results from this review were obtained from contact with the review authors).¹⁵⁰ A copy of this tool is available in Appendix A.3.5, but in summary, results were assessed for bias arising from weak instruments, genetic and other confounding, pleiotropy and population stratification. Acceptable judgements for each of the 5 domains in the tool included “low”, “moderate” and “high” risk of bias.

Risk of bias due to missing evidence

In addition to assessing the risk of bias within each result contributing to a synthesis, I also assessed risk of bias due to missing evidence at the analysis level. This assessment examines evidence missing due to selective non-reporting - as distinct from the selective reporting of a single result from multiple planned - and was performed using the forthcoming RoB-ME (Risk of Bias due to Missing Evidence in a synthesis) tool.¹⁵¹ 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 largely because the tool did not exist when the protocol was registered.

3.3.8 Analysis methods

An initial qualitative synthesis of evidence was performed, summarising the data extracted from studies stratified by study design. Where individual studies were deemed comparable, they were incorporated into a quantitative analysis or “meta-analysis”.

Results were not combined across different study designs (i.e. RCTs were not combined in a meta-analysis with results from observational studies). The summary effect estimates produced by meta-analysis of individual study designs are discussed, but are compared and contrasted more fully as part of the triangulation exercise presented in Chapter 8.

Random-effects meta-analysis

For results examining the effect of binary or continuous exposure (as opposed to categorical/dose-response, as discussed in the next section) on any dementia outcome, a random-effects meta-analysis model was used. Random-effects meta-analysis does not assume one true underlying effect, but rather allows for a distribution of true effects with variance τ^2 . The weight (w) assigned to each result (denoted as y_i) is then given as the inverse of variance of that result plus the estimate of between-result variance, denoted as $w_i = \frac{1}{v_i + \tau^2}$ for result i .

Once the weights are calculated for each result, the overall estimate (\hat{y}) and variance ($Var(\hat{y})$) can be estimated:

$$\hat{y} = \frac{\sum y_i w_i}{\sum w_i} \quad (3.5)$$

$$Var(\hat{y}) = \frac{1}{\sum w_i} \quad (3.6)$$

Results were stratified into subgroups on the basis of the overall risk of bias assessment, and summary estimates for each subgroup, in addition to an overall effect estimate, are displayed in each forest plot. Additional descriptive statistics are presented, while prediction intervals are shown as a dotted line banding the overall effect estimate. Finally, where at least 10 results are available, a test subgroup differences between studies at different levels of risk of bias was performed (see subsequent Section 3.3.8). All models were implemented using the `metafor` R package.

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 contain in the interim groups. In order to address this limitation, I performed a dose-response meta-analysis in those studies reporting more than two categories for lipid levels. This marks a departure from the published protocol, as I was unaware that such a large number of studies would report dementia risk across multiple lipid categories.

Studies were excluded from this analysis if the number of categories was less than three or if the necessary information for synthesis (cut-off points, number of participants and number of events per category) was not available. A restricted cubic 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 the exposure can have different effects versus a “normal” reference dose. The locations of the knots in the model wer identified using fixed percentiles (25th, 50th, 75th) of the exposure data. Reference doses were defined *a priori* as the cut-off of the “Normal”/“Optimal”

3.3 - Methods

categories for each fraction, as detailed in Table 1.2. 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.

When the highest reported category was open ended (e.g LDL-c ≥ 200 mg/dL), I calculated the category midpoint by assuming the width of the highest 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 naturally).

Additional analyses

Where there was evidence of heterogeneity between results included in a meta-analysis, I investigated this further using meta-regression against reported characteristics. *A priori*, I was interested in the effect that the age at baseline, sex and risk-of-bias judgement had on the results. Syntheses with greater than 10 studies were assessed for heterogeneity across these covariates.

Finally, I investigated the potential for small study effects, which may be caused by publication bias, both visually using funnel plots and formally using Egger's regression test.

Visualisation of results

Evidence maps are useful way to explore the distribution of research cohorts included in a systematic review.¹⁵² As part of the initial descriptive synthesis, the location of each individual study contributing to the evidence base was quantified and visualised on a world map.

One of the limitations of current risk-of-bias assessments in systematic reviews is that they are often divorced from the results to which they refer, and are infrequently

incorporated into the analysis. In response to this criticism, I developed a new visualisation tool was designed to allow for the production of “paired” forest plots, as recommended by the ROB2 publication, where a risk-of-bias assessment is presented alongside it’s corresponding numerical results.¹⁴⁵ This tool was developed as an adjunct to this thesis to aid in creating standardised risk-of-bias figures,¹⁵³ and the “paired” forest plot functionality grew out of a collaboration with other researchers to design a modular method for creating custom forest plots.¹⁵⁴ A summary of this tool is contained in Appendix B.2, and all forest plots presented in this Chapter were created using this tool.

Assessment of added value of including preprints

[Note: Julian, I am particularly interested in your feedback on this section, and the corresponding results section (Section 3.4.12), as I am not convinced on the language I am using]

Preprints are considered a valuable evidence source within this thesis (see Introduction, Section 1.6.1). As an adjunct analysis to this review, I explored the additional evidential value of including preprints in each meta-analysis performed, assessed using the fixed effect weight from a standard meta-analysis.

Additionally, I followed identified preprints up after a two-year lag to investigate whether they had been subsequently published (in which case preprints provide a snapshot into the future, and a systematic review update would capture these reports) or not (in which case preprints provide a distinct evidence source to conventional bibliographic databases).

3.4 Results

3.4.1 Initial search and validation of search filters

The database search identified 23,447 records, of which 7,338 were duplicated records. Of the random sample of 500 records screened to ensure the accuracy of the study design filters, no eligible records were identified. Many of those excluded by the filters were commentaries/educational articles, or described basic science studies.

3.4.2 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.

The PRISMA flow diagram presented in Figure 3.1, 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 left of the diagram) and those captured by the search tool described in the previous chapter (presented in grey on the right of the diagram).

Common reasons for exclusion at the full text-stage included studies that reported on the wrong exposures ($n = 70$; most commonly a ineligible lipid fraction), used the wrong study design ($n= 56$), or reported on a wrong outcome ($n=24$; e.g. change in cognitive scores).

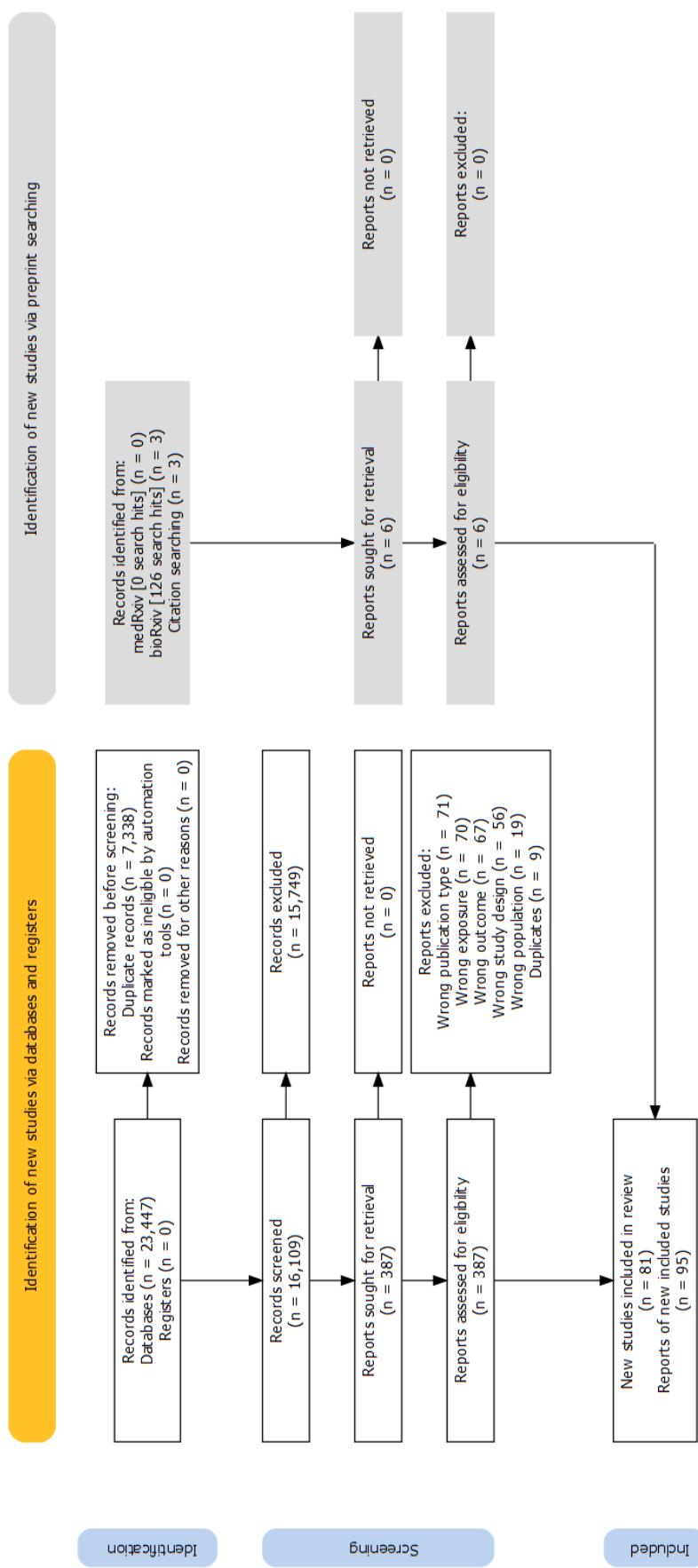


Figure 3.1: PRISMA flow diagram illustrating how records moved through the systematic review process. The different contributions of standard bibliographic databases and preprint servers to the review are indicated.

3.4.3 Validation of screening

For the assessment of the intra-/inter-rater reliability, the estimated values of $AC1$ were interpreted against the categories presented in Table 3.2. For the inter-rater reliability, agreement was “almost perfect” ($AC1 = 0.97$, $kappa = 0.54$, Table 3.3). Similarly for intra-rater reliability, agreement was “almost perfect” ($AC1 = 0.99$, $kappa = 0.65$, Table 3.4). The discrepancy between the $AC1$ and $kappa$ coefficients illustrates the sensitivity of $kappa$ to imbalanced marginals, caused in this sample by a large imbalance towards exclusion.¹⁵⁵

Table 3.3: 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 3.4: Intra-rater agreement on subset of records, indicating high consistency.

		Initial screening decision		
		Exclude	Include	Total
Same reviewer decision (with 3 month lag)	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 3.3) were investigated. This discrepancy between the two reviewers was explained in all cases by differing interpretations of the inclusion criteria, most commonly around the definition of cognitive decline versus dementia and the definition of eligible lipids fractions.

3.4.4 Characteristics of included studies

Following full-text screening, 81 unique studies (described across 95 reports) met the criteria for inclusion in the review.^{49,52,54,156–233} Table 3.5 presents a summary of the characteristics of each study.

The majority of included studies described non-randomised analyses, with the sole two included randomised controlled trials (the Heart Protection Study/British Heart Foundation trial¹⁵⁶ and the JUPITER trial¹⁵⁷) both reporting on the effect of statin use on all-cause dementia in older adults. A similarly small number of Mendelian randomisation studies were identified, several of which employed a two-sample approach using summary statistics from the same published genome wide association studies (GWAS) leading to complications in the synthesis (see Section 3.4.7).

Of the 31 non-randomised studies examining treatments that modify lipid levels, all examined statin use ($n = 31$; 100%), while a small number also reported on other non-statins agents such as fibrates ($n = 2$; 6.5%). In the 43 non-randomised studies of exposure, hypercholesterolemia ($n = 19$; 44.2%) and total cholesterol levels ($n = 21$; 48.8%) were the most frequently reported risk factors.

In terms of outcomes, the vast majority of studies examined either all-cause dementia ($n = 61$; 75.3%) or Alzheimer's disease ($n = 51$; 63%), with only a small proportion examining vascular dementia ($n = 16$; 19.8%). Some other outcome classifications such as vascular-component or mixed dementia were also investigated, but were much rarer.

Three included reports were preprinted (denoted in the Table 3.5 using an asterisk),^{157,232,234} one of which had subsequently been published and was captured by the primary literature search.²³³ All three included preprints were obtained from the bioRxiv preprint server and all described a Mendelian randomisation analysis.

(ref:studyCharacteristics-first-cite)¹⁵⁶

Table 3.5: Characteristics of included studies, stratified by study design. Note that three studies reported on multiple analytical designs within a single study report, and these have been duplicated across the relevant sub-sections.

Study	Location	N	Age at baseline	Female (%)	Exposures	Outcomes	Diagnostic criteria
Randomised controlled trials							
HPS 2002	United Kingdom	20536	>70	24.8	Simvastatin	Dementia	NR
JUPITER 2009	Multiple	17902	66 (median) 60-71 (range)	38	Rosuvastatin	Dementia	NR
Non-randomised studies of interventions							
Ancelin 2012	France	7056	NR	67	Fibrate; Statin	Dementia; AD	DSM-IV; NINCDS-ADRDA
Arvanitakis 2008	United States	929	74.9 (NR)	68.7	Statin	AD	Consortium to Establish a Registry for Alzheimer's Disease (CERAD)
Bettermann 2012	United States	3069	78.6 (3.3)	46.2	Non-statin LRA ; Statin	Dementia; AD; Vascular component	Consensus panel - criteria not reported
Beydoun 2011	United States	1604	57.6 (18.4)	38.5	Statin; TC	Dementia	DSM-III-R
Chao 2015	Taiwan	256265	73.2 (7.4)	50.3	Statin	Non-vascular dementia	ICD-9
Chen 2014	Taiwan	18100	67 (8.6)	47.9	Statin	Dementia; AD; Non-AD	ICD-9
Chitnis 2015	United States	8062	74.47 (9.21)	53.04	Statin	Dementia	ICD-9
Chou 2014	Taiwan	33398	>60	53.9	Statin	Dementia; AD; VAD; Non-vascular dementia	ICD-9
Chuang 2015	Taiwan	123300	54 (13)	49.1	Statin	Dementia	ICD-9
Cramer 2008	United States	1674	70 (6.8)	58	Statin	Dementia/CIND	DSM-IV

3.4 - Results

Table 3.5: Characteristics of included studies, stratified by study design. Note that three studies reported on multiple analytical designs within a single study report, and these have been duplicated across the relevant sub-sections. (*continued*)

Study	Location	N	Age at baseline	Female (%)	Exposures	Outcomes	Diagnostic criteria
Gnjidic 2016	Sweden	2056	>60	NR	Statin	Dementia	DSM-IV
Haag 2009	Netherlands	6992	69.4 (9.1)	60	Non-statin LRA ; Statin	AD	NINCDS-ADRDA
Hendrie 2015	United States	974	76.6 (4.9)	69.7	Statin	Dementia; AD	DSM-IV; Consortium to Establish a Registry for Alzheimer's Disease (CERAD)
Hippisley-Cox 2010	United Kingdom	2004692	46 (14)	51	Statin	Dementia	EHR codelist
Jick 2000	United Kingdom	1364	50-89	61	Non-statin LRA ; Statin	Dementia	EHR codelist
Li 2004	United States	2356	75.1 (6.1)	59.8	Non-statin LRA; Statin	Dementia; AD	DSM-IV; NINCDS-ADRDA
Li 2010	United States	3392	75 (6.2)	59	Statin	AD	NINCDS-ADRDA
Liao 2013	NR	5221	NR	NR	Statin	Dementia	NR
Liu 2019	Taiwan	2012	74 (7.5)	NR	Statin	Dementia	ICD
Pan 2018	Taiwan	14807	65 (13)	43	Statin	Dementia	ICD-9
Parikh 2011	United States	377838	75.53 (6.07)	2	Statin	Dementia	ICD-9
Rea 2005	United States	2798	NR	NR	Non-statin LRA ; Statin	Dementia; AD; Mixed; VaD	NINCDS; NINCDS-ADRDA; Combination; State of California Alzheimer's Disease Diagnostic and Treatment Centers
Redelmier 2019	Canada	28815	76 (NR)	61.3	Statin	Dementia	ICD-9
Reitz 2004	United States	1168	78.4 (6.2)	68.3	HDL-c; LDL-c; Non-HDL-c; Statin; TC; TG	VaD; AD	Cohort criteria; NINCDS-ADRDA

3.4 - Results

Table 3.5: Characteristics of included studies, stratified by study design. Note that three studies reported on multiple analytical designs within a single study report, and these have been duplicated across the relevant sub-sections. (*continued*)

Study	Location	N	Age at baseline	Female (%)	Exposures	Outcomes	Diagnostic criteria
Smeeth 2009	United Kingdom	729529	50 (NR)	40-81	Statin	Dementia; AD; Non-AD	EHR codelist
Solomon 2010	Finland	17597	68 (5.8)	57	Statin	Dementia	EHR codelist
Sparks 2008	United States	2068	75 (3.8)	54	Statin	AD	NINCDS-ADRDA
Szwast 2007	United States	1416	77.3 (5.3)	69.3	Statin	Dementia	DSM-IV
Yang 2015	Taiwan	45973	82 (5.3)	48	Fibrate; LRA (exc. statin + fibrates); Statin	Dementia	ICD-9
Zamrini 2004	United States	3397	73 (NR)	0	Statin	AD	ICD-9
Zandi 2005	United States	3308	NR	NR	Non-statin LRA; Statin	Dementia; AD	DSM-III-R; NINCDS-ADRDA
Non-randomised studies of exposures							
Ancelin 2013	France	7053	74 (5.3)	61.1	Hypercholesterolemia	AD; Dementia	NINCDS-ADRDA; DSM-IV
Batty 2014	United Kingdom	103764	47.3 (18.1)	55	Non-HDL-c; Hypercholesterolemia	Dementia	ICD
Benn 2017	NR	111194	56 (median) 46-66 (range)	55	HMGCR; LDL-c; PCSK-9	AD; VaD; Dementia	NR; ICD; ICD-10
Beydoun 2011	United States	1604	57.6 (18.4)	38.5	Statin; TC	Dementia	DSM-III-R
Bruce 2017	Australia	217	63.6 (8.4)	45.6	HDL-c; TC; TG	Dementia	NR
Chiang 2007	Taiwan	785	58 (7.4)	41.4	TC; TG	Dementia; AD; VaD	ICD-9; NR
Dodge 2011	United States	822	71.6 (4.7)	64.4	Hypercholesterolemia	Dementia; AD	DSM-III-R; Consortium to Establish a Registry for Alzheimer's Disease (CERAD)

Table 3.5: Characteristics of included studies, stratified by study design. Note that three studies reported on multiple analytical designs within a single study report, and these have been duplicated across the relevant sub-sections. (*continued*)

Study	Location	N	Age at baseline	Female (%)	Exposures	Outcomes	Diagnostic criteria
Forti 2010	Italy	749	73 (6.1)	53	Hypercholesterolemia; TG	Dementia; AD; VaD	DSM-IV; NINCDS-ADRDA; NINCDS-AIREN
Gottesman 2017	United States	15407	54.2 (5.8)	55	TC	Dementia	Combination
Gustafson 2012	Sweden	NR	NR	100	TC	AD	NR
Hayden 2006	United States	3308	74.0 (6.4)	58.2	Hypercholesterolemia	Dementia; AD; VaD	DSM-III-R; NINCDS-ADRDA; NINCDS-AIREN
Kimm 2011	South Korea	848505	53 (9.3)	42.2	TC	AD; VaD; Dementia	ICD-10
Kivipelto 2001	Finland	1499	50.4 (6.0)	62	Hypercholesterolemia	AD	NINCDS-ADRDA
Kivipelto 2005	Finland	1449	50.6 (6.0)	62	Hypercholesterolemia	Dementia	DSM-IV
Kuo 2015	Taiwan	67066	62.1(11.4)	48.4	Hypercholesterolemia	Dementia	ICD-9
Li 2005	United States	2141	74.9 (5.9)	60.5	HDL-c; TC	Dementia; AD	DSM-IV; NINCDS-ADRDA
Mainous 2005	United States	6558	NR	NR	Hypercholesterolemia	Dementia; AD	ICD-9
Mielke 2005	Sweden	382	NR	70	TC; TG	Dementia	DSM-III-R
Mielke 2010	France	1460	38-60 (range)	100	Hypercholesterolemia; TG	Dementia; AD	DSM-III-R; NINCDS-ADRDA
Mielke 2012	United States	99	74 (2.5)	100	HDL-c; TC; TG	Dementia; AD	DSM-IV; NINCDS-ADRDA
Muller 2007	United States	542	NR	NR	HDL-c; TG	Dementia; AD	DSM-IV; NINCDS-ADRDA
Noale 2013	Italy	5632	71.3(5.3)	56.3	Hypercholesterolemia; TG	Dementia	DSM-III-R
Notkola 1998	Finland	444	40-59 (range)	0	Hypercholesterolemia	AD	Combination
Peters 2009	Multiple	3336	>80	60.4	HDL-c; TC	Dementia	DSM-IV

3.4 - Results

Table 3.5: Characteristics of included studies, stratified by study design. Note that three studies reported on multiple analytical designs within a single study report, and these have been duplicated across the relevant sub-sections. (*continued*)

Study	Location	N	Age at baseline	Female (%)	Exposures	Outcomes	Diagnostic criteria
Raffaitin 2009	France	7087	73.4 (4.9)	61	Hypercholesterolemia; TG	Dementia; AD; VaD	DSM-IV; NINCDS-ADRDA; Combination
Rantanen 2017	Finland	3309	42 (median) 39–46 (range)	0	TC; Hypercholesterolemia	Dementia; AD; VaD	NR
Reitz 2004	United States	1168	78.4 (6.2)	68.3	HDL-c; LDL-c; Non-HDL-c; Statin; TC; TG	VaD; AD	Cohort criteria; NINCDS-ADRDA
Reitz 2010	United States	1130	75.7 (6.3)	65.7	HDL-c; LDL-c; TC	AD	NINCDS-ADRDA
Ronnemaa 2011	United States	2268	49.6 (0.6)	0	Hypercholesterolemia	AD; VaD; Dementia	NINCDS-ADRDA; ADDTC; DSM-IV
Schilling 2017	France	9294	73.8 (5.3)	61	HDL-c; LDL-c; TC; TG	Dementia; AD; Mixed	DSM-IV; NINCDS-ADRDA; NINCDS-AIREN
Solomon 2007	Finland	1449	50.4 (6.0)	62.1	Hypercholesterolemia	Dementia	NR
Solomon 2009	United States	9844	43 (1.7)	54	TC	AD; VaD	ICD-9
Strand 2013	Norway	48793	42.6 (4.3)	49	TC	Dementia; AD	ICD
Su 2017	United Kingdom	212085	NR	NR	NR	Dementia	NR
Svensson 2019	Japan	781	54.1 (5.6)	NR	HDL-c; Hypercholesterolemia	Dementia	DSM-IV
Tan 2003	United States	1026	76.1 (5.3)	63	HDL-C; TC	AD	NINCDS-ADRDA
Tynkkynen 2016	Finland	13725	48.4 (13.3)	51.6	HDL-c	Dementia; AD	ICD-10
Tynkkynen 2018	Multiple	22623	57 (9.2)	47	HDL-c; LDL-c; TC; TG	Dementia; AD	ICD-10
Wang 2012	Taiwan	1230400	60 (13)	52	Hypercholesterolemia	AD	ICD-9
Whitmer 2005	United States	8845	68 (2.6)	53.7	Hypercholesterolemia	Dementia	ICD-9

Table 3.5: Characteristics of included studies, stratified by study design. Note that three studies reported on multiple analytical designs within a single study report, and these have been duplicated across the relevant sub-sections. (*continued*)

Study	Location	N	Age at baseline	Female (%)	Exposures	Outcomes	Diagnostic criteria
Yamada 2009	Japan	1637	>60	100	NR	Dementia; VaD	NR
Yoshitake 1995	Japan	828	74 (5.9)	59.5	HDL-c; LDL-c; TC; TG	VaD; AD	NINCDS-AIREN; NINCDS-ADRDA
Zimetbaum 1992	United States	350	79 (median) 75-85 (range)	64.5	HDL-c; LDL-c; TC; TG	Dementia	DSM-III-R
Mendelian randomisation studies							
Andrews 2019*	Multiple	54162	NR	NR	HDL-c; LDL-c; TC; TG	AD	NR
Benn 2017	Multiple	111194; 54162	NR	NR	HMGCR; LDL-c; PCSK-9	AD; VaD; Dementia	NR; ICD; ICD-10
Burgess 2017	Multiple	21165	NR	NR	HDL-c; LDL-c; TG	AD	NR
Larsson 2017	Multiple	54162	NR	NR	LDL-c	AD	NR
Mukherjee 2013	Multiple	54162	NR	NR	HDL-c; LDL-c; TG	AD	NR
Ostergaard 2017	Multiple	54162	NR	NR	HDL-c; LDL-c; TC; TG	AD	NR
So 2017*	Multiple	54162	NR	NR	HMGCR	AD	NR
Zhu 2018*	Multiple	54162	NR	NR	HDL-c; LDL-c; TG	AD	NR

* Denotes preprinted study.

Abbreviations: AD - Alzheimer's disease; DSM - Diagnostic and Statistical Manual (Roman numerals indicate edition); EHR - Electronic code list; ICD - International Classification of Disease (numbers indicate edition); HDL-c - high density lipoprotein cholesterol; LDL-c - low density lipoprotein cholesterol; NINCDS-ADRDA - National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; NINCDS-AIREN - National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neurosciences; NR - Not reported; TC - total cholesterol; TG - triglycerides; VaD - vascular dementia.

3.4 - Results

As illustrated in Figure 3.2, the majority of reports described studies conducted in high-income countries, with the most high represented region being North America. Of interest, several of the included studies were conducted in Taiwan ($n = 10$; 12.35%), all but one of which made use of the Taiwan National Health Insurance database.



Figure 3.2: Geographical distribution of study cohorts

Finally, there were several eligible studies reported as conference abstracts that did not present numerical results. These reports were included in the analysis to enable assessment of risk of bias due to missing evidence (see Section 3.3.7).

3.4.5 Risk of bias

As discussed above, the risk-of-bias assessments are presented alongside their corresponding numerical result. A more detailed discussion of the sources and directions of bias is presented in Chapter 7, and so this section presents a brief summary of the biases observed in each study design.

For the two randomised controlled trials, both were judged to be at low risk of bias. In contrast, many of the non-randomised studies of statin use were at serious risk of bias due poor controlling for confounding, immortal time bias, and missing outcome data (as noted above, a judgement of low risk of bias for a non-randomised study is rare). Similarly, non-randomised studies of exposures suffered from incomplete adjustment for potentially important confounders, and concerns over the selection of the reported result from among several analyses (e.g. examination of lipids as a binary or continuous variable). Finally, bias was introduced into Mendelian randomisation studies via the potential for horizontal pleiotropy and population stratification.⁶²

Following best practice, any result judged to be at critical risk of bias should be excluded from any quantitative analyses. Four observational studies were excluded on this basis, predominantly due to a lack of adjustment for any potentially important confounders (i.e. the study reported unadjusted estimates).^{202,203,209}

3.4.6 All-cause dementia

Statins

The two randomised controlled trials provided very weak evidence (OR: 1.07, 95%CI: 0.70-1.66) of an effect on statin use on all-cause dementia risk (Figure 3.3).

3.4 - Results

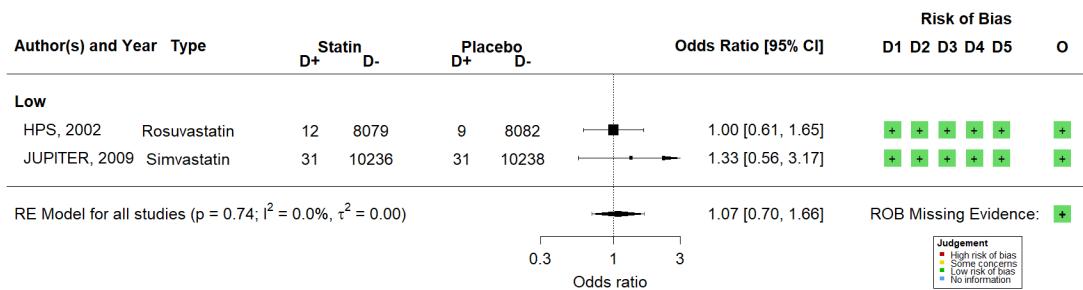


Figure 3.3: Random-effects meta-analysis of randomised controlled trials examining statin statins on all-cause dementia

In contrast, a meta-analysis of 15 prospective observational studies provided some

evidence of a protective effect of statins use on all-cause dementia risk (HR: 0.77, 95%CI: 0.65-0.90, Figure 3.4).

3.4 - Results

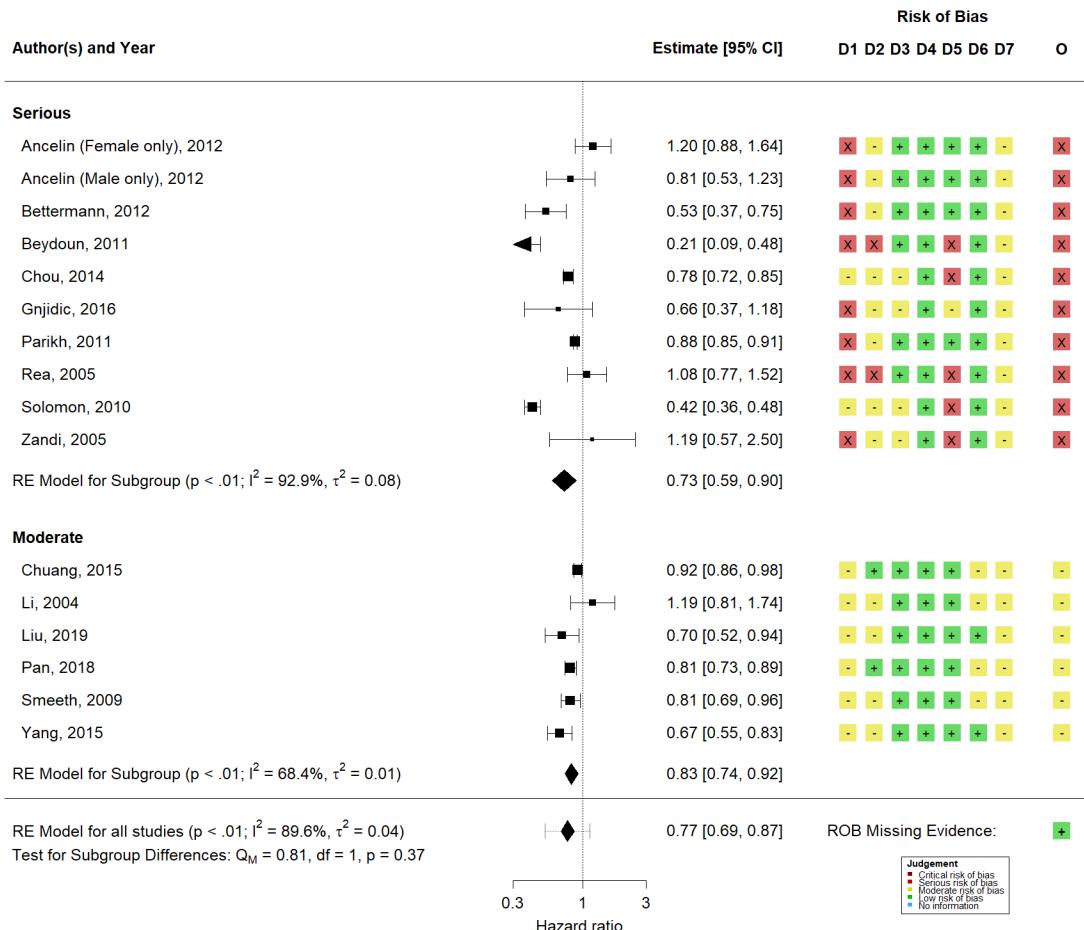


Figure 3.4: Random-effects meta-analysis of non-randomised studies examining the effect of statin use on all-cause dementia

Finally, a single Mendelian randomisation analysis was identified examining the effect of lowered LDL-c levels on the risk of all-cause dementia via genetic inhibition of the 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), emulating statin treatment (see Section 1.4.2 for more details of the statin mechanism of action). This analysis provided weak evidence for an effect (RR: 0.90, 95%CI: 0.29-2.81).

3.4 - Results

Fibrates

Two studies examined the effect of fibrate use on all-cause dementia and found very weak evidence for an effect (HR: 0.90, 95%CI: 0.75-1.07, Figure 3.5).

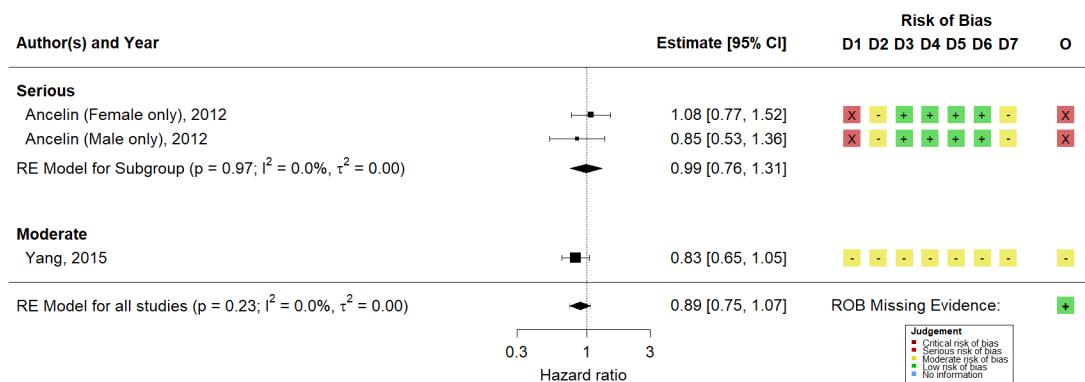


Figure 3.5: Random-effects meta-analysis of non-randomised studies examining the effect of fibrate use on all-cause dementia

Lipids

Across all outcomes, lipid levels were categorised in a number of ways. The most common categorisation was hypercholesterolemia at baseline, defined most frequently as a total cholesterol measurement of greater than 6.5 mmol/L.

Eleven studies reported on the association of hypercholesterolemia with all-cause dementia and provided weak evidence for an effect (HR: 1.11, 95%CI: 0.99-1.24, Figure 3.6).

3.4 - Results

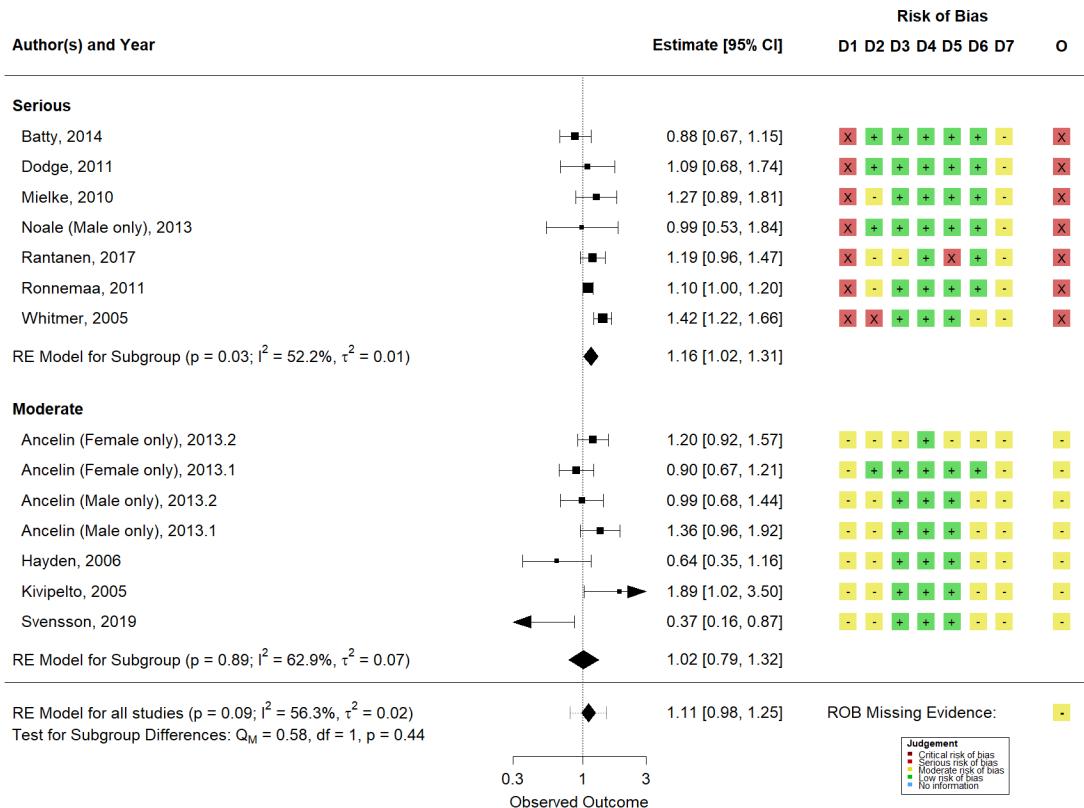


Figure 3.6: Random-effects meta-analysis of non-randomised studies examining the effect of hypercholesterolemia on all-cause dementia

Several studies analysed individual lipid fractions by estimating the risk of dementia

per 1 standard deviation increase in that fraction (Figure 3.7). Very weak evidence

for an effect on all-cause dementia was found for total cholesterol ($N = 5$; HR: 0.97,

95%CI: 0.87-1.07), LDL-c ($N = 2$; HR: 0.96, 95%CI: 0.86-1.08), HDL-c ($N = 4$; HR:

1.04, 95%CI: 0.96-1.13) and triglycerides ($N = 3$; HR: 0.90, 95%CI: 0.75-1.08).

3.4 - Results



Figure 3.7: Random-effects meta-analysis of four lipid fractions (total cholesterol, HDL, LDL, and triglycerides) on all-cause dementia risk, standardised per 1-SD increase in the lipid fraction.

Finally, there were no identified Mendelian randomisation analysis examining the effect of lower lipid levels, as determined by any genetic instrument, on all-cause dementia risk.

3.4.7 Alzheimer's disease

Statins

There were no randomised trials of the use of statins or any other lipid regulating agents on Alzheimer's disease, though several observational studies reported on this outcome and provided evidence for a protective effect ($N = 11$; HR: 0.83, 95%CI: 0.69-1.00; Figure 3.8).

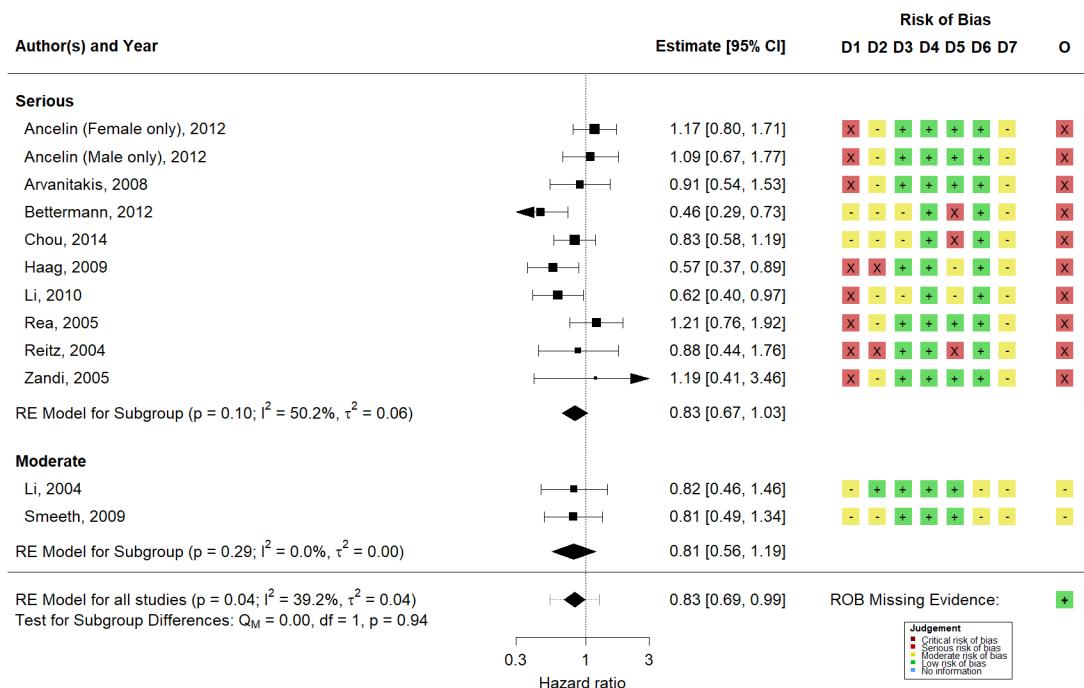


Figure 3.8: Random-effects meta-analysis of non-randomised studies examining the effect of statin use on Alzheimer's disease

Two Mendelian randomisation studies looked at specifically as a result of HMGCR inhibition, mediated by SNPs in the gene (rs172338484 and rs12916).^{191,232} The

3.4 - Results

first used a one sample approach (SNP-exposure and SNP-outcome associations are estimated using the same dataset) in a large Copenhagen-based cohort, while the second made use of summary level data obtained from the Global Lipids Genetic Consortium (SNP-exposure) and the International Genomics of Alzheimer's Project (SNP-outcome). Meta-analysis of these estimates provided weak evidence of an effect (RR: 0.76, 95%CI: 0.51-1.14, Figure 3.9).

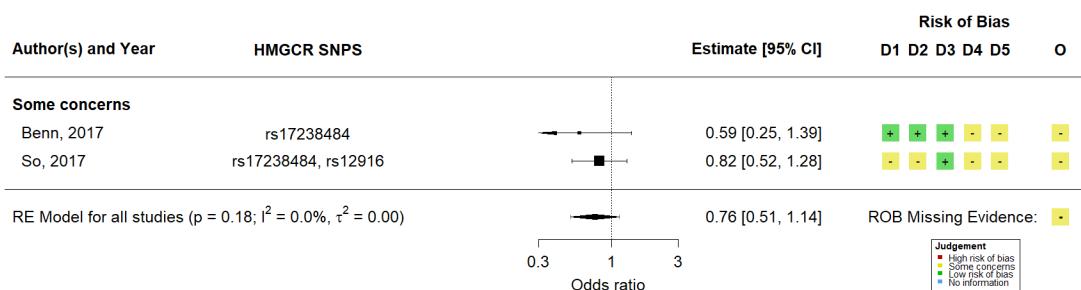


Figure 3.9: Random-effects meta-analysis of genetically lowered LDL-c via HMGCR inhibition on Alzheimer's disease

Lipids

9 studies reported on the association of hypercholesterolemia with all-cause dementia and provided weak evidence for an effect (HR: 0.99, 95%CI: 0.78-1.25, Figure 3.10)

3.4 - Results

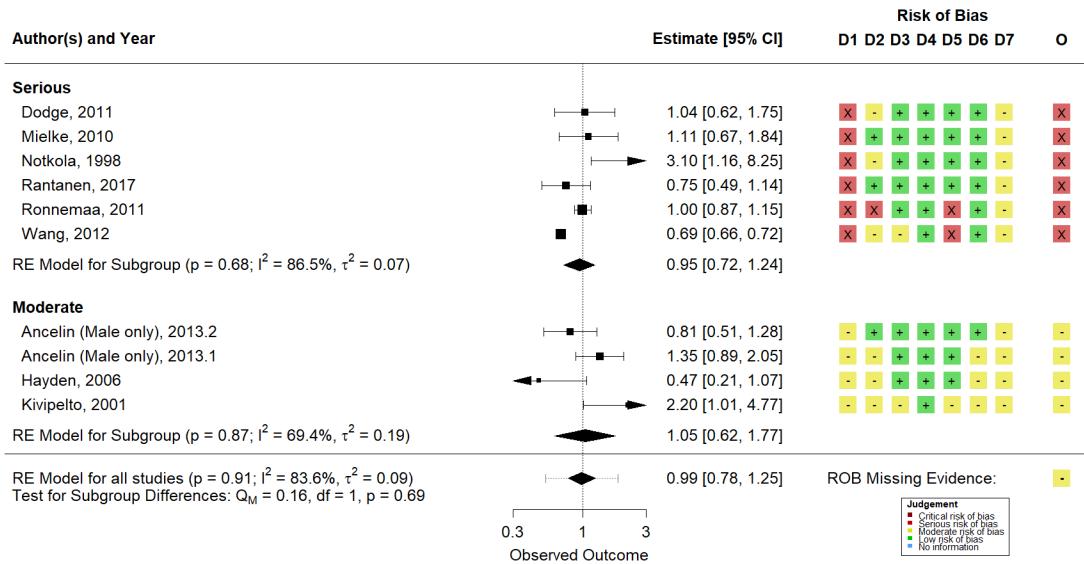


Figure 3.10: Random-effects meta-analysis of non-randomised studies examining the effect of hypercholesterolemia on Alzheimer's disease

Similarly to all-cause dementia, several studies analysed individual lipid fractions

by estimating the risk of dementia per 1 standard deviation increase in that fraction

(Figure 3.11). Very weak evidence for an effect on all-cause dementia was found

for total cholesterol ($N = 5$; HR: 1.00, 95%CI: 0.93-1.07), LDL-c ($N = 3$; HR: 1.03,

95%CI: 0.91-1.16), HDL-c ($N = 4$; HR: 0.99, 95%CI: 0.91-1.07) or triglycerides

($N = 3$; HR: 0.99, 95%CI: 0.83-1.19).

3.4 - Results

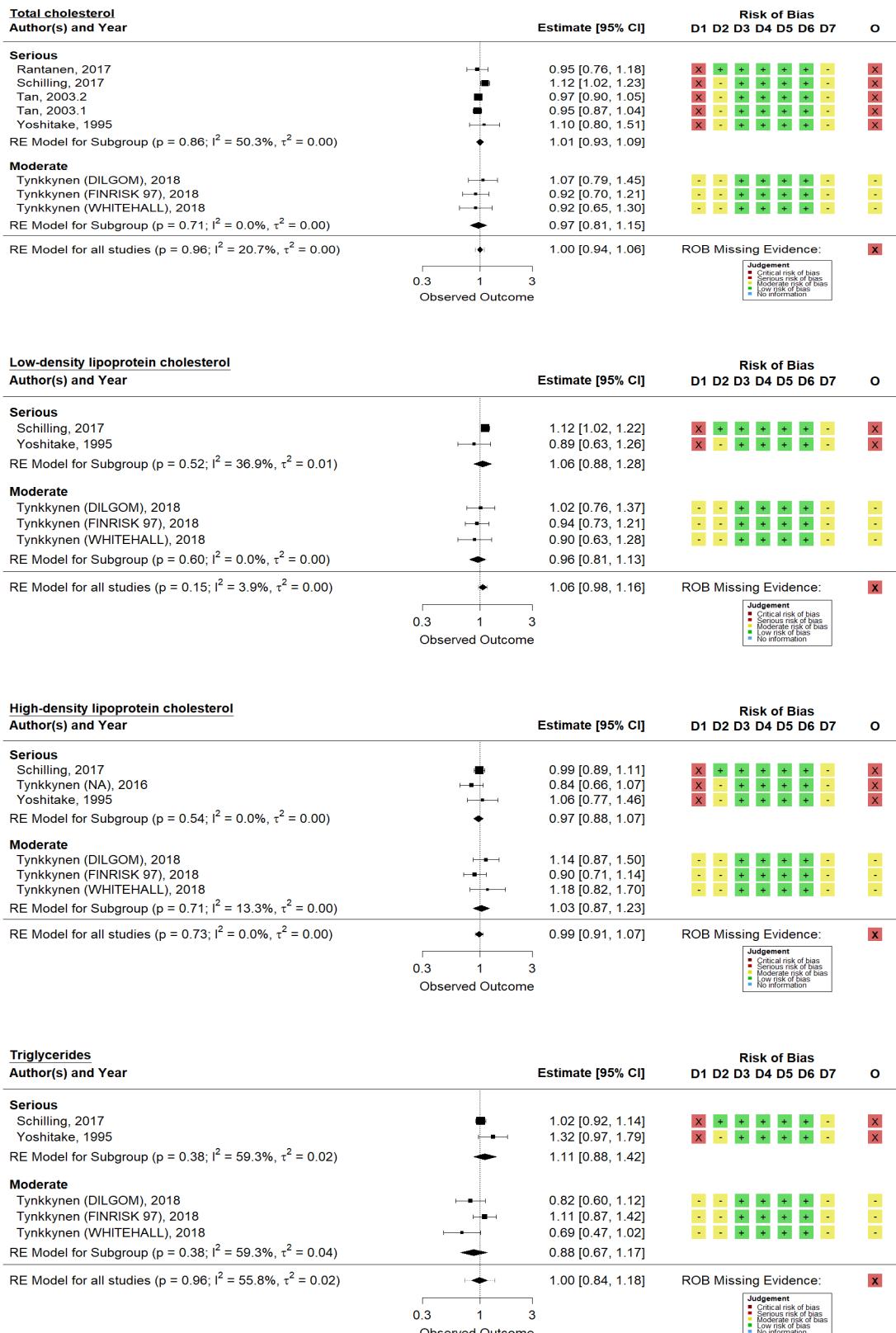


Figure 3.11: Random-effects meta-analysis of four lipid fractions (total cholesterol, HDL, LDL, and triglycerides) on Alzheimer's disease risk, standardised per 1-SD increase in the lipid fraction.

Finally, there were several identified Mendelian randomisation studies examining the effect of genetically lowered LDL-c on Alzheimer's disease risk (Figure 3.12). However, all of these studies used a two-sample approach, making use of summary statistics from the GLGC and IGAP consortia. Due to this overlap, which would result in a falsely precise estimate caused by multiple counting of the same participants, no meta-analysis of these studies was performed. However, across all analysis, using varying number of SNPs, no evidence for an effect was observed (Figure 3.12).

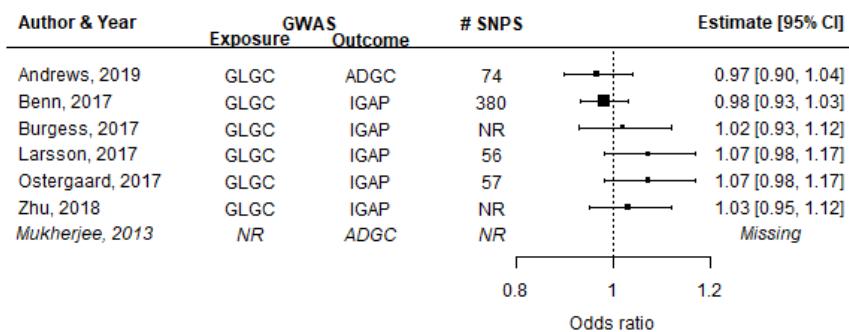


Figure 3.12: Summary of duplication across Mendelian randomisation studies which used summary statistics from the Global Lipid Genetics Consortium (GLGC) and the International Genomics of Alzheimer's Project (IGAP). Note that the Alzheimer's Disease Genetics Consortium (ADGC) is a sub-cohort within IGAP.

3.4.8 Vascular dementia

Statins

As noted in Section 3.4.4 above, there was substantially less literature available on the association of my risk factors of interest and vascular dementia. There were no available randomised trials for this outcome. Three prospective cohort studies examined statin use and vascular dementia, though meta-analysis of these

3.4 - Results

studies provided very weak evidence for an effect ($N = 3$; HR: 1.02, 95%CI: 0.77-1.36; Figure 3.13).

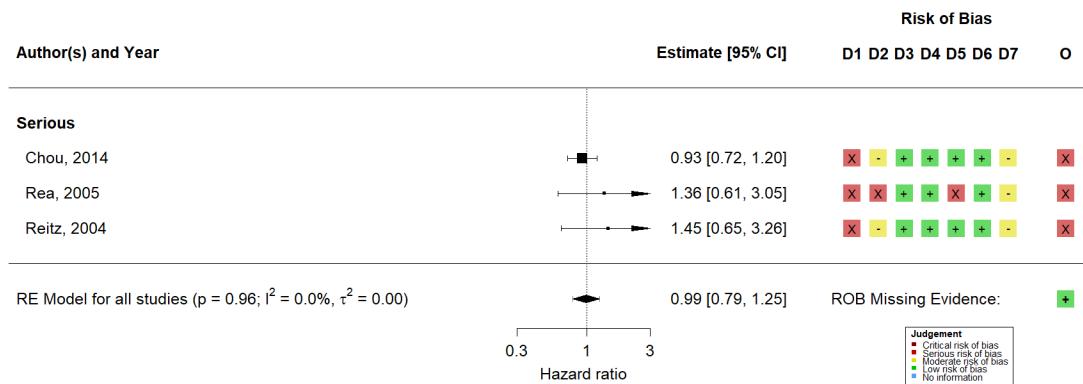


Figure 3.13: Random-effects meta-analysis of non-randomised studies examining effect of statin use on vascular dementia

A single Mendelian randomisation analysis was identified that examining the effect of genetically lowered LDL-c levels via HMGCR inhibition on the risk of vascular dementia, which provided weak evidence for an effect (RR: 0.44, 95%CI: 0.21-0.91).

Lipids

Three studies reported on the association of hypercholesterolemia with vascular dementia and provided weak evidence for an effect (HR: 1.20, 95%CI: 1.00-1.44, Figure 3.14)

3.4 - Results

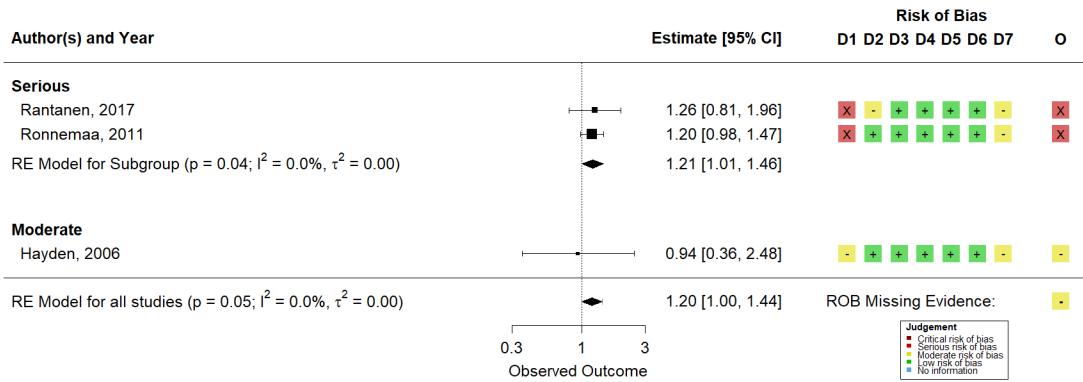


Figure 3.14: Random-effects meta-analysis of non-randomised studies examining the effect of hypercholesterolemia on vascular dementia

Few studies investigated the effect of individual lipid fractions. Only for total cholesterol was there greater than one result reported (Figure 3.15), and a meta-analysis of these found very weak evidence for an effect ($N = 2$; HR: 1.05, 95%CI: 0.79-1.41). A single study provided evidence on the other three fractions,²²⁵ and similar found minimal evidence of an effect LDL-c ($N = 1$; HR: 1.12, 95%CI: 0.83-1.51), HDL-c ($N = 1$; HR: 0.83, 95%CI: 0.60-1.14) or triglycerides ($N = 1$; HR: 1.00, 95%CI: 0.75-1.34).

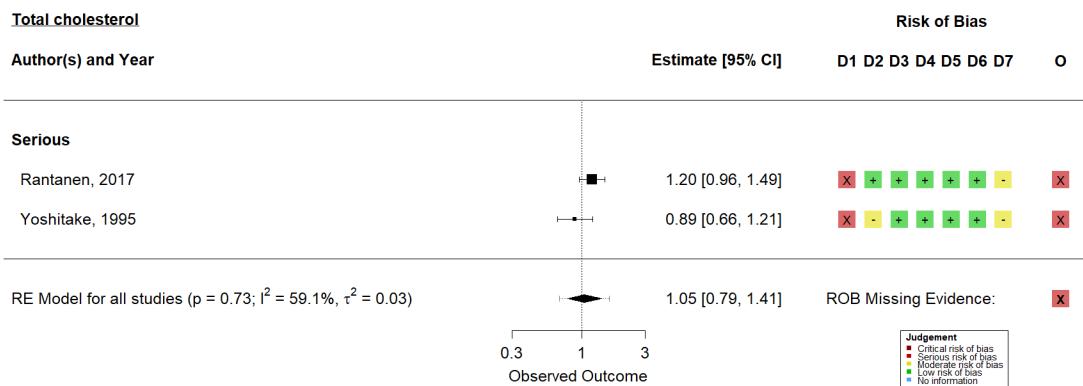


Figure 3.15: Random-effects meta-analysis of four lipid fractions (total cholesterol, HDL, LDL, and triglycerides) on Alzheimer's disease risk, standardised per 1-SD increase in the lipid fraction.

Finally, there were no identified Mendelian randomisation analysis examining the

effect of genetically determined lipid levels on vascular dementia risk.

3.4.9 Dose response meta-analysis of lipid levels

There were 13 studies that were eligible for dose-response meta-analysis as they provide several categories of lipid exposure. However, following data extract, five were excluded as they did not reported the relevant information needed (most commonly, the cut-off measures for each category).

Across the remaining eight studies, a sufficient number of studies ($n \geq 3$) were identified only for the total cholesterol-Alzheimer's, LDL-Alzheimer's, and total cholesterol-dementia strata. This analysis provided very weak evidence for a non-linear effect of lipid levels on dementia outcomes, and Figure 3.16 illustrates this for the total cholesterol-Alzheimer's strata. Similar figures for the other analysed lipid-outcome strata are presented in Appendix ??.

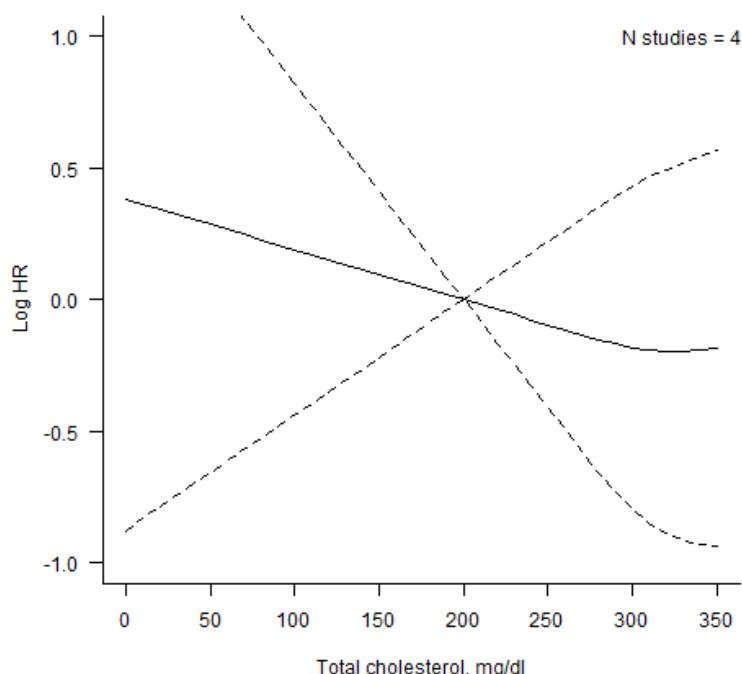


Figure 3.16: Dose-response meta-analysis of total cholesterol on Alzheimer's disease

3.4.10 Additional analyses

Investigation of potential sources of heterogeneity was complicated by two factors. In the first instance, few meta-analysis included more than 10 results, the recommended minimum required for meta-regression. Secondly, poor reporting of summary statistics including education level and baseline cognitive ability precluded the use of several results in a meta-regression analysis.

Age and sex were assessed as potential causes of heterogeneity in the meta-analyses of statin use and hypercholesterolemia on all cause dementia and Alzheimer's disease, but I found very weak evidence for variation in the observed effect estimates due to these factors.

Similarly, assessment of small-study effects, for which publication bias could be one potential reason, was hindered by the relatively small number of results included in a given meta-analysis. However, all analyses assessed provided weak evidence of small-study effects.

[Note: Julian, is this sufficient or should I report statistics for each? Additionally, are there any other sensitivity analyses I should be presenting here?]

3.4.11 Missing evidence

The risk of bias due to missing evidence in each synthesis is shown beside the overall summary diamond in each forest plot presented above. For randomised controlled trials and non-randomised studies of interventions, the risk of bias due to missing evidence was assessed to be minimal. However, there was substantial evidence that results were selectively reported in studies examining the effect of lipid fractions on dementia outcomes (Figures 3.7, 3.11 and 3.15).

3.4.12 Added evidential value of including preprints

As show in Figure 3.1, the number of hits returned by the preprint searching was not substantial (bioRxiv = 256, medRxiv = 0). From these hits, three preprinted reports of eligible studies were included in the review, of which two described unique studies not captured by the main search.^{232,235}

Including preprints did provide useful additional evidential value in a number of meta-analysis. To demonstrate this, the effect of HMGCR SNPs on Alzheimer's disease (Figure 3.9) in Mendelian randomisation studies was re-analysed using a fixed-effect model. Examination of the weight assigned to each result in the analysis illustrates that a large proportion (78%) of the weight is given to the preprinted result.

Investigation of the publication status of the two preprints reporting studies not also identified by the main search after a two-year lag found that one had been subsequently published in late 2019.^{227,235} The final preprint has not yet been published.²³²

3.5 Discussion

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. This discussion seeks to summarise the key findings in terms of literature sources and results as reported. A detailed comparison across the evidence sources, exposure measures and sources of bias reported here is presented as part of the triangulation exercise in Chapter 7.

3.5.1 Summary of findings

There was some evidence of protective effect of statins on all-cause and Alzheimer's disease dementia when looking at solely at observational studies. This finding was not supported by evidence from the two available RCTs, or by studies that emulated statin treatment using a genetic proxy, suggesting that these findings may be a result of heterogeneity in exposure (e.g. mid-life in studies of lipids with late-life lipid reduction in RCTs) or alternatively due to biases within the non-randomised studies.

The majority of studies were non-randomised studies of lipids, or treatments that affect lipid levels such as statins. This distribution of evidence between analytical designs is to be expected. Randomised controlled trials of dementia are particularly challenging, as the long follow-up made necessary by the long latent period of the condition, makes trials logistically challenging and financial expensive. Similarly, Mendelian randomisation is a comparatively new study design (as illustrated in Figure 3.17), which appears in the literature in recent years, driven by the availability of summary genome wide association studies (GWAS) that form the basis of the two-sample Mendelian randomisation approach.

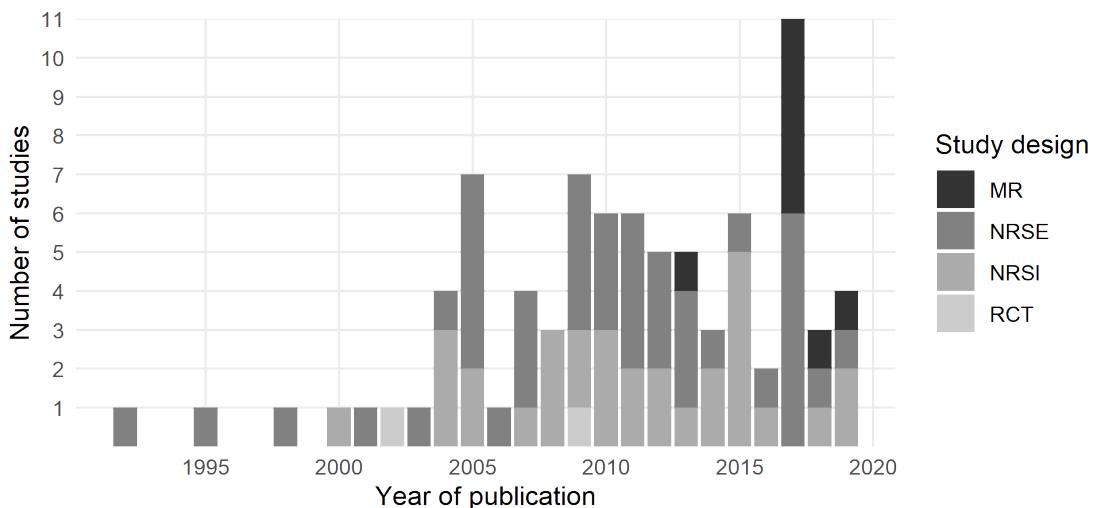


Figure 3.17: Included study designs by year of publication

3.5 - Discussion

A common theme was an absence of studies examining vascular dementia as an outcome, most noticeable when comparing the evidence base for statins in dementia/Alzheimer's (Figures 3.4 & 3.8) with that available for vascular dementia (Figure 3.13). This is particularly interesting given that lipids and statins are strongly related to the prevention of vascular disease. A potential explanation for this observation may be publication bias or the "file-drawer effect",⁷⁶ though there was very weak evidence of a small-study effects for this outcome (of which publication bias is a potential cause).²³⁶ Similarly, only one Mendelian randomisation studies examined this outcome, primarily because of the absence (until recently) of vascular dementia GWAS which precludes a two-sample approach.

Of note, this review did not include the commonly cited PROSPER RCT, which examined the effect of pravastatin on CVD risk, reporting on cognitive outcomes as one of several secondary outcomes.²³⁷ While widely cited in relation to the effect of statins on dementia risk and included in the Cochrane review of RCTs on this topic,²³⁸ the trial reported solely on the change in a range of cognitive measures (MMSE, Stroop test, Picture-Word Learning test and others) over follow-up. Though an useful indicator of general cognitive decline, it is not equivalent to a dementia diagnosis using recognised criteria as cognitive tests should feed into a broader diagnostic pathway (see Section 1.3.2). As such, this trial did not met the inclusion criteria for this review.

Risk of bias across the individual results was generally quite high, and the causes and expected directions of these biases are discussed in more detail in Chapter 7. Of particular interest to this chapter, however, was the high risk of bias due to missing evidence observed for observational studies of statin levels. In many cases, estimates were know to be missing from meta-analyses not at random due to preferential reporting of significant results observed in a number of analysis, leading to the high-risk judgement. These missing estimates were most commonly identified via analysis of conference abstract/final publication pairs.^{239,240} In addition, some authors stated outright that non-significant results were not reported (e.g. "The other lipid variables not significantly associated with dementia and Alzheimer's

disease ... were not reported in the Table.”).¹⁸⁹ However, as all identified missing results are likely to be non-significant, they would not be expected to have a substantial impact on their respective meta-analysis (which provided very weak evidence for an effect), other than increasing the precision of the summary estimate.

Finally in terms of generalisability, despite a large proportion of included studies being conducted in the Western world (Figure 3.2), the applicability of the results to other populations is aided by the inclusion of studies which made use of data from the Taiwan health insurance database.

3.5.2 Comparison with previous reviews

While conducting this review, I identified several previous systematic reviews of this topic.^{57,241–244,244} However, this review is the first to use established domain based assessments tools (for example, the RoB 2 tool for randomized controlled trials)¹⁴⁵ to assess the risk of bias in included studies. The majority of the highly cited reviews on this topic either do not formally consider risk of bias in the observational studies they include^{57,245} or used a non-domain-based assessment tool (e.g. the Newcastle-Ottowa Scale).²⁴³

I identified one previous review of Mendelian randomisation studies examining risk factors for Alzheimer’s disease. However, this review was conducted prior to the majority of Mendelian randomisation studies included in this review being published and extracted results including SNPs in the APoE4 genetic region (see the following section for a discussion of the bias this introduces).

Despite these differences in time-scales and 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.^{246,247} This approach would have enabled more efficient identification of relevant primary

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

3.5.3 Inclusion of Mendelian randomisations studies

One of the particular strengths of this review is the inclusion and critical assessment of Mendelian randomisation studies as a source of evidence.

Mendelian randomisation is a powerful analytical technique, using natural variation in participants genomes to identify causal links between a genetically determined risk factor and an outcome, given that the core assumptions detailed in Figure 3.18 are valid. However, inclusion of Mendelian randomisation as an acceptable study design in this review was complicated by a number of factors.

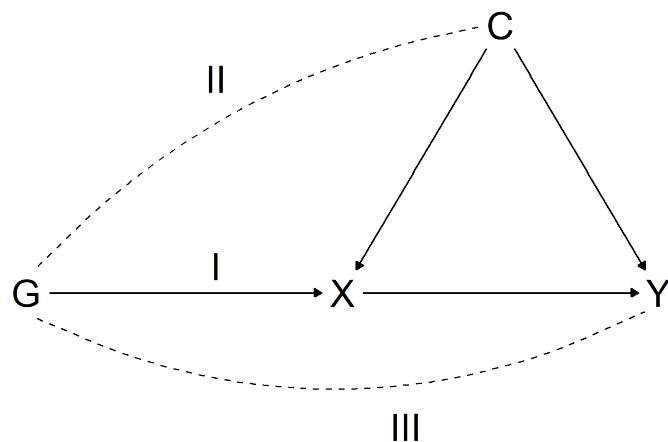


Figure 3.18: Summary of the assumptions in Mendelian randomisation analyses: (I) *relevance* - the genetic variant associates with the risk factor of interest; (II) *independence* - the variant-exposure association has no unmeasured confounders; and (III) *exclusion restriction* - variants affect the outcome only through their effect on the risk factor of interest (i.e. there is no horizontal pleiotropy).

Firstly, this study design is relatively new, particularly when compared to randomised trials or cohort studies. Figure 3.17 demonstrates that Mendelian randomisation studies only begin to appear in the evidence base much later than NRSE/NRSI,

3.5 - Discussion

likely due to the limited availability of large scale GWAS datasets needed for two-sample Mendelian randomisation analyses. As such, the process and tools for systematically assessing this study design 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 its application, which lead some including general terms for instrumental variable analyses in my search.

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 3.3.7. 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.²⁴⁸ While this tool assesses the underlying GWAS used, it does not assess the additional methodological considerations of the Mendelian randomisation analysis itself. For this review, I utilised the best available author-devised tool, sourced from 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.²⁴⁹ On review of this paper, the search would not have been expected to find it given the absence of any lipid-related 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,

3.5 - Discussion

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.^{250,251} 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 for this study design by a panel of methodologists and analysts would be of substantial benefit to the field.

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 introduction (see Section 1.5.1), an increasing number of ApoE4 alleles is a major independent risk factor for Alzheimer’s disease, and so violates the exclusion restriction criteria (Figure 3.18). In all cases, excluding these variants attenuates the observed effect to the null. A clear example of this is Benn *et al* (2017), where the ApoE variants were not sufficient identified and excluded and the published paper detailed evidence for a protective effect of LDL-c was identified (RR: 0.83, 95%CI: 0.75-0.92).⁶⁶ Following several rapid responses, the data was re-analysed excluding a larger area around ApoE4 which attenuated the finding to the null.²⁵²

3.5.4 Inclusion of preprints

As highlighted in Section 1.6.1, this review explicitly sought to synthesize evidence

across different publication statuses (preprinted vs. published). Using the tool

described in Chapter 2.2, two preprint serves related to health and biomedical

sciences were search as part of this review. The small number of studies returned

by the searches (or the absence of any relevant hits in the medRxiv database - see

Figure 3.1) is due to the timing of the preprint searches. The searches for this

review were performed in mid-July 2019, but the medRxiv repository, an offshoot of

the Epidemiology and Clinical Trials categories of the bioRxiv preprint server, only

registered its first preprint 25th June 2019. As such, at the point it was searched,

the medRxiv database contained only a very small number of records (n=148).

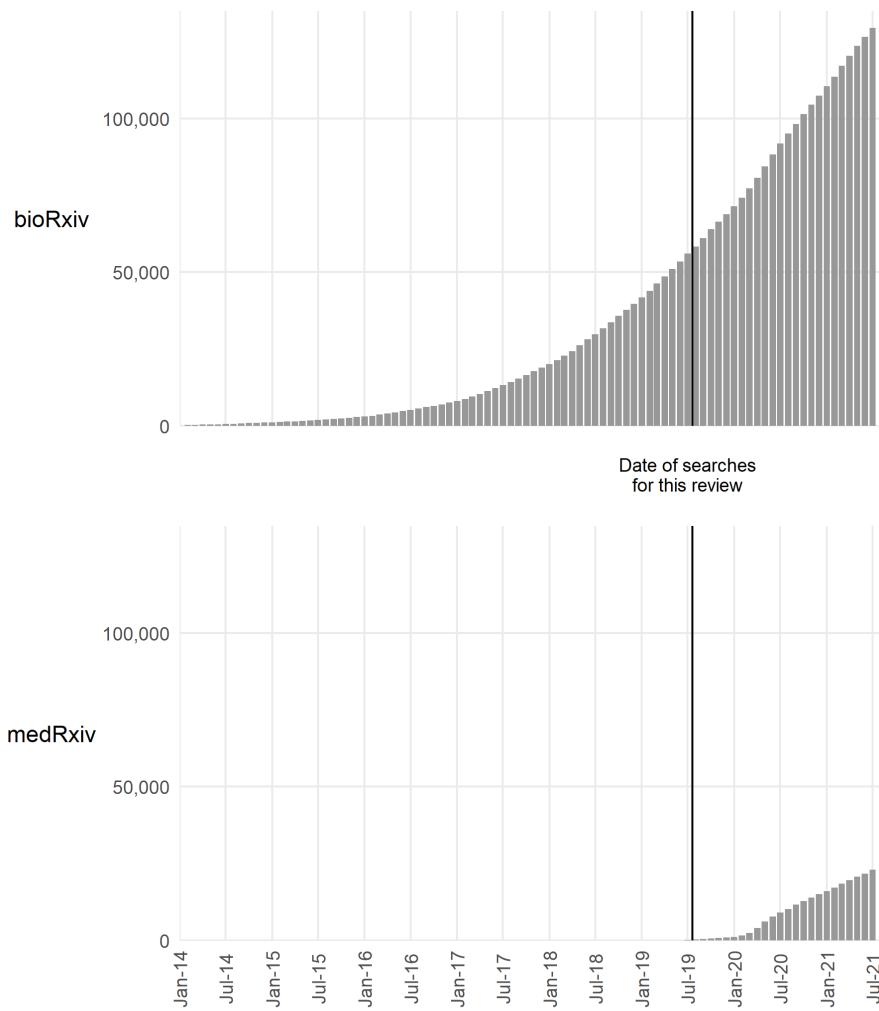


Figure 3.19: Growth of preprint repositories over time. Given the relative sizes of the preprint repositories at the time the searches for this review were conducted (bioRxiv n= 56,007, medRxiv n = 148), the number of hits returned by each is expected.

Three relevant preprints from the bioRxiv hits were identified. The added evidential value of including these preprints was described in Section 3.4.12, and indicated that results available only via preprinted reports can contribute substantially to a meta-analysis. Of note, all three included preprints described Mendelian randomisation studies, potentially indicating that more biologically-focused study designs are over-represented in the bioRxiv repository.

Of the three identified preprints, two were subsequently published as of September 2020. This fits well with the analysis presented in Chapter 2 that, allowing for a

two-year lag, approximately two-thirds of preprints are published, and also nicely illustrates the dual advantages of preprinted reports to evidence syntheses.

Firstly, preprints provide an advance snapshot of the literature, capturing articles that will eventually be published but were not available at the time of the main search. Consider the example of one eligible preprint in this review was initially posted on bioRxiv in July 2019²⁵³ and was subsequently published in 2021 following peer-review.²²⁷ Secondly, inclusion of preprints allows for results that may never be formally published to be included in an evidence synthesis exercise as is the case with the second preprint included in this review.²³² Both of these aspects illustrate that, 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.

More recently, inclusion of preprints in systematic reviews has become significantly more widespread. This is largely 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.⁷⁵ How well this adoption of preprints will transfer to other less-urgent topics, where the speed of research does not put the same focus on preprinted articles, is currently unknown.

3.5.5 Open data sharing

As discussed in Section 3.4.9, many primary studies did not report important information required for the dose-response meta-analysis, and so could not be included in the synthesis. This limitation was compounded by the expected low response rate to requests for further information from primary authors. While contacting authors is worthwhile, as it can substantially change the conclusion of a systematic review²⁵⁴ and is not too costly to systematic reviewers,²⁵⁵ a far preferable option is that the authors of primary studies readily deposit all relevant study data at the point of publication.

3.5 - Discussion

Based on my experience of extracting data for this review, I co-authored a 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. A copy of this publication is available in Appendix B.4.²⁵⁶ In an attempt to apply my own guidance, I have invested a substantial amount of time and effort into making the data obtained by this review openly available to other researchers. [Zenodo repository will be cited here]

3.5.6 Strengths and limitations

Strengths

I believe there are several aspects where this review is distinct from those reviews already available in the published literature. While several reviews of this research topic exist,^{57,241–243} 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, this review employed a structured approach to risk-of-bias assessment using a domain-based tool. This repres important strength of this review, as the detailed risk of bias assessments are used in

Thirdly, as discussed at length in the section above, in contrast to other available reviews and enabled by the tool described in Chapter 2, this review systematically searched preprinted health-related preprints. Finally, as a secondary element, I used this review to pilot new research synthesis methodologies, in particular a new visualisation approach for risk-of-bias assessments and a forthcoming tool for assessing the risk-of-bias due to missing evidence.

Limitations

The primary limitation of this review is that several included studies used data from EHR databases, which come with serious concerns regarding validity^{257 258,259}. 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 this 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 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.²⁶⁰

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.

3.6 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 preceding 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).

3.7 References

— Archibald Cochrane, 2000¹²⁵ — Archibald Cochrane, 2000¹²⁵

4

Systematic review of all available evidence
on the association between blood lipids
and dementia outcomes

5

Primary analysis of lipid-regulating agents and dementia outcomes

Contents

5.1	Lay summary	111
5.2	Introduction	112
5.3	Methods	113
5.4	Results	127
5.5	Discussion	143
5.6	References	152

5.1 Lay summary

Electronic health record (EHR) databases are large collections of medical 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 standardised coding system. 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 England. Using this data, the analysis presented in this chapter examined whether treatments which lower cholesterol levels (also known as lipid-regulating agents or LRA) of which statins are a prime example, affect the risk of all-cause dementia and related outcomes (Alzheimer's disease, vascular dementia and other dementias).

Little 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 lipi-regulatin agent called fibrates. In contrast, I found an increased risk of vascular and other (i.e. non-Alzheimer's) dementia was associated with lipid-regulating agent use.

This increased risk in outcomes with a vascular element (e.g. vascular dementia) 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, which makes it appear as if statins are harmful. 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 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 through use of a Cox proportional hazards analysis, incorporating a time-varying treatment indicator.²⁶² This approach provides a evidence source at risk of a distinct bias due to the alternative analytical strategy. The results from this analysis will be incorporated into the triangulation exercise presented in Chapter 8.

This chapter represents an extended version of a preprint manuscript, a copy of which is available in Appendix B.4.

5.3 Methods

5.3.1 Study protocol

An *a priori* protocol for this study was published,²⁶³ and amendments to this are recorded in Appendix A.4.1.²⁶⁴

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.²⁶⁵ The database has been collecting primary care data from

5.3 - Methods

participating practices across England since 1987.^{266,267} 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.^{265,268}

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.²⁶⁹ 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

The index events, exposures and outcomes used in this analysis 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).

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 cut-off were included in the original CPRD extract obtained for this analysis. However, these were excluded from the analysis as data quality and reliability are thought to be higher after this date.²⁷⁰ 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.²⁶³ The blood lipid cut-offs were based on NIHR-recommended levels at the time the protocol was written. 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.

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: (a) an outcome of interest; (b) death; (c) end of follow-up (29 February 2016); (d) last registration date with their GP practice; or (e) the last CPRD collection date for their practice. Participants were ineligible for our cohort if they were less than 40 years of age (as these patients are less likely to be prescribed a LRA), 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),²⁷¹ namely: statins, fibrates, bile acid sequestrants, ezetimibe, nicotinic acid groups, ezetimibe and statin (representing one treatment containing both drugs, rather than the two classes being prescribed concurrently), and omega-3 fatty acid groups.

A participants 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 dementias, 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 (Figure 5.1). The diagnosis date of the outcome was determined by the first record of a relevant code.

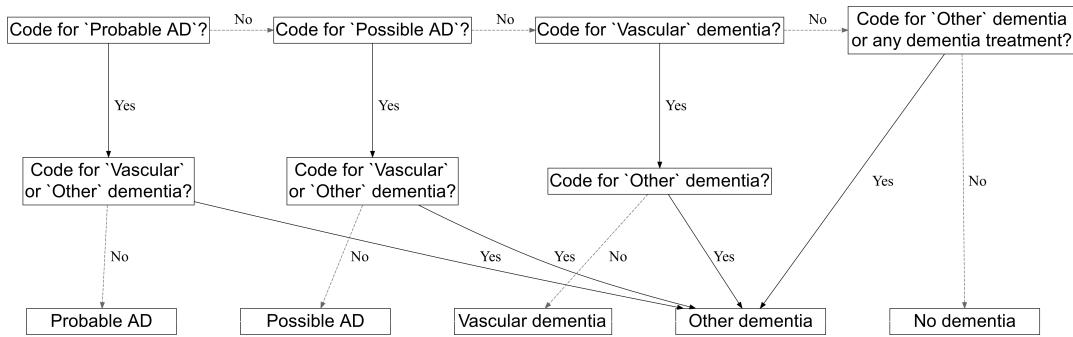


Figure 5.1: Decision tree for assigning dementia subtypes, based on the presence of Read codes in the patient's record. Note that an outcome of "probable" or "Possible" Alzheimer's disease (AD) requires the absence of any vascular outcome codes.

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. These are discussed in detail below and summarised in Table 8.1.

Charlson index implemented using Read code lists.²⁷³ Code lists based on those by Taylor et al.²⁷⁴

5.3 - Methods

Table 5.2: Definition of covariates adjusted for in the full-adjusted model. The codelists used to define the majority of these covariates were originally created for use in a previously published analysis,²⁷² while others were built on or adapted from previous published work.^{273–275}

Covariate	How was the covariate defined?
Previous history of coronary arterial disease	Presence of one or more relevant Read codes on record.
Previous history of coronary bypass surgery	Presence of one or more relevant Read codes on record.
Previous history of cerebrovascular disease (including stroke)	Presence of one or more relevant Read codes on record.
Chronic illness, including cancer and arthritis	Charlson index implemented using Read code lists. ²⁷³ Code lists based on those by Taylor et al. ²⁷⁴
Socioeconomic position	2010 English Index of Multiple Deprivation (IMD) at the twentile level, where 1 represents the least deprived and 20 the most deprived.
Consultation rate	Calculated by dividing the total number of clinic visits by the length of the patient record prior to the index date to give an average annual rate.
Alcohol status	Recorded value (current, former or never).
Smoking status	Most recent of recorded value (current, former or never) or Read code indicating a recorded value. Code lists based on those by Wright et al. ²⁷⁵
Body Mass Index	Recorded value if available, or a calculated value using the last recorded height and weight measurements. Measurements taken before the age of 25 were excluded to ensure adult measurements were used.
Peripheral arterial disease	Presence of one or more relevant Read codes on record.
Hypertension	Presence of one or more relevant Read codes on record.
Baseline total cholesterol	Continuous value recorded as test result ("enttype==163 & test_data1==3")
Baseline LDL cholesterol	Continuous value recorded as test result ("enttype==177 & test_data1==3")
Chronic kidney disease	Presence of one or more relevant Read codes on record.
Type 1 Diabetes	Presence of one or more relevant Read codes on record.
Type 2 Diabetes	Presence of one or more relevant Read codes on record.

Demographic covariates adjusted for included age and gender. Age was calculated at date of entry into the cohort, and adjusted for via its use as the time axis for the Cox model (see Section 5.3.10). Socioeconomic status was proxied using

5.3 - Methods

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 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 , 2001-2005, 2006-2010, > 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.²⁷⁶ 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,²⁷⁷ 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.²⁷⁸ Variables with missing observations were identified for inclusion in the imputation model. 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.²⁷⁹ Using the MICE (Multiple Imputation by Chained Equations) command in STATA16, 20 imputed datasets were created.

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.²⁷⁷

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.

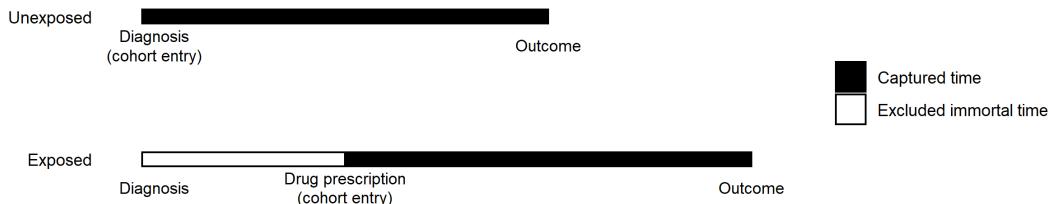
A Cox PR model was chosen for this analysis as it inherently accounts for the length of time participants spend in each exposure group. Using this approach, time-at-risk can be properly attributed to the appropriate exposure group, thus mitigating the impact of immortal time bias. This is discussed in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

5.3.9 Immortal time bias and time-varying treatment indicators

Immortal time bias describes 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.²⁸⁰ Following the example displayed in the figure, the unexposed group are followed from the date of a diagnosis, while the exposed group is followed from date of a prescription. In this scenario, the time between diagnosis and prescription for the exposure group is missing, and any events that occur in the exposed group prior to the prescription 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 prescription, and any events occurring within it, is inappropriately assigned to the exposed group.

A



B

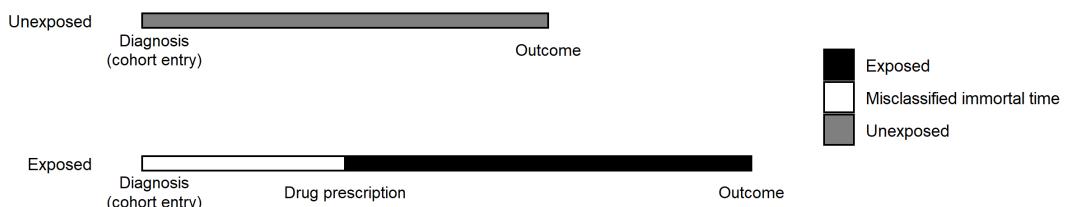


Figure 5.2: Diagram illustrating the two presentations of immortal time bias, as a selection bias (Panel A) and a misclassification bias (Panal B).

This second presentation appears to be common in the existing literature, as several of the studies included in the systematic review presented in Chapter ?? were identified as being at risk of immortal time bias following formal risk of bias assessment using the ROBINS-I tool (see Section ??). Both presentations of immortal time bias appear in the literature on the relationship of statins and dementia, as identified by the systematic review presented in Chapters 3/4

This analysis attempted to address this issue by following all participants from a common index date (defined as earliest of: (a) date of raised cholesterol test results; (b) hypercholesterolemia diagnosis; or (c) LRA prescription). Following a recommended approach to addressing the second form of immortal time bias, I employed a time-varying indicator of treatment status to correctly allocate time-at-risk to the exposed and unexposed groups.²⁸⁰

Under this approach, all patients start in the unexposed group and 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).

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.^{281–283} A model using age as the time axis inherently accounts, or adjusts, for participants age as a potential confounder of the exposure-outcome relationship. As such, 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.

Complete case vs imputed data

Using multiple imputation to handle missing data is an alternative to a “complete case” approach,²⁸⁴ where participants missing any covariate are dropped from the dataset. As a recommended sensitivity analysis,²⁸⁵ I preformed and compared the results of both methods, to investigate the impact of multiple imputation on the results.

Control outcomes

In addition to the primary outcomes of interest (described in Section 5.3.5), I extracted data on three additional control outcomes. The inclusion of control outcomes in observational analyses are a useful technique to assess the strength of uncontrolled confounding,²⁸⁶ and these outcomes are usually class as either “negative” or “positive” outcomes.

Negative outcomes are defined as those without a likely causal path between the exposure and outcome (see Figure 5.3 for a directed acyclic graph, or DAG, 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, I chose three control outcomes were chosen in reference to this drug class: back pain (negative control), ischaemic heart disease (positive protective control), and Type 2 diabetes (positive harmful control).

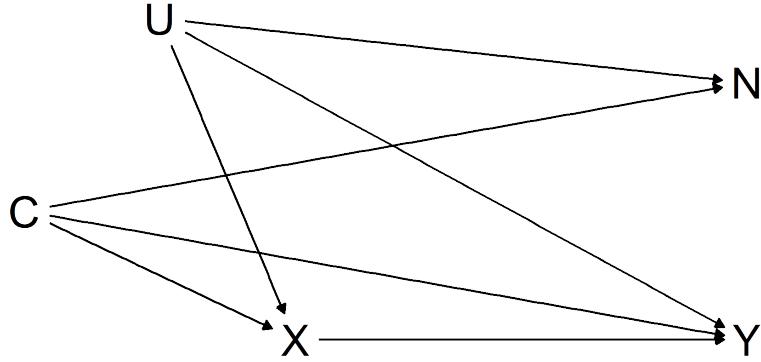


Figure 5.3: Causal diagram (directed acyclic graph) showing relationship between exposure X , outcome Y , confounders (measured C and unmeasured U) and an ideal negative outcome N . Note the absence of any arrow between X and N . In this scenario, any association observed between X and N is due to the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

Despite observational analyses suggesting a link between statins and muscular pain (as opposed to more serious complications such as myopathy),²⁸⁷ systematic reviews of the adverse events of statin use³⁰ and N-of-1 trials explicitly exploring the association of statin use with muscle pain²⁸⁸ have found little evidence supporting an effect. This suggests that muscular backpain would be suitable for use as a negative control outcome in this analysis. Under this approach, if statin use is found to be associated with muscular backpain in this analysis, this suggests the presence of residual confounding and reduces my confidence in the results for the dementia outcomes.

Similarly, incident ischemic heart disease and Type 2 diabetes were included as a protective and harmful positive control outcome, respectively. The protective effect of lipid-lowering treatment, via statins, on the risk of ischemic heart disease is well-established,³⁰ while there is growing evidence for an increased risk of Type 2 diabetes with statin use.^{30,289,290} Similar to the negative outcome, if the analysis strategy can reproduce these known associations, this will provide evidence that potential confounders have been sufficiently adjusted for.

Impact of additional covariates

To observe the effect of adjusting for additional covariates, I ran two additional models unadjusted except for: (a) age; and (b) age and gender. The results of these models was then compared the results with the fully 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) to assess the robustness of the estimates, as statins are contraindicated in pregnancy.²⁹¹

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 1.4.2).³³ To investigate this, I further stratified the statin exposure group into lipophilic (Atorvastatin, Lovastatin, Simvastatin, Cerivastatin) and hydrophilic (Pravastatin, Rosuvastatin, Fluvastatin) statins.

Comparing dementia codelists

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 published study by Smeeth *et al.*²⁹²

This previous analysis used a propensity matching approach to estimate the association of statins with a range of outcomes in The Health Improvement Network (THIN) database, a alternative source of English electronic health records which has substantial overlap with the CPRD.²⁹³ The code lists used in this analysis 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

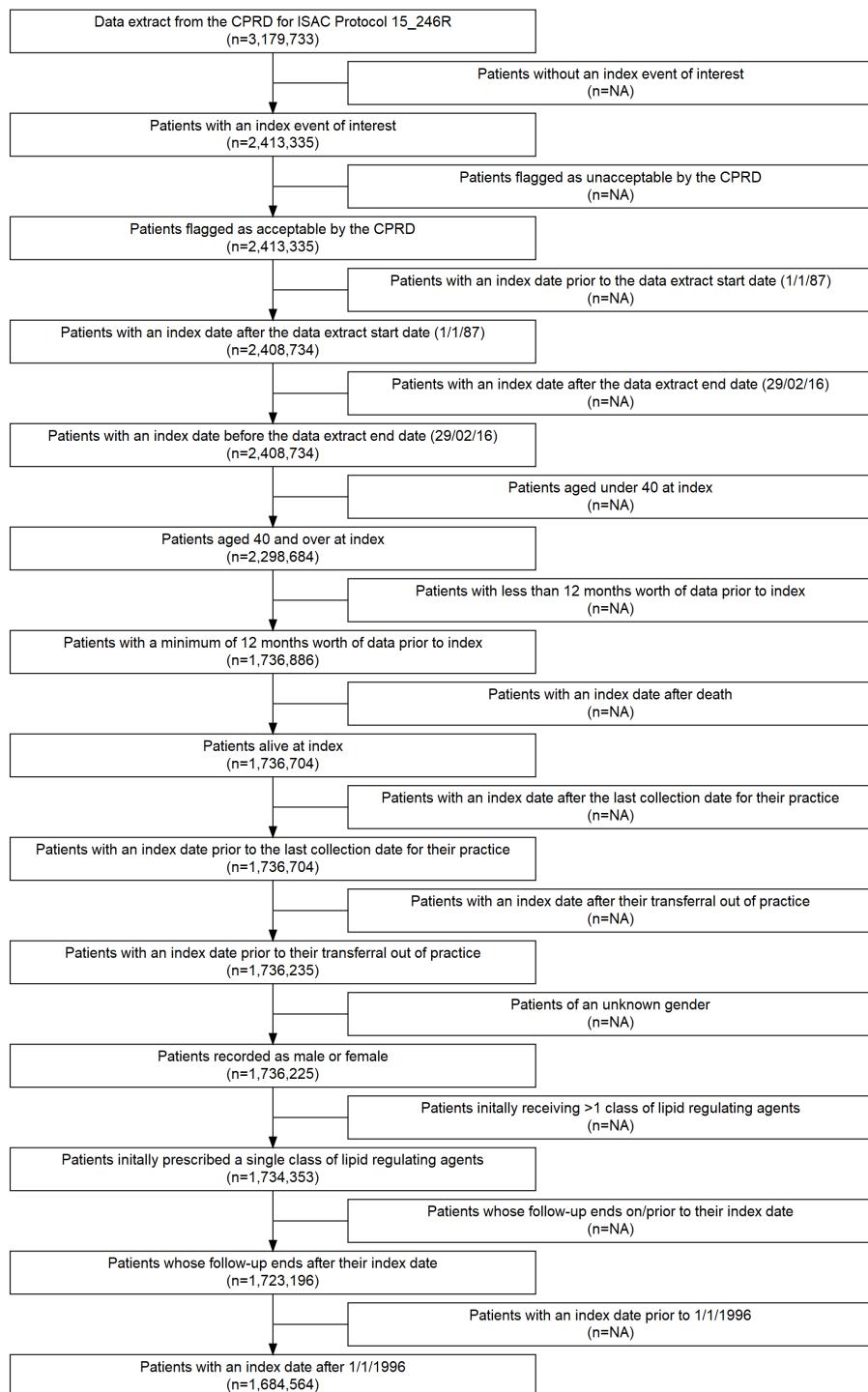


Figure 5.4: Attrition of CPROD participants as the eligibility criteria were applied. The largest cause of attrition was the absence of an index event of interest.

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 dementias).

The number of events, time-at-risk and crude rates for each drug class, tabulated by dementia outcome, are shown in Table 5.3. A substantial majority (98.1%) of participants prescribed a lipid-regulating agent were prescribed a statin. I excluded the “Ezetimibe and statins” (n=127) and “Nicotinic acid groups” (n= 165) classes from subsequent class-based subgroup analyses based on the extremely small number of participants in these groups. 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.

5.4 - Results

Table 5.3: Summary of number of events, years at risk and crude rates per 100,000 participant-years-at-risk stratified by dementia outcome and drug class of interest.

Exposure Group	Any dementia			Possible AD			Probable AD			Vascular dementia			Other dementia		
	Events	PYAR	Rate *	Events	PYAR	Rate *	Events	PYAR	Rate *	Events	PYAR	Rate *	Events	PYAR	Rate *
No LRA (unexposed)	18,608	5,872,717	317	6,368	5,818,047	109	2,637	5,800,964	45	4,813	5,811,594	83	4,790	5,808,285	82
By drug class															
Statins	22,920	4,871,568	470	6,190	4,758,526	130	5,773	4,753,437	121	5,871	4,755,258	123	5,086	4,747,237	107
Omega-3 FAs	19	8,034	236	4	7,927	50	7	7,950	88	4	7,938	50	4	7,925	50
Fibrates	141	38,003	371	49	37,102	132	21	36,835	57	36	37,001	97	35	36,983	95
Ezetimibe	32	6,604	485	8	6,429	124	7	6,425	109	12	6,444	186	5	6,393	78
BAS	106	36,370	291	28	35,808	78	19	35,726	53	26	35,768	73	33	35,808	92
Ezetimibe + Statins †	0	986	-	0	986	-	0	986	-	0	986	-	0	986	-
NAG	4	1,403	-	0	1,379	-	2	1,391	-	1	1,389	-	1	1,382	-
Total	41,830	10,835,686	386	12,647	10,666,205	119	8,466	10,643,714	80	10,763	10,656,378	101	9,954	10,644,999	94

* Crude rate per 100,000 participant-years-at-risk

† One treatment containing both drugs, rather than the two classes being prescribed concurrently

Abbreviations: AD - Alzheimer's disease; BAS - Bile acid sequestrants; NAG - Nicotinic acid groups; Omega-3 FGs - Omega-3 Fatty acid groups; PYAR - Participant-years-at-risk.

5.4 - Results

Table 5.4: 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	Fibrates	Omega-3 Fatty Acid Groups
Sample size (N)	1,684,564	1,087,704	585,528	5,396	763	3,889	992
Year of cohort entry (median)	2006	2007	2004	2005	2004	2001	2005
Female	53.0% (893174)	56.2% (610950)	47.1% (276043)	66.4% (3585)	54.5% (416)	38.6% (1500)	52.6% (522)
Age at cohort entry (median)	57	54	62	57	60	58	56
CAD	0.4% (7133)	0.1% (589)	1.1% (6465)	0.1% (6)	0.9% (7)	1.4% (53)	1.3% (13)
CBS	0.3% (5699)	0.1% (682)	0.8% (4926)	0.1% (4)	0.4% (3)	2.0% (78)	0.6% (6)
CVD	2.1% (34899)	1.1% (11619)	3.9% (22977)	1.6% (86)	2.6% (20)	4.4% (170)	1.7% (17)
Charlson (ever > 0)	30.6% (516135)	25.1% (272642)	40.7% (238403)	42.5% (2292)	41.7% (318)	50.8% (1976)	40.4% (401)
IMD-2010 (median)	9	8	9	8	9	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)	7.1 (6.2)	8.0 (8.0)
Alcohol (ever)	85.9% (1447151)	86.6% (941648)	84.7% (496110)	82.8% (4468)	84.0% (641)	82.9% (3223)	82.0% (813)
Smoking (ever)	51.1% (861355)	47.1% (511826)	58.6% (343074)	55.2% (2978)	57.5% (439)	60.2% (2341)	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)	29.0 (5.2)	26.9 (5.5)
PAD	0.7% (12613)	0.4% (4039)	1.4% (8424)	0.9% (47)	0.9% (7)	1.9% (75)	1.0% (10)
Hypertension	16.0% (269804)	11.5% (124604)	24.4% (143101)	12.8% (692)	23.9% (182)	25.8% (1002)	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)	6.4 (5.6)	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)	3.3 (1.8)	3.2 (1.0)
CKD	0.1% (1295)	0.1% (740)	0.1% (545)	0.1% (6)	0.1% (1)	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% (31)	0.1% (1)
Type 2 Diabetes	2.9% (48557)	1.1% (11797)	6.1% (35941)	2.3% (123)	5.4% (41)	15.8% (614)	2.8% (28)

Note: The ‘Nicotinic acid groups’ (n=165) and ‘Ezetimibe and Statins’ (n=127) subgroups are not shown, but are included in the whole sample column

Abbreviations: BMI - Body mass index; CAD - Coronary arterial disease; CBS - Coronary bypass surgery; CKD - Chronic kidney disease; CVD - Index of multiple deprivation; LRA - Lipid regulating agent; PAD - Peripheral arterial disease; SD - Standard deviation.

5.4 - Results

The distribution of baseline characteristics across the remaining seven drug classes can be seen in Table 5.4. Note due to the experimental design, the median year of entry is expected to be later for those not prescribed an LRA, as this exposure group is more likely to include those who entered into the cohort towards the end of study window (as these people had less follow-up time in which to be prescribed an LRA).

The stopping, addition and switching of drug classes was common across all drug classes (Table 5.5).

Table 5.5: 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.

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, 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

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

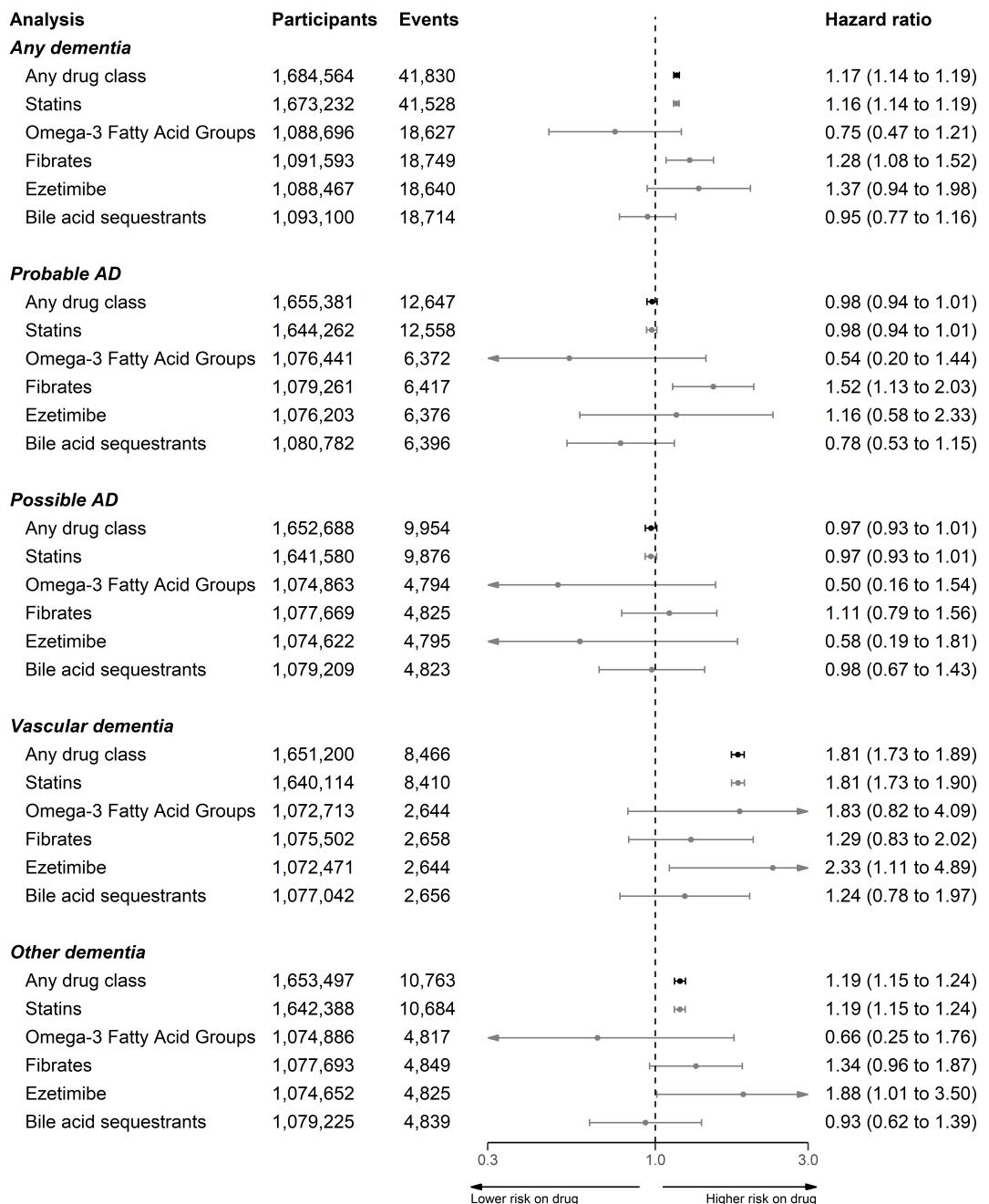


Figure 5.5: Results from primary analyses of CPRD data using the fully adjusted model and participant age as the time scale.

Alzheimer's disease

My results show little 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 to no treatment, with the sole exception of fibrates on probable Alzheimer's disease (HR:1.52, 95%CI:1.13-2.03).

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 dementias (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, we found treatment with a lipid-regulating agent was associated with a slightly increased risk (HR:1.17, 95%CI:1.14-1.19), which lies between the associations for the Alzheimer and non-Alzheimer dementia outcomes as would be expected. 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.²⁷⁸

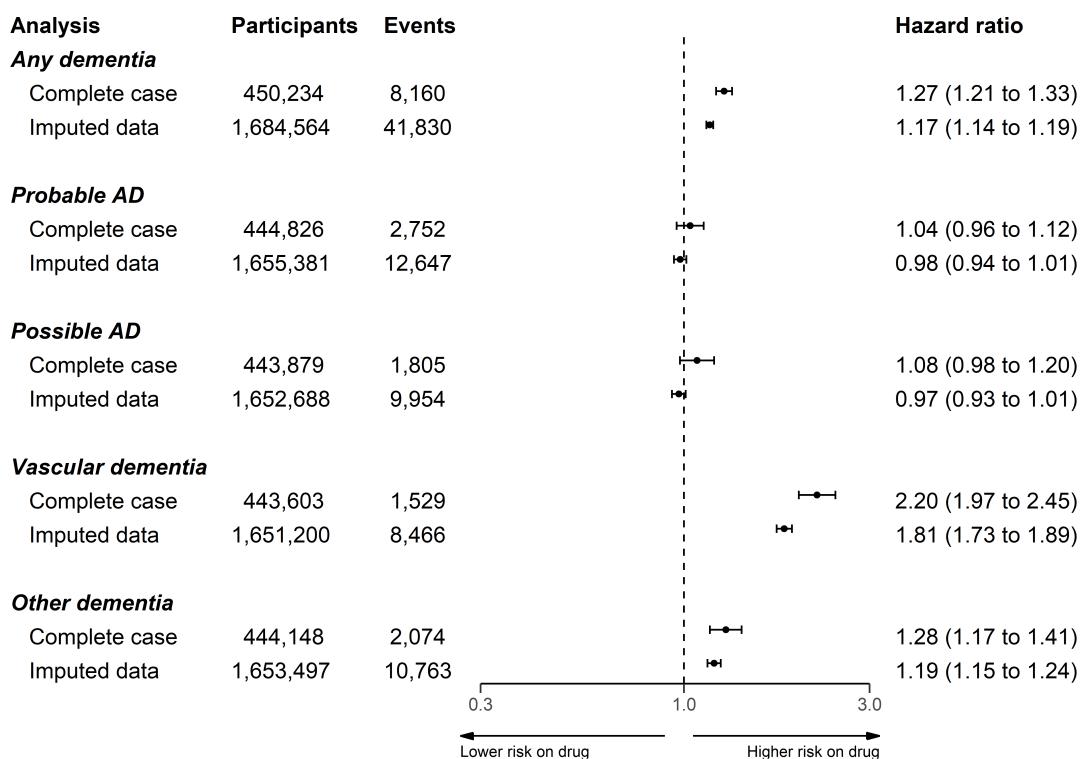


Figure 5.6: Comparison of analyses using the complete case versus imputed cohorts.

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) and Type 2 diabetes (HR: 1.50, 95%CI: 1.48-1.51).

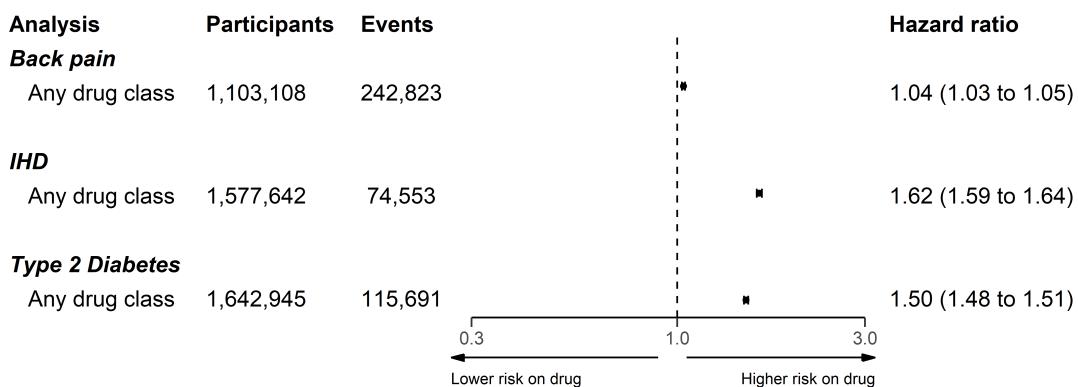


Figure 5.7: Results of control outcome analysis.

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 an 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

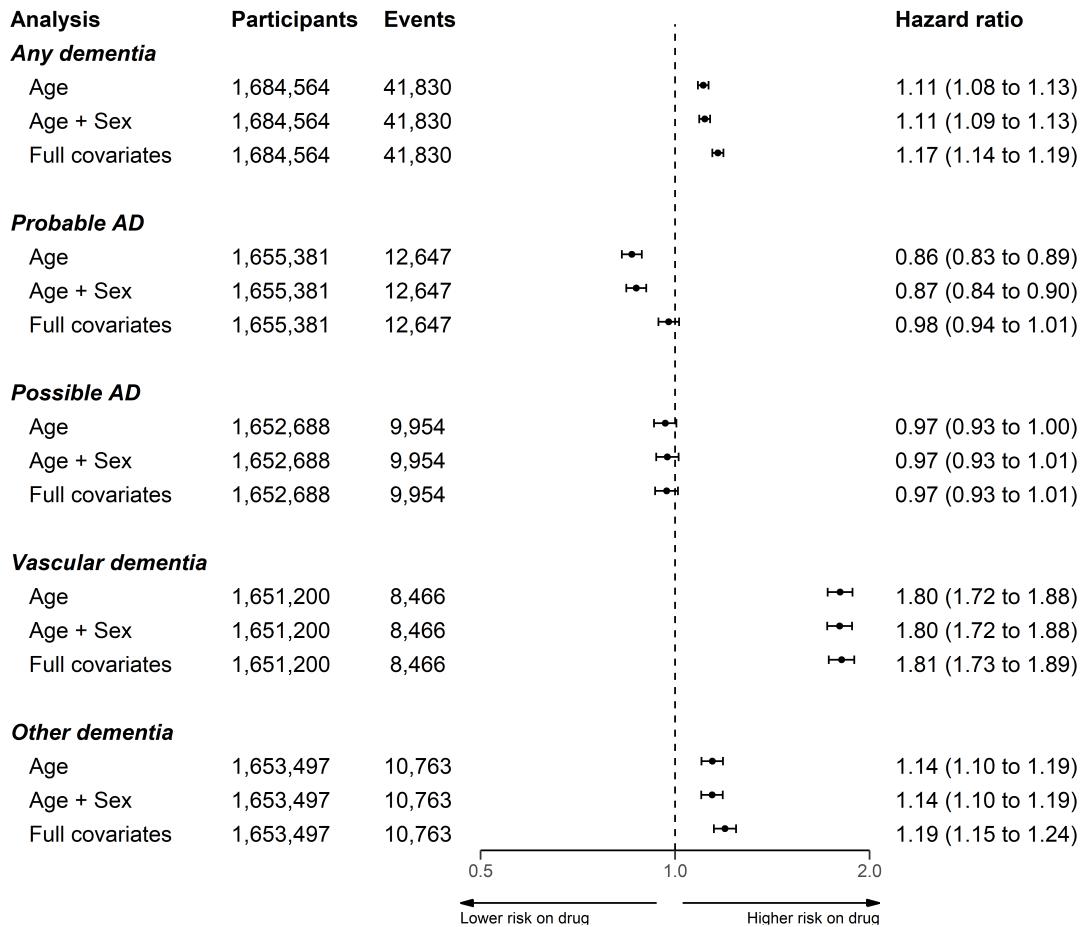


Figure 5.8: Results of three models adjusting for a different set of covariates.

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).

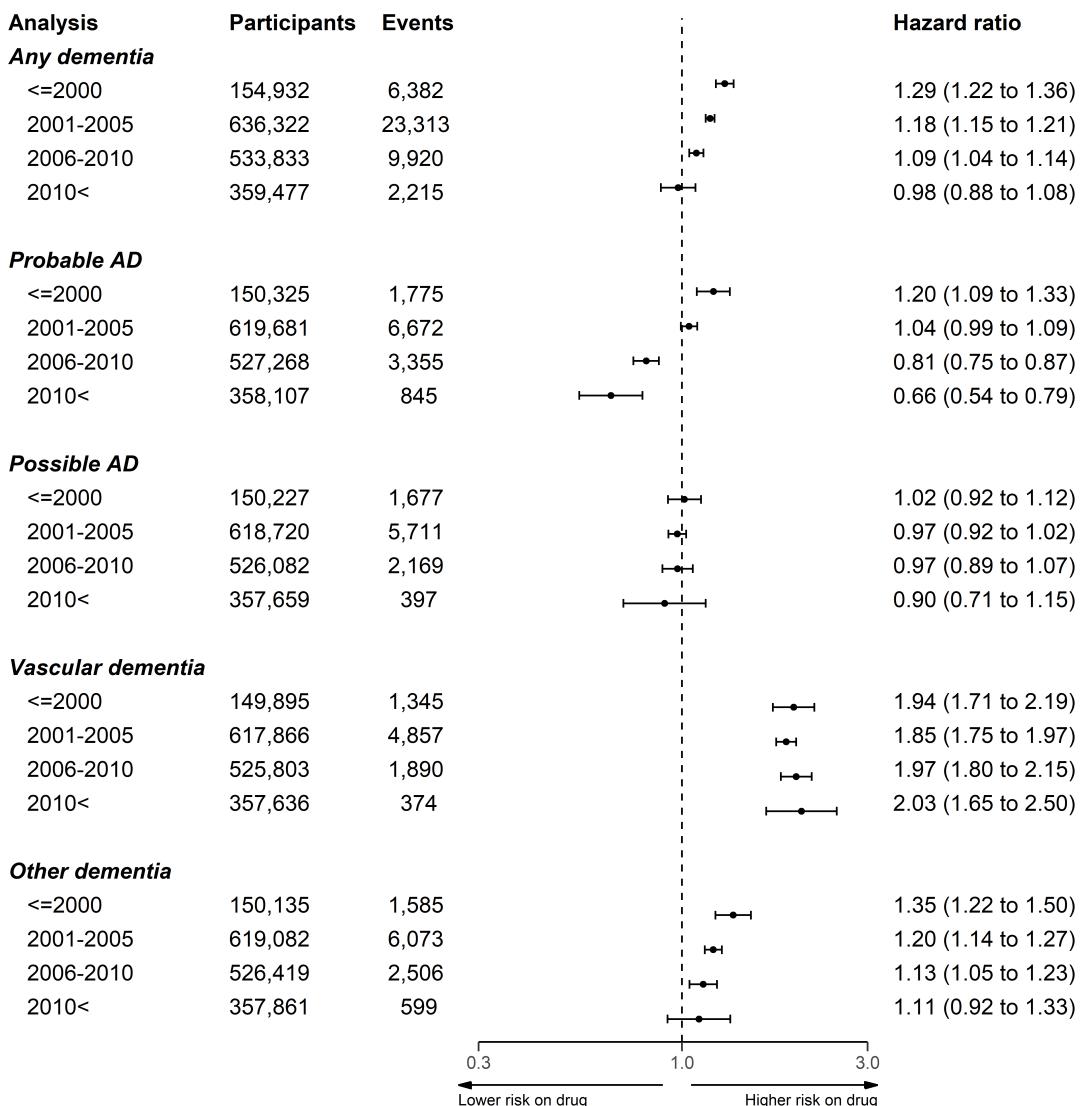


Figure 5.9: Analysis of any lipid-regulating agent on probable AD outcome, stratified by grouped year of cohort entry.

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 5.6). While the frequency of outcomes declines in more recent strata, likely due to the limited follow-up inherent to these groups, this decline in frequency is relatively constant across the dementia subtypes.

Table 5.6: Frequency of diagnoses by grouped year of cohort entry

Year of cohort entry	No dementia	Probable AD	Possible AD	Vascular dementia	Other dementia	Total
<=2000	148550 (95.9%)	1775 (1.1%)	1677 (1.1%)	1345 (0.9%)	1585 (1.0%)	154932
2001-2005	613009 (96.3%)	6672 (1.0%)	5711 (0.9%)	4857 (0.8%)	6073 (1.0%)	636322
2006-2010	523913 (98.1%)	3355 (0.6%)	2169 (0.4%)	1890 (0.4%)	2506 (0.5%)	533833
2010<	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).

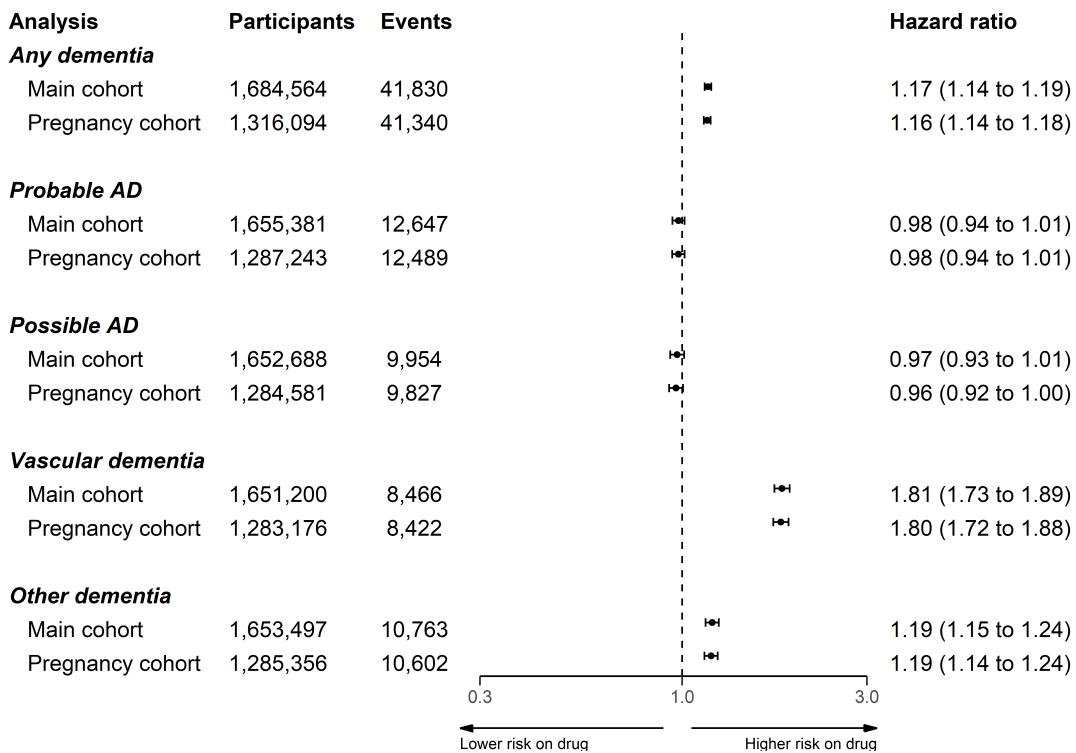


Figure 5.10: Comparison of analysis using main cohort and a cohort with potentially pregnancy women removed.

Statin properties

In the cohort, statins with lipophilic properties were much more frequently prescribed than hydrophilic statins (Table 5.7). 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.

Table 5.7: Summary of statin properties (lipophilicity vs hydrophilicity) by grouped year of prescription.

Prescription Year Group	Hydrophilic	Lipophilic	Total
<=2000	7037 (18.2%)	31531 (81.8%)	38568
2001-2005	21427 (10.3%)	187018 (89.7%)	208445
2006-2010	3566 (1.6%)	217726 (98.4%)	221292
2010<	1115 (0.9%)	119035 (99.1%)	120150

5.4 - Results

When stratifying by statin properties, hydrophilic statins were less harmful in the any, vascular and other dementias outcomes compared to lipophilic statins (Figure 5.11). Additionally, in the AD outcomes, hydrophilic statins were associated with a small reduction in risk, compared to the weak evidence for an effect for lipophilic statins.

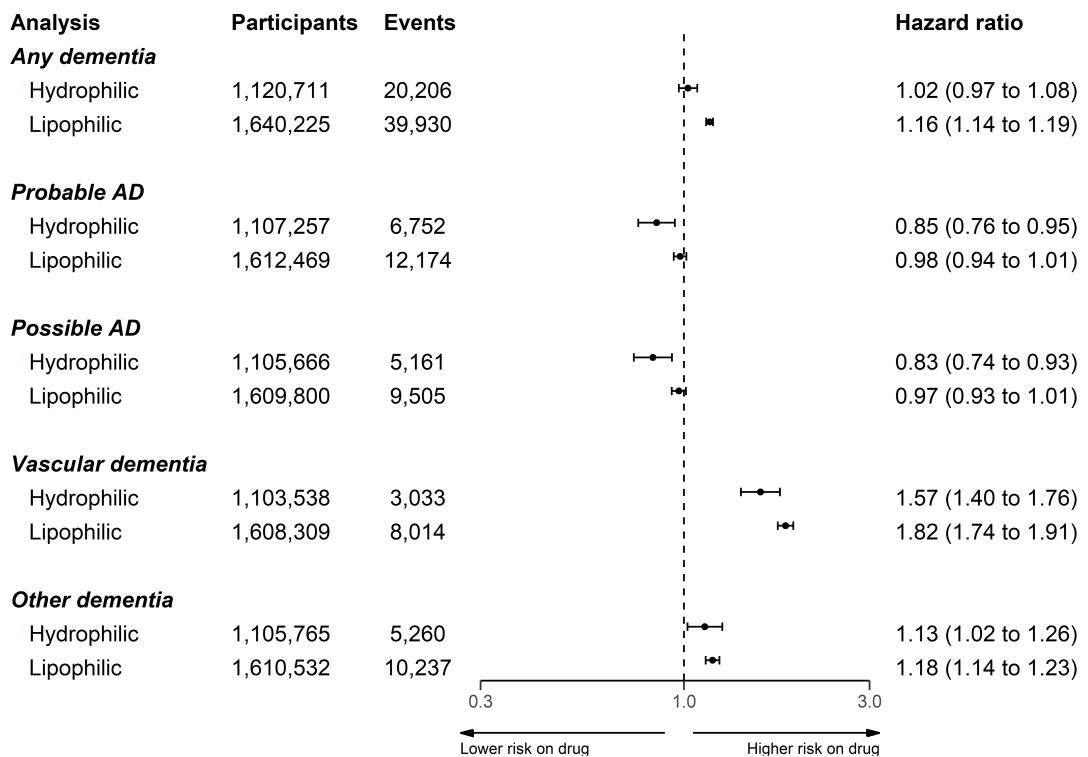


Figure 5.11: Analysis stratified by statin properties (hydrophilic vs lipophilic)

Impact of dementia code lists

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 ("other") dementia outcomes, respectively.

5.5 Discussion

5.5.1 Summary of findings

Lipid-regulating agents showed little evidence of an association with probable and possible Alzheimer's disease when compared to no treatment, but were associated with increased risk of an all-cause dementia, vascular dementia and other dementias diagnosis. The estimate 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 dementias, 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 dementias 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 detailed comparison of the result presented above with the existing evidence base identified by the systematic review (Chapter ??), see Chapter 7.

[Feedback requested: How much should I be cross-referencing with the systematic review at this stage? Or should I keep it self-contained, and

then provide a comparison across the review/CPRD/IPD analyses as the first part of the triangulation chapter?]

Confounding by indication

A likely explanation for the observed increased risk of dementia outcomes with a vascular component (e.g. vascular and other dementias) with lipid-regulating agent use, and an important limitation of this analysis, is residual “confounding by indication”. While the term has been used to describe different source of bias in epidemiological analyses,²⁹⁴ it is used here to described the role of risk factors that both prompt treatment and increase the risk of the outcome, thus causing a distorted positive association between the treatment and outcome (see Figure 5.12). In causal inference nomenclature, statins and dementia are said to be *d*-connected as there is an open “backdoor” path between them via the uncontrolled confounders.²⁹⁵ In the context of this analysis, the confounding variable (or, more likely, variables) would prompt prescription of statins but also represent a vascular risk factor that contributes to the development of the vascular dementia.

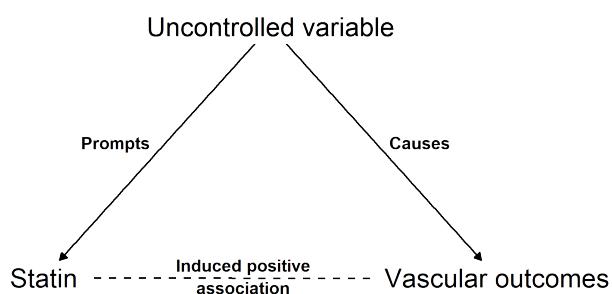


Figure 5.12: Causal diagram (directed acyclic graph) illustrating how confounding by indication could induce a positive association between statins and vascular dementia.

5.5 - Discussion

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. 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.

In line with this, an increasingly harmful effect is observed when moving from the probable and possible Alzheimer’s disease outcome to the other dementias outcome, and finally to the vascular dementia outcomes. This pattern suggests that the strength of the residual confounding by indication increases as the proportion of cases with a vascular component in an outcome definition increases. Given confounding related to vascular factors, this pattern is also expected given the decision tree for assigning outcomes in the presence of greater than one dementia code. Under this system, the Alzheimer’s disease outcomes require a “pure” condition and the presence of any vascular or other dementias 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 outcomes) is not unusual. Using a conventional epidemiological technique, a previous analysis also found an increased risk of coronary heart disease (analogous to the ischemic heart disease outcome used in this analysis) in those taking statins (HR: 1.31, 95%CI: 1.04-1.66).²⁹⁶ Controlling for confounding by indication in that study through the use of a trial emulation analysis gave an estimate of 0.89 (95%CI: 0.73-1.09), a comparable though less precise estimate to that observed in RCTs of statin use (0.73, 95%CI: 0.67-0.80).²⁹⁷

Given the absence of vascular dementia in the published literature, as highlighted in the previous chapter (see Section ??), the unexpected increase in vascular dementia risk with statin use is particularly interesting. It is possible that previous research encountered similar methodological issues to this analysis, and via a publication

bias mechanism where unexpected or assumedly incorrect results are less likely to be published, the results from these analyses did not make it into the evidence base.

Statin properties

This analysis found that hydrophilic statins were less harmful in the any, vascular and other dementias outcomes compared to lipophilic statins, and were associated with a small reduction in the risk of the probable and possible AD outcomes. The relative precision of the estimates in each group is expected, as the two most commonly prescribed statins are lipophilic (simvastatin and atorvastatin).²⁹⁸

A widely discussed concept in the literature surrounding statin use and cognitive outcomes is the fact that lipophilic statins are more likely to be able to cross the blood brain barrier, and so have a more potent protective effect by directly lowering brain cholesterol.²⁹⁹ My findings that hydrophilic statins appear to be more protective/less harmful than their lipophilic counterparts runs counter to this assertion.

An initial interpretation of the the different association observed in the two groups was that the lipophilic statins may be more potent, and so are prescribed to patients with a higher underlying vascular load, leading to increased confounding by indication in this group. However, the statin with the strongest lipid lowering effect that is available via the NHS, rosuvastatin, is hydrophobic.

[Yoav, are there are particular indications/contraindications for statins with different properties? And any general comments on the interpretation of these findings would be welcome!]

Impact of code lists

As part of a sensitivity analysis exploring the impact of outcome code-lists, I used definitions for Alzheimer's disease and other dementias obtained from a previously published paper (Smeeth *et al.*).²⁹² Using these lists in my analytical set-up, I found a harmful association of statin use with both outcomes.

This finding disagrees with the results of the original analysis, which found evidence for a protective effect of statin use on all-cause dementia (HR: 0.81, 95%CI: 0.69-0.96) and non-AD dementia (HR: 0.82, 95%CI: 0.69-0.97), but little evidence of an effect on AD (HR: 0.81, 95%CI: 0.49-1.35).

However, comparison of the results obtained using the two sets of code lists was deemed less useful following a detailed comparison of the codes used. While all of the codes used to define Alzheimer's in the Smeeth *et al.* paper are included in the probable Alzheimer's code-list (see Figure 5.13), 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.

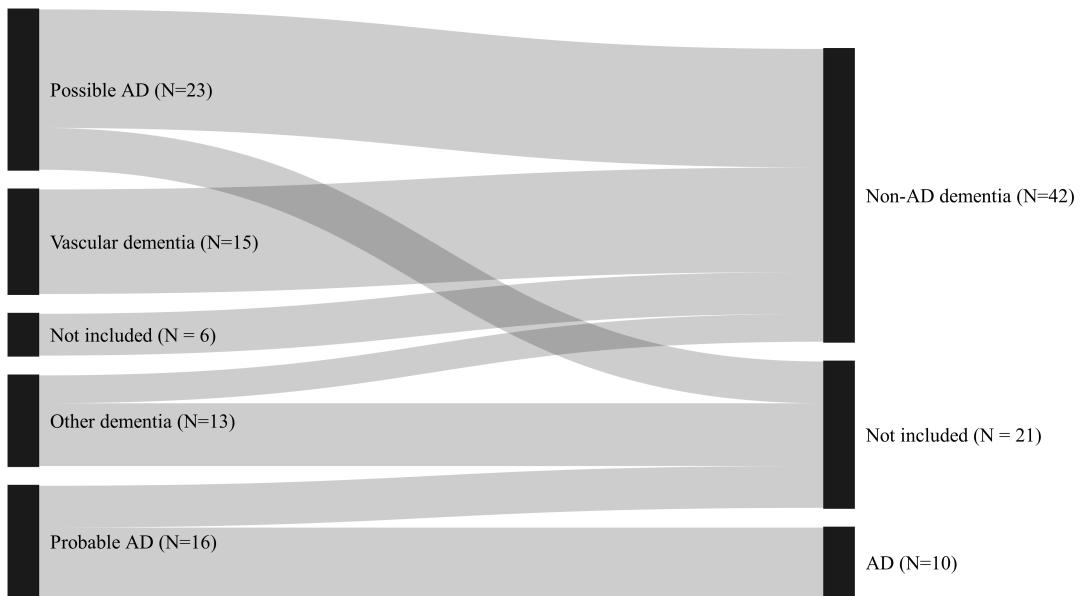


Figure 5.13: Sankey diagram comparing the codes used in this analysis with those used in the Smeeth *et al* paper.²⁹² The outcomes and number of codes contributing to each are presented (the Smeeth *et al.* outcomes are on the right hand side of the figure). The joining lines showing the overlap between 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 used to define competing outcomes (e.g. AD vs non-AD dementia). This different codes used by Smeeth *et al.*, in addition to an analytical approach that adjusted for covariates defined after the index date, may go some way to explaining why our analysis obtained different results despite the substantial overlap in the data sources used.

5.5.3 Strengths and limitations

The primary strength of this analysis is the relative size of the CPRD and length of follow-up. Having reviewed the existing literature, as identified by the systematic review in Chapter ??, 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

5.5 - Discussion

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. Finally, it is one of the few that provide evidence on the association of lipid-regulating agent use and vascular and other dementias.

However, the findings of this analysis are subject to several limitations in addition to the confounding by indication discussed above. There is a strong possibility of differential misclassification of dementia-related conditions based on the exposure.³⁰⁰ 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 contains prescriptions for lipid-regulating agents. Further, there is a potential for general non-differential misclassification of the outcome due to the varying positive predictive value of electronic health record code lists to identify dementia cases.^{258,259}

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, as illustrated by the discrepancy between the results when using the codes lists defined for this study and those used by Smeeth *et al.* This presents a particular challenge in comparing research across different time-periods and coding systems, particularly the unexpected results I obtained for the vascular and other dementias outcomes.

A further limitation stems from the possibility of uncontrolled confounding due to genetic factors. Number of ApoE $\epsilon 4$ alleles represents the strongest genetic risk factor for Alzheimer's disease, but also substantially increases LDL-c levels,³⁰¹ potentially prompting treatment with a statin or other lipid regulating agent. I was unable to control for ApoE genotype in this analysis as I did not have access to genetic data on participants. As a result, any protective association between LRA use and the Alzheimer's disease outcomes may be masked by residual negative confounding by ApoE.

Finally, as with many studies of dementia, there is a risk of reverse causation in my analysis. Dementia and associated conditions have a long prodromal period, during which preclinical disease could cause indications for the prescription of a lipid-regulating agent.

5.5.4 Enabling easy synthesis of this analysis

The raw data supporting this analysis is not publicly available, as access to the CPRD data is controlled by a data monitoring committee. However, when data are not readily available, sharing the analysis code and summary statistics represents a way for readers to validate the findings.³⁰²

In light of this and my own experiences in attempting to extract information for papers assessing preventative treatments, as documented in Section 3.5.5, the outputs from this analysis have been made readily available. All code, Read code lists and summary statistics (namely the tables presented in this chapter plus summary tables of effect estimates) can be downloaded in a machine readable format from the archived repository for this project (Zotero DOI: **TBC**). This open approach should enable easy inclusion of this analysis in future evidence synthesis exercises, allowing new work to build on that presented here.

5.5.5 Conclusions

This chapter has provided new evidence on the association of lipid-regulating agents with incidence 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 potential immortal time bias.

5.5 - Discussion

It found little 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 ???. 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. The presence of bias 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 dementia outcomes with a vascular component. 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.²⁹⁶

Regardless, this analysis has provided an additional source of evidence for the triangulation exercise presented in Chapter 8. In the following Chapter, the dataset described here is incorporated along with several other datasets as part of an individual patient data analysis to investigate the association of blood lipid levels with incidence of dementia outcomes directly, rather than via the proxy of lipid-regulating treatment.

5.6 References

6

Individual participant data meta-analysis

Contents

6.1	Lay Summary	153
6.2	Introduction	154
6.3	Methods	155
6.4	Results	160
6.5	Discussion	166
6.6	Conclusion	169
6.7	References	170

6.1 Lay Summary

As part of a broader investigation into the relationship between lipid levels and dementia risk, I applied for access to 37 unique data sources containing information on these variables. However, only a small proportion of these data sources ($n = 3$, 8.1%) provided me with access.

The resulting analysis did not suggest an effect of any lipid fraction (total cholesterol, LDL-c, HDL-c, or triglycerides) on any dementia outcome. In addition, patients characteristics such as age and sex did not have any impact on the relationship of lipid fractions and dementia outcomes.

Finally, the reasons for the low response rate are explored and future studies to investigate how data access rates could be improved are suggested.

6.2 Introduction

Individual patient data meta-analyses are considered to be the gold standard form of evidence synthesis, allowing for the application of a common selection model and analytical approach across all identified cohorts. They are particularly useful for investigating the impact of patient-level characteristics, something that is not possible with aggregate data unless the results have been explicitly stratified by the characteristic of interest.

While the systematic review presented in Chapter ?? did not provide strong evidence for an effect of age at lipid measurement or patient sex on the relationship between lipids and dementia, the number of included studies in the lipid fraction meta-analyses were very small, due to both the relatively small number of studies reporting on this exposure and the poor reporting of summary statistics of patient characteristics.

As such, the aims of this analysis were two fold: firstly, to perform a individual patient data meta-analysis across identified cohorts to examine the impact of patient age-at-measurement and sex on the effect of lipids on dementia; and secondly, to expand the evidence base for the effect of lipids on dementia outcomes by incorporating previously unanalysed cohorts contained in the Dementia Platform UK, a large consortia of dementia cohorts.

6.3 Methods

6.3.1 Eligibility criteria

Eligible data sources were defined as those that contained lipid measurements and lipids reported/available as a continuous measure. Data sources which were cross-sectional, either by design or due to the available data (e.g. a study conducted multiple waves, but only data from a single wave could be accessed) were excluded. Similarly, due to the limited time-frame, studies making use of population-level electronic health records, which often require an extensive project proposals in order to gain access to the data, were ineligible due to the time and cost involved in applying. The one exception to this was data from the CPRD, which I already had access to as part of the analysis reported in Chapter ??cprd-analysis-heading).

6.3.2 Applying for data access

Potentially eligible data sources were identified via two mechanisms with a distinct focus. For each approach, the number of cohorts responding to the request, and where applicable, the reasons given for a lack of data sharing were recorded.

Systematic review

The first approach focused on previously analysed data sources. The systematic review detailed in Chapter ?? identified several observational prospective cohort studies examining the effect of blood lipid levels on dementia outcomes. The data sources used in each of these cohort analyses were screened against the criteria listed in the previous section, and eligible cohorts were approached for data access. In the first instance, the first/corresponding author of the publication was emailed (see Appendix A.5.1 for a copy of the text and documentation sent to potentially collaborators) in Autumn 2020. If this approach did not elicit a response within two months, the last author was contacted, on the basis that the

first/corresponding author may have been a more junior member of the research group who has moved to a different institution.

Dementia Platform UK

The second approach focused on incorporating previously unanalysed data, thus providing additional evidence on the relationship between statins and dementia risk.

This was achieved through the Dementia Platform UK (DPUK), a collaborative grouping of existing dementia cohorts established by the MRC which works with data owners to make their data readily accessible for secondary analysis.⁹⁹ It provides access to 42 cohorts with over 3 million participants, and makes use of a central streamlined application processs for all cohorts with the intent of making it easier to access data from existing data sources.

Cohorts included in the DPUK were assessed against the eligibility criteria, and in Autumn 2020, an application for access to a subset of 16 cohorts was made via the common-access procedure.

6.3.3 Risk of bias assessment

Risk of bias assessment was performed for each of the included cohorts using the risk-of-bias assessment tool for non-randomised studies of exposures introduced previously (for a more detailed discussion of this tool, see Section 3.3.7).

Due to the inherent conflict of interest in assessing my own analysis, the result were also assessed by two independent reviewers.

6.3.4 Missing data

In this analysis, missing data were handled using a multiple imputation approach. Variables with missing observations were identified, and 20 imputed datasets were

created.²⁷⁸ 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.²⁷⁹ Imputation was performed using the MICE (Multiple Imputation by Chained Equations) using the XXX command in R.

6.3.5 Variable definition

Exposure

To aid interpretability of the outcome, all analyses were standardised to refer to a 1-standard deviation increase in the lipid fraction under investigation.

Outcomes

Similar to the analyses presented in other chapters, the endpoints in this analysis was incidence of dementia and related sub-types (AD, vascular dementia).

Covariates

Relevant covariates were adjusted for in all models. These included age, sex, education and comorbidities. For information on how these were defined across cohorts, see Appendix ??.

6.3.6 Statistical analysis

Descriptive statistics

All IPD obtained was examined and recoded as necessary to ensure they comply with the coding guideline for the IPD database. As recommended, summary statistics for each data source were calculated and compared to publicly available

6.3 - Methods

statistics to ensure no errors were introduced in the data cleaning process.³⁰³ Any discrepancies were investigated.

Data cleaning and harmonisation

Across all cohorts, data cleaning was performed in a similar manner, using commonly named variables, so that a single model could be applied using functional programming. The one exception to this is the CPRD data, which was held in a different system to the rest. However, an identical model was applied within this cohort. The advantage of this approach is that it reduces the likelihood of errors in model mis-specification if needing to change variables names from cohort to cohort.

Individual patient data analysis

In this analysis, an two-stage individual patient data model was used out of necessity, with the one-stage model being precluded by the included datasets being in protected data silos. It also allows for each studys' data to be analysed according to a common analysis plan, which is particularly important in this topic, due to the differing definitions between studies of the serum concentrations above or below which lipids are deemed to be at abnormal levels.

Under a two-stage approach, estimates for the effect of each lipid fraction on available dementia outcomes in each study were first calculated for each data source. Effect estimates for the associations between lipid fractions and incident dementia, adjusted for identical covariates, were obtained by applying a common model to each data source. More specifically, hazard ratios were used to quantify the effect of a 1-SD lower exposure to a lipid fraction on dementia outcomes. A discrete proportional hazards model was employed to account for the interval censoring introduced by design of the longitudinal cohort studies.³⁰⁴ An overall effect estimate was then produced by combining data-source-specific estimtes in a random-effects meta-analysis (see Section ?? in Chapter ?? for a broader discussion of meta-analysis methods). Following best practice, clustering within studies was accounted for.³⁰⁵

Investigating effect of patient-level covariates

In order to investigate the interaction of patient-level characteristics with lipid levels, interaction terms for lipid-covariate terms were included in the model above. These were extracted and synthesised using a random effects meta-analysis.

Combining published aggregate data and individual participant data

Where a data source did not provided their individual participant data, but stratified by a patient characteristic of interest in the published data, the aggregate data and individual data were combined.

6.3.7 Comparison with published findings

Where an included data source had previously been analysed and results for the association between lipid levels and dementia outcomes reported, these were compared to the results of this analysis. If discrepancies were identified, these were investigated.

6.3.8 Secondary analyses

Amount of information added

The proportion of new information added to the evidence base via incorporation of the unanalysed datasets was quantified, by addding the additional results to the lipid-dementia meta-analyses presented in Section ??, and extracting the combined weights assigned to the new data in a fixed-effect model.

Data availability statements

As a sensitivity analysis, I examined the data sharing statements for each using a predefined categories used in a similar previous analysis.³⁰⁶

6.4 Results

6.4.1 Data access

Of the 37 studies to which I applied for data access, only three (8.1%) provided data and were included in the analysis. Figure 6.1 details whether the cohorts eventually included in the review were identified by the systematic review or via DPUK portal. In addition, the reasons for cohorts not being included in the analysis are presented, stratified by application approach.

In summary, the requests for data from cohorts identified by the systematic review was characterised by a very low response rate. For the minority who did respond, common reasons given by corresponding authors for not sharing the data included that they no longer worked with the same group and did not know if the data were available or how to obtain it, had access to the data, or that they were currently or intended to perform a similar analysis as the one proposed.

With respect to the application to DPUK cohorts, where a dedicated project manager liaises with data owners on the applicants' behalf, the overall response rate was higher resulting in eventual access to three cohorts (one of which was also identified by the systematic review). However, even using this streamlined approach, a response was obtained for only half of the approach cohorts.

```
in line 14 near 'data'
```

Figure 6.1: Flowchart of included cohorts, stratified by identification method (systematic review vs DPUK).

As highlighted in Figure 6.1, there was little overlap between cohorts identified by the systematic review and those contained in the DPUK ($n = 1$, 2.7%), indicating that the DPUK is a useful source of unanalysed data with respect to this question.

6.4.2 Included data sources

The four data sources used in this analysis are summarised in Table 6.1 and are described in detail in the following sections. Of note, all included data sources were based in the United Kingdom. This is likely due to the majority of included datasets coming from the Dementia Platform UK route (Figure 6.1), which as implied by the name, has a geographical focus on cohorts performed in the United Kingdom.⁹⁹

Table 6.1: Summary of cohorts for which data were available

Cohort	N	Dementia events (all-cause)	Age (mean)	Male (%)
CaPS	2512	1034	52	100
CPRD	X	X	X	X
EPIC	1001	5	52	45
Whitehall II	8022	181	50	69

Caerphilly Prospective Study

The Caerphilly Prospective Study is a longitudinal study of men in the Cholesterol measures (total, LDL-c, HDL-c and triglyceride) were measured at baseline in 1979-1983. As the study population has aged, additional outcomes. Of particular relevance to this analysis, from Phase III (1989-1993) onwards, a battery of cognitive tests were introduced.

CPRD

The Clinical Practice Research Datalink (CPRD) is a large population-based, electronic health record (EHR) database.²⁶⁵ containing 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.^{265,268}

The CPRD is introduced more fully in Chapter @ref(). Briefly, a similar approach to the cohort definition as used in Chapter

Participants were included from the first date of lipid measurement, so no issues with immortal time bias as discussed in Chapter ...

6.4 - Results

Additionally, other number of participants is larger in this analysis as there is no restriction on the level which lipids should be in order to be included in the analysis.

the

Epic Norfolk

The European Prospective Investigation of Cancer - Norfolk is a^{307,308}

The added evidental value of the EPIC cohort is small, given the relatively small number of participants and the fact that the cohort has only 8 dementia events.

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

This data source was analysed in one of the included studies identified by the systematic review presented in Chapter ??,²²¹ meaning that a comparison between the published result and the analysis reported here was was possible.

6.4.3 Excluded data sources

Cohorts identified the systematic review but excluded from the IPD analysis

The main reason for exclusion within this records was use of an electronic health record database. Examples included the Taiwan National Health Insurance database and studies run in the US making use of Medicare data.

Cohorts identified in the DPUK but excluded from the IPD analysis

Several additional cohorts are present in the DPUK collection but were not formally approached with a request for data access. This was primarily due to the information obtained via the online data dictionary system or via initial contact with the cohort owners indicating that the study was relatively new (and so was yet to collect more than 1 wave of data) or that the exposure variables of interest were not recorded.

Cohorts providing data but ultimately excluded

As highlighted in Figure 6.1, several cohorts from the DPUK responding positively but on inspection of data provided these cohorts were excluded for a range of reason. The exclusion reason for each cohort for which data were provided is illustrated in Table 6.2.

"“, "Cohort“, "Reason" “1”,“NICOLA”,“Cross-sectional - only one wave of data available” “2”,“TRACK HD”,“Participants carried HTT gene (i.e. premanifest Huntington’s Disease). Cohort owners indicated that this is likely to overshadow”

Table 6.2: Exclusion reasons for cohorts providing data

Cohort	Reason
NICOLA	Cross-sectional - only one wave of data available
TRACK HD	Participants carried HTT gene (i.e. premanifest Huntington's Disease). Cohort owners indicated that this is likely to overshadow

BRACE is a memory-complaint-based cohort, suggesting that it is unlikely they were dementia-free at baseline). Similarly, the Memento study was excluded as its criteria for entry required participants to be outpatients from a memory clinic. The Generation Scotland provided only cross-sectional data only for the variables of interest, as did the NICOLA study. Finally, the ELSA study was excluded based on the definition of the exposure variable. This study provided a “high cholesterol” a binary variable rather than the raw continuous lipid measurement data.

Cohorts approach but received no answer

This was particularly common among the cohorts identified via the systematic review and included studies with a dedicated access process or data access committee. A cohort of particular interest, having been used in multiple studies identified by the systematic review, is the Three City Study in France. Despite multiple attempts to contact the team, there was no response received. This study was approached at multiple points over the course of this analysis, and no answer was received.

This was also a problem for cohorts approach via the DPUK - half of the cohorts applied to did not respond to the application within a year.

6.4.4 Analytical results

There was no evidence for an effect of any lipid fraction on any dementia outcome (Figure 6.2). Additionally, meta-analysis of the interaction coefficients for age-at-lipid-measurement and participant sex did not reveal an evidence for an varying effect across these variables.

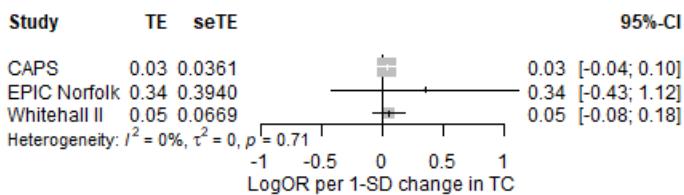


Figure 6.2: Random effects meta-analysis of estimates from included data sources:

6.5 Discussion

6.5.1 Summary of findings

This analysis requested data from 37 data sources, but only obtained data from $n = 3$, 8.1%, all of which were based in the United Kingdom. No evidence for an effect of lipids on the risk of dementia or related outcomes was identified. Similalarly, there was no evidence for an effect of age-at-measurement or sex on the lipid-dementia relationship.

Additionally analyses revealed that authors of cohorts indentified by the systematic review rarely stated that the data underpinning their analyses would be freely available, while a review of corresponding authors of these publications demonstrated that for many, the institutional email address provided was out of date.

6.5.2 Limitations

Low response rate to request for data

(although, in hindsight, the form of contact used (email) has been shown to be less successful in eliciting responses from authors when compared with telephoning³⁰⁹)

A key limitation of this analysis is the very low response rate to requests for data access. This finding is not unexpected however, given that a review of IPD studies published between 1987 and 2015 found that fewer than half managed to obtain data from greater than 80% of studies, and that in many cases, the exact percentage of studies for which data were obtained was not accurately reported.

There are likely several likely reasons for this low response rate. In general terms, there are several well-documented reasons why data are not made readily available, including concerns regarding participant privacy, concerns around scoping or “parasitic” behaviour, and a lack of trust between primary and secondary researchers.

More specific to this analysis, individual participant data meta-analysis including studies other than randomised controlled trials have less success in obtaining individual participant data from studies.⁹⁴ Additionally, while no evidence is available on whether the characteristics of the researcher requesting data access influences the response rate and eventual decision, there is the possibility that my position as a PhD student meant I was less likely to elicit the response rate than a well-known professor might. The timing of the requests for data access, coinciding with a global pandemic, may also have affected the response rate as researchers prioritised other work.

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.³¹⁰⁻³¹² 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.⁹⁷ Several of the included studies are greater than 10 years

6.5 - Discussion

old, and so obtaining data from these cohorts is less likely compared with more recently published studies.

One of the reasons for this may be that the email addresses reported on publications are more likely to be out of date for older publications. Qualitatively, investigation of a subset of cohorts revealed that several corresponding/first authors were no longer at the same institution as when the study was reported, and as a result, were unlikely to have access to the institutional email address listed on the study publication.

These challenges are in theory what the DPUK was built to address. However, even with the help of the streamlined application process afforded by the DPUK, accessing sufficient data was a challenge. The response rate among DPUK cohorts a year after application was just 50%. In addition, some cohorts responded that that the proposed study question was already under investigation by another group, and that they would not share the data on this basis. In light of this, a centralised database of ongoing analysis being performed in DPUK would be of enormous help. Finally, the DPUK should more clearly distinguish between those cohorts that are “DPUK native” versus externally hosted, as the response time for externally hosted cohorts is likely to be much longer.

Whether or not to press ahead with an IPD analysis in the absence of all (or even most) data is a personal decision, and some previous analyses have highlighted where they decided not to pursue an IPD study.³¹³

For the purposes of this thesis, conducting the IPD analysis was useful as it afforded me the opportunity to gain experience in applying the methodology. In addition, it allowed me to analyse two additional previously unanalysed cohort studies which were then utilised in the triangulation study detailed in Chapter 7.

Risk of bias assessment

There is some concerns about performing risk of bias assessments on your own analysis, and so two secondary reviewers

6.5.3 Strengths

While this analysis did not manage to systematically obtain and analyse a large proportion of identified data, it did enable an analysis of two previously unanalysed datasets - the CaPS study and the EPIC Norfolk study - providing additional data that is incorporated into the triangulation exercise reported in Chapter 7.

6.5.4 Reflections on the process

In hindsight, attempting to undertake a large-scale IPD meta-analysis as part of a larger PhD project may have been overly ambitious. Data harmonization between cohorts in an IPD analysis is an often under-appreciated challenge,³⁰³ and in line with this, data cleaning for this analysis took substantially longer than expected.

6.6 Conclusion

This analysis provided very weak evidence for an effect of lipid fractions on dementia outcomes. However, it does provide an analysis two previous unanalysed data sources that are subsequently used in Chapter 7.

6.7 References

Next on my list of features to be specially considered I would place the consistency of the observed association. Has it been repeatedly observed by different persons, in different places, circumstances and times?

— Sir Austin Bradford Hill, 1965³¹⁴

7

Aetiological triangulation across new and existing evidence sources

Contents

7.1	Lay summary	171
7.2	Introduction	172
7.3	Methods	173
7.4	Results	183
7.5	Discussion	186
7.6	Conclusion	193
7.7	References	194

7.1 Lay summary

Triangulation is the practice of using multiple sources of evidence to provide more reliable answers to a research question. Different sources of evidence will have different limitations. If the result from each source points towards the same answer, this improves our confidence in the conclusion.

7.2 Introduction

This chapter will attempt to triangulate the evidence identified by the systematic review in Chapter ??, with the results from the analysis of CPRD data (Chapter ?? and IPD 6.

7.2.1 Overview of triangulation

Aetiological triangulation, or simply triangulation, is the process of comparing and contrasting across result in order to .

It has been widely used in the literature, and is comparable to Bradford-Hill's criteria of "consistency", that is the replication of an observed relationship across several different contexts.³¹⁴ In this framework,

7.2.2 Terminology

Different types of triangulation

Qualitative triangulation is describing the different results. It regularly appears though not under that name, in the discussion of results section under the name "comparison with other literature".

Previous examples of this approach have used criteria such as

In contrast, *quantitative triangulation* attempts to obtain a overall numerical result by synthesising the different types of evidence.

However, previous attempts at triangulation ignore the potential for varying and opposing sources of bias within a evidence source. For example,

The terminology used in describing different mechanisms by which an observed result may complicate interpretation. This section clearly defines how each are used in h

Here, internal bias is used in terms of

In contrast, external bias (also called indirectness, or)

Additionally, potential biases can be additive or proportional. Pro

Talk about problems of scale - assuming bias direction at the study level works well if you only have one study per design category as in the example by Ference et al (e.g. IPD analysis of)

7.3 Methods

7.3.1 Data sources

As outlined in the thesis overview presented in Section 1.7, this triangulation exercise draws on the research produced in each of the preceding chapters. More specifically, this chapter builds on the comprehensive systematic review presented in chapter @re(sys-rev-heading), and incorporates the results of the analyses presented in Chapters ?? (association of statins with dementia outcomes in the CPRD) & 6 (association of lipids with dementia outcomes in IPD).

Table 7.1 illustrates the causal question each approach has attempted to address, along with the exposures and outcomes for each.

Table 7.1: Summary of studies included in this thesis, and used as evidence sources in the triangulation exercise. Note, Chapter 2 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 3/4	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 5	Are lipid regulating agents associated with dementia risk in a large scale electronic health record database?	Lipid regulating agents (statins, ezetimibe, fibrates, etc.)	Seven classes of lipid regulating agents	Provides additional observational data on vascular dementia (under-represented in the literature)
Chapter 6	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 a source of observational evidence created using a method with distinct sources of bias to those identified by the systematic review
				Allows for stratification of effect by variables of interest (e.g. sex)

7.3.2 Qualitative triangulation

As a first step to triangulating the available evidence, all information sources were grouped by outcome, and systematically compared and contrasted. Potential reasons for heterogeneity between study designs were examined with specific reference to the risk-of-bias assessments performed.

This analysis built on the detailed risk of bias assessments performed and presented as part of the systematic review (Chapter @ref()/@ref()). Similarly, risk-of-bias assessments were performed for each new source of evidence presented in this thesis, details of which are available in Appendix @ref(?).

7.3.3 Quantitative triangulation

In addition to the qualitative discussion of evidence, I attempted to integrate the numerical results of the multiple approaches. This proof-of-concept approach incorporated advancements in the way that bias in results is assessed (domain-based assessment tools) and existing methods for bias-adjusted meta-analysis³¹⁵ to illustrate how causal questions could be addressed under this quantitative triangulation framework.

Several steps were used to accomplish this approach, as described in the following sections. In summary these steps are:

1. Define the causal question of interest
2. Identify of relevant evidence sources
3. Assess(and visualise) the extent and direction of bias in each result
4. Define modifying terms for bias and indirectness in each result
5. Perform bias-adjusted meta-analysis using bias-/indirectness-adjusted results

The approach was implemented for two causal questions of interest, as defined below.

Definition of the causal questions of interest (case-studies)

This was also guided by the forthcoming ROBINS-E tool, which has

- Effect of average exposure to LDL-c during midlife on AD
- Effect of late-life LDL-c on all-cause dementia

Identify relevant studies

Once the causal question of interest had been defined, studies related to it were individual results were obtained from the general evidence sources described in Section 7.3.1

In order to illustrate the bias-/indirectness adjustment process detailed in the subsequent sections, the process for calculating the adjusted estimate for a single study is described in detail for a single study. Following this, the process was applied to the two subsets of studies related to the causal questions of interest.

Assess risk and direction of bias in each result

The risk of bias tools used to assess each result are described in detail in Section 3.3.7. In summary, randomized controlled trials were assessed using the RoB 2 tool;¹⁴⁵ non-randomised studies of interventions (NRSI), were assessed using the ROBINS-I tool;¹⁴²

For the sake of consistency, I mapped the different acceptable judgements in each risk-of-bias tool to a harmonised set of judgements: “Low”, “Moderate”, “High”. The exact mapping can be seen in Table 7.2. Of note, no mapping was performed for the critical risk-of-bias levels present in the tools for non-randomised studies.

This is because current best practice in evidence synthesis is to exclude all studies at critical risk of bias for further quantitative synthesis.

Table 7.2: robLevelsMapping

Harmonised	RoB2	ROBINS-I / ROBINS-E	MR
Low	Low	Low	Low
Moderate	Some concerns	Moderate	Moderate
High	High	Serious	High
-	-	Critical	-

In addition to assessing the extent of bias in each domain, as defined by the three levels detailed above, I attempted to records the direction of the bias, so that it could feed into the triangulation.

This was achieved via an additional question at the end of each bias domain, which asked assessors to predict the expected direction of bias in that domain. Note that this question is only apply to existing tools (ROB2/ROBINS-I/ROBINS-E), but I employed an identical approach during assessment of Mendelian randomisation studies. This optional question response were:

- “Favours experimental”/“Favours comparator” (additive bias)
- “Towards null”/“Away from null” (proportional bias)
- “Unpredictable”

As indicated in the options above, the response to this prompt is used both to define the direction and type (additive or proportional) of bias.

Additive biases (identified by a “Favours experimental”/“Favours comparator” response) are XXXX. A key example of additive bias is immortal time bias, which will always favour the intervention.

7.3 - Methods

Proportional biases (identified by a “Towards null”/“Away from null” response) are those related to the magnitude of the effect. A key example of this type of bias is non-differential misclassification bias, which will bias the effect estimate towards the null.

When predicting the direction of effect

From these responses the direction of effect was calculated. For additive biases, this was simply a case of shifting the effect estimate to the left or right. For proportional biases, decide

For example, the direction of the bias when the direction of bias is thought to favour the intervention/exposure will be the same regardless of the position of the point estimate. The same is not true for the “Towards the null”/“Away from the null” options.

For example, if the effect estimate represents a protective effect (below the null), then bias towards the null would be represented by a right pointing arrow (Figure ??). In contrast, if the effect of the intervention is harmful (effect estimate above the null), bias towards the null will be represented by a right-pointing arrow. The program to create these figures thus requires knowledge of both the predicted direction of bias and the effect estimate.

However,

S

Highlight that this is slightly different for the confounding domain in non-randomised studies, which only allows for the “Favours experimental”, “Favours comparator”, and “Unpredictable” options.

Of note, the guidance for the tools suggests that if assessors are not able to articulate why the bias in a given domain will “pull” the observed estimate in a specific direction, they should not try to guess.

To assess the usefulness of this system, the number of domains for which bias was expected (i.e. judged to be a more than Low risk of bias) but for which no expected

direction could be identified were expressed as a percentage. If this percentage is high, the helpfulness of the approach is

Visualisation of extent and direction of bias

To aid with the triangulation exercise and to act as, I created a new method of presentation of the extent/direction bias assessment results to enable detailed comparison across different studies contributing to the synthesis. While these figures do not

Similar to the standard paired forest plots introduced in Chapter ??, the level of bias across the domains in each study is reported using coloured blocks. However, in contrast to the previous plots, arrows are used in place to illustrate the predicted direction of bias in that domain, categorised as pulling towards the left or the right.

This second aspect required some additional consideration, given the distinct and competing options available to assessors.

Tha

Helper functions in the forthcoming `triangulate` package - see Appendix @ref() - help to preprocess risk of bias data for use with both bias-direction plots and adjusted meta-analyses. These bias direction plots themselves were created using the risk of bias isualisation tool described in Appendix ??).

Where the direction of bias was unclear/could not be determined from the report, this is indicated by a dash.

Assign modifying values to risk/direction of bias

For example, for a “High” additive bias in j th domain of the i th study is deifined as $\delta_{i,j}^{\text{High}}$ and uncertainty around this estimate is given by a distribution:

$$\delta_{i,j}^{\text{High}} = f(\mu_{i,j}^{\text{High}}, \sigma_{i,j}^{\text{High}})$$

7.3 - Methods

The sign of $\mu_{i,j}^{\text{High}}$ is defined by the relative direction of the bias in relation to the point estimate. If the bias is expected to pull the effect to the left, the position is given a negative sign, and if it is expected to pull the effect to the right, it is given a positive sign.

One key feature of the domain-based risk-of-bias domains is that the domains are considered interchangeable - i.e. a “high” bias in one domain is equivalent to a “high” bias in any other. As such, for this analysis, I assumed that the distributions assigned to each level of bias are identical across all values of j , that is, are identical for all domains in the risk of bias assessment tool.

$$\delta_{i,1}^{\text{High}} = \delta_{i,2}^{\text{High}} = \dots = \delta_{i,j}^{\text{High}}$$

and similarly

$$\delta_{i,1}^{\text{Moderate}} = \delta_{i,2}^{\text{Moderate}} = \dots = \delta_{i,j}^{\text{Moderate}}$$

For the purposes of this example analysis,

High Some concerns

These figures were sense-checked against

However, this also allows us to a priori define common modifying values across all domains of bias. In a bias-adjusted meta-analysis using this approach, a high risk of bias in the will have the same modifying effect on the results.

As such, a simulation approach was taken to explore the impacts of several different strategies

Table 7.3: Distributions for additive bias assigned to each bias judgement under two sensitivity scenarios, on the log scale.

Bias Level	Scenario 1	Scenario 2
Low	-	-
Moderate	N(.3, .12)	N(.3, .12)
High	N(.3, .12)	N(.3, .12)

Assign modifying values to indirectness

As I am attempting to identify the causal effect across different study designs, I needed to . This is not necessary if just summarising the literature.

A similar approach was used to define prior distributions for external bias or indirectness. I used a similar enforced scale of “Low/”Moderate“/”High" indirectness versus each of the 4 aspects of the target causal question. (recap).

In the original method, this was assessed against 4 categories of internal bias.

In this

“One important difference between these approaches that will need to be addressed in the triangulation is that RCTs relate to short-term modification of lipid levels, whereas Mendelian randomisation studies typically refer to lifelong exposures; and traditional observational epidemiology approaches may refer to short or longterm exposure (although are often not explicit about this). The triangulation will take these into account, for example by superimposing the studies on a hypothetical model for exposure-outcome relationship across the life course.”

Combine in a bias-adjusted meta-analysis

Once the total internal and external additive and proportional biases for each result were computed, they were used to define an adjusted estimate for each result included in the synthesis

Using the method defined in Turner *et al*,³¹⁵ the adjusted estimate for each result is defined as:

$$\hat{\theta} = \frac{y_i - \mu_{i\delta}^I - \mu_{i\beta}^E \mu_{i\delta}^E}{\mu_{i\beta}^I \mu_{i\delta}^E}$$

where $\mu_{i\delta}$ denotes total additive and $\mu_{i\beta}$ the total proportional bias, and the ^I and ^E superscripts denote internal and external (indirectness) biases respectively.

using the same terms, the standard error of this estimate, $\hat{\theta}$, is then calculated as:

$$SE(\hat{\theta}) = \left(\frac{1}{\mu_{i\beta}^I \mu_{i\delta}^E} \right)^2 \left(s_i^2 + (\mu_{i\beta}^I)^2 + \sigma_{i\beta}^I (\hat{\theta}^2 \sigma_{i\beta}^{E2} + \sigma_{i\delta}^{E2}) + \sigma_{i\beta}^I (\hat{\theta} \mu_{i\beta}^E + \mu_{i\delta}^E)^2 + \sigma_{i\delta}^{I2} \right)$$

Once each individual result has been adjusted, they are combined n a standard random-effects meta-analysis

Software

Personal communication with the authors of the original method resulted in the STATA code to implement the

As detailed in the introduction to this section, the model for defining bias-indirectness-adjusted results was available as STATA code. The has since been generalised as part of the `triangulate` R package to allow for users to apply the approach detailed here.

The package allows for preprocessing of domain-based risk-of-bias data to correctly identify the direction of the bias relative to the effect estimate, calculate

bias/indirectness-adjusted estimates for each result, and synthesise these adjusted results in a standard random-effects meta-analysis.

All meta-analyses conducted within the package are implemented using the `metafor` R package.

7.4 Results

7.4.1 Qualitative triangulation

All-cause dementia

Statins and all-cause dementia

This conflicts with the findings of our analysis, where statin use was associated with an increased risk of all-cause dementia (HR:1.17, 95%CI:1.14-1.19). Some of the included studies in the meta-analysis specifically exclude vascular dementia from the definition of all-cause dementia,³¹⁶ which may limit the ability for comparison with our findings for the all-cause dementia outcome.

Additionally, a previous analysis of the THIN EHR database using a propensity-score matched analysis found a protective effect of statins on all-cause dementia (HR:0.81, 95%CI:0.69-0.96)²⁹².

Alzheimer's disease

Statins and Alzheimer's disease

Our results are broadly in line with the findings of two distinct approaches examining the effect of statin treatment on subsequent Alzheimer's disease. No randomized trials of statins for the prevention of Alzheimer's disease have been reported, but a recent meta-analysis of 20 observational studies found statins were associated with a reduced risk of Alzheimer's disease (RR 0.69, 95% CI 0.60–0.80) with stronger

7.4 - Results

evidence than observed in our analysis.²⁴³ This review included case-control studies and analyses likely to be at risk of immortal time bias, which may account for the discrepancy with our findings. Additionally, a recent Mendelian randomization study examining the effect of genetic inhibition of HMGCR on Alzheimer's disease (a genetic proxy for statin treatment) provided equivocal evidence (OR: 0.91, 95%CI: 0.63-1.31) but was consistent with our results.³¹⁷

Our additional analyses stratified by statin properties found little evidence of differences in associations of lipophilic and hydrophilic statins and incidence of Alzheimer's disease, consistent with a recent meta-analysis of observational studies.⁵⁷

Vascular dementia

Statins and vascular/other dementia

Far fewer studies have tested the association between lipid-regulating agents and vascular dementia or other dementias. A recent review found four observational studies examining the association of statins and vascular dementia found limited evidence for an effect (RR:0.93, 95% CI 0.74–1.16).²⁴³ This contrasts with the harmful association found in our analysis (HR:1.81, 95%CI:1.73-1.89). When stratifying by lipid properties, lipophilic statins were more harmful than hydrophilic statins in vascular dementia, potentially due to their ability to cross the blood brain barrier.

Other drug classes

Apart from statins, few studies examining a lipid-regulating agent have been reported (Chapter 3/4).

One of the few classes for which evidence was available were fibrates, which found little evidence of an association with all-cause dementia was identified,¹⁵⁸ inconsistent with our finding that patients prescribed fibrates had higher all-cause dementia risk than those prescribed other lipid lowering agents.

A previous Mendelian randomization study found little evidence that genetic variants that proxy for ezetimibe affect risk of Alzheimer's disease (OR: 1.17, 95%CI: 0.73-1.87),³¹⁷ consistent with our findings. Note that this study was published in 2020 and so was not included in the review.

7.4.2 Quantitative triangulation

Single study

Explain how example study was chosen

Causal question #1

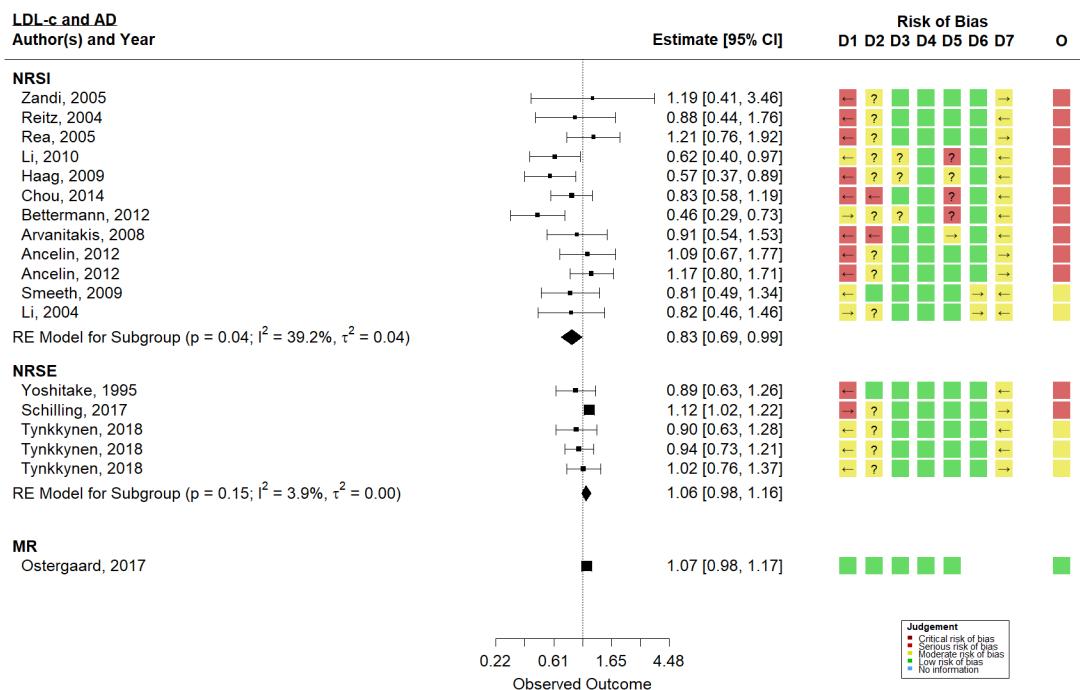


Figure 7.1: Bias direction plot, summarising internal biases related to effect of LDL-c on Alzheimer's disease :

Including sensitivity analyses

and citations of studies - going to be a bit of pain

As expected

Causal question #2

7.5 Discussion

Summary of risk-of-bias/triangulation results

Percentage of domains for which a direction of bias could be assigned?

Main takeway is that heterogeneity between study designs is substantially reduced

Compare bias-adjusted and

See what happens if using additive ($l=1, m=2, s=3, c=4$) vs multiplicative/log scale ($l=1, m=2, s=4, c=8$). Present adjusted results and meta-analysis of adjusted results.

Applicability/indirectness as an issue Compare with GRADE and cite George's example of 40% of effect predicted by MR seen when using statins for 5 years

7.5.1 Summary of findings

This chapter has demonstrated how new developments in systematic risk of bias assessment.

Future quantitative triangulation should move beyond the concept of comparing results at the approach level, and instead focus on the inherent threats to internal validity present in each specific result relevant to the causal question.

While this makes the process both more labour intensive and complex, to assume that, for example, all retrospective studies are equally at risk is to overlook. While

It could be true that all retrospective studies share some *minimal level* of bias, there are several opportunities. Failing to separate these individual results loses information, particularly given the now widespread availability of tools to systematically assess the biases present in each result.

In this chapter, I have presented a new method of visualising both the level and expected direction of bias inherent to each result, and suggested how this could be incorporated with a bias-adjusted meta-analysis approach to work towards a generalised quantitative triangulation framework.

I've highlighted the limitations of this method, as the correct distributions of modifying values in each risk-of-bias domain should be driven by meta-epidemiological studies which do not yet exist, framing them as opportunities for future research.

7.5.2 Limitations and strengths

Problems with quantitative triangulation approach

Missing natural variation obtained from elicitation approach. Trade-off between systematic approach presented here and personal variation obtained from several reviewers.

A potential combination between the two would be to use the variation in risk-of-bias assessments

Elicitation and bias-adjusted not as favoured versus weighting³¹⁸ though weighting at the overall risk-of-bias level loses some of the information contained in domain-level based assessments.

No good information on what the potential distributions for different levels of bias (High vs Moderate vs Low) should be.

Meta-epidemiological studies on the effect of different methodological issues on results exist for randomised controlled trials,^{319,320} suggesting that X, Y and Z.

However, there is substantially less , arguably because of the absence of clear Many meta-epi studies of non-randomised studies assume bias at the study-design level, similar to the early triangulation frame-work, rather than considering the actual biases deemed to be present in a given result. In this scenario, study

Potential future meta-epidemiology studies should begin to create these datasets.

A key example of a useful future study in this domain would be the mining of maximally adjusted vs. unadjusted estimates from abstracts from primary studies to assess the impact of insufficient confounding by topic. However, these datasets could also be built from systematic reviews of a topic, as they would already be gruped by reserach domain and provided the risk of bias data is shared, provide a ready source of information

For example, if it was . New software for performing risk of bias assessments, being developed by the Bristol Appraisal and review of Research group has this as a secondary methodological aim.

This approach has a number of advantages. One of the key ones is modifying distributions of each level of bias could be defined *a priori* in contrast to the .

In addition, the generalised framework will allow for

Subjectivity of risk of bias assessments

By definition, . In this case of this proof-of-concept analysis, the direction of bias was assessed by a single assessor. If attempting to impact policy, the

Defining appropriate levels of bias

Previous work on assessing the impact of different sources of bias on effect estimates . For example, a study examining the effect of found that studies with

Previous work examining

7.5 - Discussion

A further issue is the “Low” risk of bias domain. In this proof-of-concept analysis, I assumed that domains at “Low” risk of bias did not require any adjustment. However, “Low” risk of bias is not equivalent to the absence of bias, and potentially should still be adjusted for. In this case, defining the predicted direction of bias would be challenging in the absence of obvious sources of bias (particularly given how challenging it can be when sources of bias are evident within a domain).

In previous attempts at bias/indirectness-adjusted meta-analysis, the extent to each was assessed via a elicitation process using a number of experts. This approach is potentially subject to misclassification of the impact of bias on the basis of the results, as there is no way to ensure that results at a similar level of bias for confounding (for example) are being adjusted by the same amount. For example, if experts are influenced by the absolute effect estimate, then a result demonstrating a stronger protective effect may receive greater modification than a study at the same “risk” of confounding but with a more modest effect. This is particularly true where experts have preconceived ideas about the true effectiveness of the intervention.

In this case, separating the assessment of bias from the assignment of modifying values to each level of bias will limit the potential for. Alternatively, reasonable modifying distributions for each level of bias could be defined *a priori* by the study team, similar to how important confounders and co-interventions are defined in advance when performing ROBINS-I/ROBINS-E assessments.

In the meantime, the extent of the impact of knowledge of the results could be analysed using existing data. For example, a re-analysis of the data presented in a recent paper where both expert elicitation and structured

Ideally, the modifying values

Similar there are several examples of usmeta to define priors for other statistical terms in meta-analysis. For example, previous work has looked at het

Like other studies examining the prior distributions of heterogeneity across different topic areas

7.5 - Discussion

Talk about different ways to define bias terms, using expert-elicitation vs .³²¹

Also, talk about the fact that much of the bias-adjusted analysis work to date has been done on trials. Describe reasons for this but highlight that for triangulation, there needs to be better coverage of

Exclusion of studies at critical risk of bias

One potential - however, the estimation of an appropriate adjustment value for these studies would be substantially more challenging.

In any case, it represented a small number of studies.s

Type of bias (proportional vs additive) not shown on figure

In contrast

The graphics shown in Section @ref() do not show the type of bias (additive or proportional), only the direction. While this is acceptable, future work could assess, potentially by using different arrow types to denote the different types of bias.

A key strength of this project is the development of a standardised way of capturing bias domains and converting the.

Helper functions in the triangulate package convert risk-of-bias assessment

Challenges of real-life data

Compare and contrast with the nice example presented in the triangulation paper - realities of non-exemplars is that it is very hard to get this right. Also highlight the issue with assigning a direct of bias in many studies

I hope this presents step forward in how researchers think about and visualise triangulation at the result level, rather than simply saying that certain

Standardisation across study designs

Standardising across different measures of ex

Want best information across lipids, and so standardised across different measures

Fixing directions of effect - get everything pointing the same way

7.5.3 Alternative approaches

This method is also different from previous implementations of the bias-adjusted meta-analysis approach, in that for each domain for each study, the bias can be either proportional or. IN previous implementations, distributions for each source of bias (pre-domain-based tools) were elicited on both the additive and proportional scales.³¹⁵ Puts more pressure on the assessor to correct sum additive and proportional biases highlighted, and answer with only a single option.

E-values

Present as an alternative to putting priors on the level of bias due to confounding.

E-values work the other way around - get result then work out how strong a biasing factor would have to be to negate the result. IN other words, they represent an estimate ho how strong the total residual confounding would have to be in order to full explain away the observed result.³²²

One major problem is that they can only be used to evaluate negating biases, not other biases.

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

Heavily criticised in a range of papers.

Also, no useful to focus on a specific bias in the context of risk of bias domains

As an example, I calculated the E-value required to attenuate the harmful association observed between statin use and vascular dementia in the CPRD analysis ('**PLACEHOLDER**') presented in Chapter ???. Assuming an outcome prevalence of <15%, an E-value of 3.31 was identified

This calculation was performed using the XXX R package.

However, this approach is only useful if the only source of bias in a result is due to uncontrolled confounding (as assessed by a domain-based risk of bias tool).

Otherwise, its interpretation is more challenging, as the biasing effect is incorrectly assigned complete to confounding, thus overestimating the expected effect of confounding.

the contribution of confounding to the bias in a result, and while it may be applicable if the only bias assessed to

7.5.4 Need for new method

As the need and appetite for synthesis of different, future iterations of the risk of bias tools should take this into account. Simple things like harmonising the risk of bias levels across tools would be a good first move towards a more integrated suite of tools. Similarly, software that implements the risk-of-bias tools should both allow for direction of bias question and carefully consider how this data will be exported.

Can I suggest any empirical studies that need to be performed? Average strength of immortal time bias/etc?

While the limitations inherent to this method are important, and limit the , they reflect a more detailed appreciation of the complexities of risk of bias assessment and triangulation as whole (i.e. not applying expected directions of effects to a).

This approach also builds on existing approaches employed previously by “exploding out” the results of a systematic review/meta-analysis and considering the effects of bias in each individual result separately. This more granular approach is to be preferred over assigning a bias judgement and direction to the summary effect produced by a meta-analysis, which may mask the different biases applicable to each unique result.

7.5.5 Filling in the gaps

Triangulation as a extension of evidence synthesis.

7.6 Conclusion

Triangulation is a promising developing field, somewhat hamstrung by the limited understanding of the impact of biases at the meta-epidemiological level.

This chapter has demonstrated a repurposed framework for quantitative triangulation, exploiting existing methods in evidence synthesis and developments in the assessment of different risks of bias.

It also introduce a new way to visualise the biases at work in a particular result, rather than making assumptions about the type and directions of bias at the study design level as suggested by previous work on this topic.

Finally, it highlighted the limitations of current methods for triangulation in the absence of informative priors for the impact of biases across the

7.7 References

— Hold — Hold

8

Discussion

Contents

8.1	Lay summary	195
8.2	Chapter overview	195
8.3	Summary of principal findings	196
8.4	Overall strengths and limitations	196
8.5	Implications of this research	197
8.6	Reproducible research	197
8.7	Public involvement and engagement	198
8.8	Future work	198
8.9	Overall conclusions	200

8.1 Lay summary

This chapter

8.2 Chapter overview

In this chapter I will summarise the principal findings from my thesis, assess the unique contributions of this thesis to the field of research,

discuss the implications for both clinical and public health policy,
describe the overall strengths and limitations of this thesis as a whole,
suggest avenues for impactful future research based on my findings and experiences.

8.3 Summary of principal findings

8.4 Overall strengths and limitations

This thesis has used multiple sources of evidence to examine the relationship between blood lipids levels and cholesterol risk.

It has highlighted the importance of including preprints in systematic review, and provided a tool to enable researchers to easily do so.

It has analysed two previously unanalysed datasets, adding useful information to the evidence base

However, the thesis as a whole is subject to some strong 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 -³²³

This thesis is limited by the low response rate to data access requests, and the The lack of large, precise studies made comparison across different types of evidence difficult, while the low number of eligible studies for a

8.5 Implications of this research

8.5.1 Clinical practice

The clinical recommendations that can be drawn from the work in this thesis are limited, due to the

however, a key limiting factor is the absence of a detailed pathological mechanism for Alzheimer's disease. Additionally, vascular dementia

However, based on the available evidence

8.5.2 Public health

8.6 Reproducible research

Reproducible 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 with best practices

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

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.

8.7 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

8.8 Future work

8.8.1 Inclusion of preprints

Normalisation of inclusion of preprints in systematic reviews. Wider assessment of the evidence base that exists solely as preprinted literature ()

Studying of factors which influence eventual publication would also help to identify the extent and strength of publication bias by providing a

Preprints also represent a key opportunity to assess the impact of peer review on qualitative changes within

8.8.2 Evidence on vascular dementia

As discussed, there is an absence of evidence on vascular dementia across the evidence base, potentially driven by

large-scale GWAS of vascular dementia should be performed, to identify associated loci that future Mendelian randomisation studies can make use of. While this

is methodologically challenging, given the difficult in diagnosing “pure” vascular dementia, it would be a worthwhile endeavour, and would allow for the assessment of many genetically driven risk factors beyond those considered in this thesis.

8.8.3 Reviewing of Mendelian randomisation studies

As a field, and as noted in Chapter ??, the evidence synthesis theory and methods have not been sufficient developed for Mendelian randomisation studies. Future work should aim to validate search strategies for this study design, paying particular attention to the range of terms used to define it (e.g. genetic instrumental variable analysis.)

Work is ongoing in this field, and new methodological work such as the release of teh MR-STROBE guidelines will aid in the future assessing by helping to improve the reporting of Mendelian randomisation studies.

Problems with reporting of MR studies will hopefully

8.8.4 Empirical estimates of bias

Finally, substantial future work should be directed towards the conduct of meta-epidemiological studies to assess the impact of different bias in non-observational studies. While frameworks such as the one employed in Chapter 7 exist to allow for the production of bias-adjusted and triangulated estimates.

8.8.5 Imporved analytical designs

Use of the real-world evidence approach needs to be traded off against the important limitations that t

Talk about the problems with real world evidence, and highlight that even when using these techniques, you still don't get the correct answers.

Tie in with GDS comments on off-targets effect.

8.9 Overall conclusions

This thesis has made unique and novel contributions to the

Illustrating that commonly employed approaches to address

Definition of covariates adjusted for in the full-adjusted model. The codelists used to define the majority of these covariates were originally created for use in a previously published analysis.²⁷² while others were built on or adapted from previous published work,^{273–275}. Definition of covariates adjusted for in the full-adjusted model. The codelists used to define the majority of these covariates were originally created for use in a previously published analysis.²⁷² while others were built on or adapted from previous published work.^{273–275}

Covariates adjusted for. Covariates adjusted for.

Charlson index implemented using Read code lists.²⁷³ Code lists based on those by Taylor et al.²⁷⁴ Charlson index implemented using Read code lists.²⁷³ Code lists based on those by Taylor et al.²⁷⁴

Most recent of recorded value (current, former or never) or Read code indicating a recorded value. Code lists based on those by Wright et al.²⁷⁵ Most recent of recorded value (current, former or never) or Read code indicating a recorded value. Code lists based on those by Wright et al.²⁷⁵

Charlson index implemented using Read code lists.²⁷³ Code lists based on those by Taylor et al.²⁷⁴

8.9 - Overall conclusions

Table 8.1: Definition of covariates adjusted for in the full-adjusted model. The codelists used to define the majority of these covariates were originally created for use in a previously published analysis.²⁷² while others were built on or adapted from previous published work,^{273–275}.

Covariate	How was the covariate defined?
Previous history of coronary arterial disease	Presence of one or more relevant Read codes on record.
Previous history of coronary bypass surgery	Presence of one or more relevant Read codes on record.
Previous history of cerebrovascular disease (including stroke)	Presence of one or more relevant Read codes on record.
Chronic illness, including cancer and arthritis	Charlson index implemented using Read code lists. ²⁷³ Code lists based on those by Taylor et al. ²⁷⁴
Socioeconomic position	2010 English Index of Multiple Deprivation (IMD) at the twentile level, where 1 represents the least deprived and 20 the most deprived.
Consultation rate	Calculated by dividing the total number of clinic visits by the length of the patient record prior to the index date to give an average annual rate.
Alcohol status	Recorded value (current, former or never).
Smoking status	Most recent of recorded value (current, former or never) or Read code indicating a recorded value. Code lists based on those by Wright et al. ²⁷⁵
Body Mass Index	Recorded value if available, or a calculated value using the last recorded height and weight measurements. Measurements taken before the age of 25 were excluded to ensure adult measurements were used.
Peripheral arterial disease	Presence of one or more relevant Read codes on record.
Hypertension	Presence of one or more relevant Read codes on record.
Baseline total cholesterol	Continuous value recorded as test result ("enttype==163 & test_data1==3")
Baseline LDL cholesterol	Continuous value recorded as test result ("enttype==177 & test_data1==3")
Chronic kidney disease	Presence of one or more relevant Read codes on record.
Type 1 Diabetes	Presence of one or more relevant Read codes on record.
Type 2 Diabetes	Presence of one or more relevant Read codes on record.

Lasciate ogne speranza, voi ch'intrate. . .

9

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 (2019).
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. Li, G. *et al.* Serum cholesterol and risk of Alzheimer disease: A community-based cohort study. *Neurology* **65**, 1045–1050 (2005).
50. 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).
51. Mielke, M. M. *et al.* The 32-year relationship between cholesterol and dementia from midlife to late life. *Neurology* **75**, 1888–1895 (2010).
52. Tan, Z. S. *et al.* Plasma total cholesterol level as a risk factor for Alzheimer disease - The Framingham study. *Archives of Internal Medicine* **163**, 1053–+ (2003).
53. Mielke, M. M. *et al.* High total cholesterol levels in late life associated with a reduced risk of dementia. *Neurology* **64**, 1689–1695 (2005).
54. Reitz, C., Tang, M. X., Luchsinger, J. & Mayeux, R. Relation of plasma lipids to Alzheimer disease and vascular dementia. *Archives of Neurology* **61**, 705–714 (2004).
55. Moroney, J. T. Low-Density Lipoprotein Cholesterol and the Risk of Dementia With Stroke. *JAMA* **282**, 254 (1999).
56. 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 (2015).
57. 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).

58. 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).
59. Trompet, S. *et al.* Pravastatin and cognitive function in the elderly. Results of the PROSPER study. *Journal of Neurology* **257**, 85–90 (2010).
60. Davey Smith, G. & Hemani, G. Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics* **23**, R89–R98 (2014).
61. Greenland, S. An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology* **29**, 722–729 (2000).
62. Davies, N. M., Holmes, M. V. & Smith, G. D. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
63. Larsson, S. C. *et al.* Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* **359**, j5375 (2017).
64. Østergaard, S. D. *et al.* Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. *PLOS Medicine* **12**, e1001841 (2015).
65. Kim, J., Basak, J. M. & Holtzman, D. M. The Role of Apolipoprotein E in Alzheimer's Disease. *Neuron* **63**, 287–303 (2009).
66. 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).
67. Donnelly, C. A. *et al.* Four principles to make evidence synthesis more useful for policy. *Nature* **558**, 361–364 (2018).
68. 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).
69. 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).
70. 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).
71. 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
72. Lefebvre, C. *et al.* Chapter 4: Searching for and selecting studies. in *Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021)* 67–107 (Cochrane, 2020).

73. Committee on Publication Ethics (COPE). *Discussion Document: preprints*. (2018).
74. Vale, R. D. & Hyman, A. A. Priority of discovery in the life sciences. *eLife* **5**, e16931 (2016).
75. Fraser, N. *et al.* Preprinting a pandemic: The role of preprints in the COVID-19 pandemic. *bioRxiv* 2020.05.22.111294 (2020). doi:10.1101/2020.05.22.111294
76. Rosenthal, R. The file drawer problem and tolerance for null results. *Psychological Bulletin* **86**, 638–641 (1979).
77. Maslove, D. M. Medical PreprintsA Debate Worth Having. *JAMA* **319**, 443–444 (2018).
78. Schalkwyk, M. C. I. van, Hird, T. R., Maani, N., Petticrew, M. & Gilmore, A. B. The perils of preprints. *BMJ* **370**, m3111 (2020).
79. Mahood, Q., Eerd, D. V. & Irvin, E. Searching for grey literature for systematic reviews: Challenges and benefits. *Research Synthesis Methods* **5**, 221–234 (2014).
80. 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).
81. 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
82. Munafò, M. R. & Smith, G. D. Robust research needs many lines of evidence. *Nature* **553**, 399–401 (2018).
83. Page, M. J. *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).
84. 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).
85. Riley, R. D., Lambert, P. C. & Abo-Zaid, G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ* **340**, c221 (2010).
86. 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).
87. 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).
88. Clayton, J. A. Applying the new SABV (sex as a biological variable) policy to research and clinical care. *Physiology & Behavior* **187**, 2–5 (2018).
89. McCartney, M., Treadwell, J., Maskrey, N. & Lehman, R. Making evidence based medicine work for individual patients. *BMJ* **353**, i2452 (2016).

90. 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).
91. Riley, R. D. *et al.* Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: Statistical recommendations for conduct and planning. *Statistics in Medicine* **39**, 2115–2137 (2020).
92. 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).
93. Tugwell, P. & Knottnerus, J. A. Advantages of individual patient data analysis in systematic reviews. *Journal of Clinical Epidemiology* **63**, 233–234 (2010).
94. Nevitt, S. J. *et al.* Exploring changes over time and characteristics associated with data retrieval across individual participant data meta-analyses: Systematic review. *The BMJ* **357**, j1390 (2017).
95. 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).
96. 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).
97. Vines, T. H. *et al.* The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* **24**, 94–97 (2014).
98. Wartenberg, D. & Thompson, W. D. Privacy Versus Public Health: The Impact of Current Confidentiality Rules. *American Journal of Public Health* **100**, 407–412 (2010).
99. Bauermeister, S. *et al.* The Dementias Platform UK (DPUK) Data Portal. *European Journal of Epidemiology* **35**, 601–611 (2020).
100. Rawlinson, C. & Bloom, T. New preprint server for medical research. *BMJ* **365**, (2019).
101. Sever, R. *et al.* *bioRxiv: The preprint server for biology.* (Scientific Communication and Education, 2019). doi:10.1101/833400
102. Bealy, C. R Programming Humour. *Stack Exchange* (2013).
103. 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).
104. 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).

105. Ropensci/medrxivr: Access and search medRxiv and bioRxiv preprint data.
106. Wateridge, J. IT projects: A basis for success. *International Journal of Project Management* **13**, 169–172 (1995).
107. Abdill, R. J. & Blekhman, R. Rxivist.org: Sorting biology preprints using social media and readership metrics. *PLOS Biology* **17**, e3000269 (2019).
108. Rxivist: Explore popular biology preprints.
109. 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).
110. 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
111. 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).
112. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2019).
113. Git scraping: Track changes over time by scraping to a Git repository.
114. Rethlefsen, M. L. *et al.* PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic reviews* **10**, 1–19 (2021).
115. 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).
116. 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
117. 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).
118. 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).
119. 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).
120. Issue #380 - medrxivr: Accessing and searching medRxiv preprint data in R · ropensci/software-review · GitHub. *GitHub* · *ropensci/software-review*

121. Abdill, R. J. & Blekhman, R. Tracking the popularity and outcomes of all bioRxiv preprints. *eLife* **8**, e45133 (2019).
122. 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
123. Song, F. *et al.* Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment* **14**, (2010).
124. 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
125. Cochrane, A. L. 1931-1971: A critical review, with particular reference to the medical profession. *Medicines for the year 2000* (1979).
126. 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
127. 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).
128. 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).
129. Hupe, M. EndNote X9. *Journal of Electronic Resources in Medical Libraries* **16**, 117–119 (2019).
130. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews* **5**, 210 (2016).
131. Organization(WHO), W. H. *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research.* (World Health Organization, 1993).
132. Román, G. C. *et al.* Vascular dementia: Diagnostic criteria for research studies: Report of the NINDS-AIREN International Workshop. *Neurology* **43**, 250–250 (1993).
133. Edition, F. & others. Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc* **21**, (2013).
134. 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).
135. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **20**, 37–46 (1960).

136. 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).
137. Brennan, P. & Silman, A. Statistical methods for assessing observer variability in clinical measures. *BMJ : British Medical Journal* **304**, 1491–1494 (1992).
138. McHugh, M. L. Interrater reliability: The kappa statistic. *Biochemia Medica* **22**, 276–282 (2012).
139. Higgins, J. P., Li, T. & Deeks, J. J. Section 6.5.2.10 : Combining groups. in *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated August 2019)* (eds. Higgins, J. P. T. et al.) (Cochrane, 2019).
140. 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).
141. McKenzie, J. E. & Brennan, S. E. Section 12.1: Why a meta-analysis of effect estimates may not be possible. in *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated August 2019)* (eds. Higgins, J. P. T. et al.) (Cochrane, 2019).
142. Sterne, J. A. *et al.* ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355**, i4919 (2016).
143. Campbell, D. T. Factors relevant to the validity of experiments in social settings. *Psychological Bulletin* **54**, 297–312 (1957).
144. Jüni, P., Altman, D. G. & Egger, M. Assessing the quality of controlled clinical trials. *BMJ* **323**, 42–46 (2001).
145. Sterne, J. A. C. *et al.* RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, (2019).
146. 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).
147. French, C. E. *et al.* Cannabis use and the risk of tuberculosis: A systematic review. *BMC Public Health* **19**, 1006 (2019).
148. Wells, G. A. *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. (2000).
149. 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).
150. Spiga, F. *et al.* Tools for the assessment of quality and risk of bias in Mendelian randomization studies: A systematic review. *medRxiv* 2021.10.21.21265126 (2021). doi:10.1101/2021.10.21.21265126

151. Risk Of Bias due to Missing Evidence (ROB-ME).
152. Saran, A. & White, H. Evidence and gap maps: A comparison of different approaches. *Campbell Systematic Reviews* **14**, 1–38 (2018).
153. 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).
154. Issue #102 - Feature request: Paired forest/risk-of-bias plot. *GitHub* · [mcguinlu/robvis](https://github.com/mcguinlu/robvis)
155. 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).
156. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20 536 high-risk individuals: A randomised placebocontrolled trial. *The Lancet* **360**, 7–22 (2002).
157. Ridker, P. M. *et al.* Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein. *New England Journal of Medicine* **359**, 2195–2207 (2008).
158. Ancelin, M.-L. *et al.* Lipid lowering agents, cognitive decline, and dementia: The three-city study. *Journal of Alzheimer's Disease* **30**, 629–637 (2012).
159. Arvanitakis, Z. *et al.* Statins, incident Alzheimer disease, change in cognitive function, and neuropathology. *Neurology* **70**, 1795–1802 (2008).
160. Bettermann, K. *et al.* Statins, Risk of Dementia, and Cognitive Function: Secondary Analysis of the Ginkgo Evaluation of Memory Study. *Journal of Stroke and Cerebrovascular Diseases* **21**, 436–444 (2012).
161. Beydoun, M. A. *et al.* Statins and serum cholesterol's associations with incident dementia and mild cognitive impairment. *Journal of Epidemiology & Community Health* **65**, 949–957 (2011).
162. Chao, T. F. *et al.* Statins and the risk of dementia in patients with atrial fibrillation: A nationwide population-based cohort study. *International Journal of Cardiology* **196**, 91–97 (2015).
163. Chen, J. M. *et al.* Effects of statins on incident dementia in patients with type 2 DM: A population-based retrospective cohort study in Taiwan. *PLoS ONE [Electronic Resource]* **9**, e88434 (2014).
164. Chitnis, A. S. *et al.* Use of statins and risk of dementia in heart failure: A retrospective cohort study. *Drugs & Aging* **32**, 743–754 (2015).
165. Chou, C. Y., Chou, Y. C., Chou, Y. J., Yang, Y. F. & Huang, N. Statin use and incident dementia: A nationwide cohort study of Taiwan. *International Journal of Cardiology* **173**, 305–310 (2014).
166. Chuang, C. S., Lin, C. L., Lin, M. C., Sung, F. C. & Kao, C. H. Decreased prevalence of dementia associated with statins: A national population-based study. *European Journal of Neurology* **22**, 912–918 (2015).

167. Cramer, C., Haan, M. N., Galea, S., Langa, K. M. & Kalbfleisch, J. D. Use of statins and incidence of dementia and cognitive impairment without dementia in a cohort study. *Neurology* **71**, 344–350 (2008).
168. Gnjidic, D. *et al.* STATIN THERAPY AND DEMENTIA IN OLDER ADULTS: ROLE OF DISEASE SEVERITY AND MULTIMORBIDITY. *Journal of the American Geriatrics Society* **64**, 223–224 (2016).
169. Haag, M. D., Hofman, A., Koudstaal, P. J., Stricker, B. H. & Breteler, M. M. Statins are associated with a reduced risk of Alzheimer disease regardless of lipophilicity. The Rotterdam Study. *Journal of Neurology, Neurosurgery & Psychiatry* **80**, 13–17 (2009).
170. Hendrie, H. C. *et al.* Statin use, incident dementia and alzheimer disease in elderly african americans. *Ethnicity & Disease* **25**, 345–354 (2015).
171. Hippisley-Cox, J. & Coupland, C. Unintended effects of statins in men and women in England and Wales: Population based cohort study using the QResearch database. *BMJ (Clinical research ed.)* **340**, c2197 (2010).
172. Jick, H., Zornberg, G. L., Jick, S. S., Seshadri, S. & Drachman, D. A. Statins and the risk of dementia. *Lancet (London, England)* **356**, 1627–1631 (2000).
173. Li, G. *et al.* Statin therapy and risk of dementia in the elderly: A community-based prospective cohort study. *Neurology* **63**, 1624–1628 (2004).
174. Li, G. *et al.* Age-varying association between statin use and incident Alzheimer's disease. *Journal of the American Geriatrics Society* **58**, 1311–1317 (2010).
175. Liao, M. T., Tsai, C. T. & Lin, J. L. Statins reduce the incidence of dementia in patients with atrial fibrillation: A nationwide cohort study. *European Heart Journal* **34**, 740–741 (2013).
176. Liu, J. M., Chen, T. H., Chuang, H. C., Wu, C. T. & Hsu, R. J. Statin reduces the risk of dementia in diabetic patients receiving androgen deprivation therapy for prostate cancer. *Prostate Cancer & Prostatic Diseases* **22**, 276–283 (2019).
177. Pan, M. L., Hsu, C. C., Chen, Y. M., Yu, H. K. & Hu, G. C. Statin use and the risk of dementia in patients with stroke: A nationwide population-based cohort study. *Journal of Stroke & Cerebrovascular Diseases* **27**, 3001–3007 (2018).
178. Parikh, N. M. *et al.* Risk factors for dementia in patients over 65 with diabetes. *International Journal of Geriatric Psychiatry* **26**, 749–757 (2011).
179. Rea, T. D. *et al.* Statin use and the risk of incident dementia: The Cardiovascular Health Study. *Archives of Neurology* **62**, 1047–1051 (2005).
180. Redelmeier, D. A., Manzoor, F. & Thiruchelvam, D. Association between statin use and risk of dementia after a concussion. *JAMA Neurology* **20**, 20 (2019).
181. Reitz, C. *et al.* Association of higher levels of high-density lipoprotein cholesterol in elderly individuals and lower risk of late-onset Alzheimer disease. *Archives of Neurology* **67**, 1491–1497 (2010).

182. 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).
183. Solomon, A. *et al.* Serum cholesterol changes after midlife and late-life cognition: Twenty-one-year follow-up study. *Neurology* **68**, 751–756 (2007).
184. Sparks, D. L. *et al.* Reduced risk of incident AD with elective statin use in a clinical trial cohort. *Current alzheimer research* **5**, 416–421 (2008).
185. Szwast, S. J. *et al.* Association of statin use with cognitive decline in elderly African Americans. *Neurology* **69**, 1873–1880 (2007).
186. Yang, Y. H. *et al.* Statins reduces the risk of dementia in patients with late-onset depression: A retrospective cohort study. *PLoS ONE [Electronic Resource]* **10**, e0137914 (2015).
187. Zamrini, E., McGwin, G. & Roseman, J. M. Association between statin use and Alzheimer's disease. *Neuroepidemiology* **23**, 94–98 (2004).
188. Zandi, P. P. *et al.* Do statins reduce risk of incident dementia and Alzheimer disease? The Cache County Study. *Archives of General Psychiatry* **62**, 217–224 (2005).
189. Ancelin, M. L. *et al.* Sex differences in the associations between lipid levels and incident dementia. *Journal of Alzheimer's Disease* **34**, 519–528 (2013).
190. Batty, G. D., Russ, T. C., Starr, J. M., Stamatakis, E. & Kivimaki, M. Modifiable cardiovascular disease risk factors as predictors of dementia death: Pooling of ten general population-based cohort studies. *Journal of Negative Results in Biomedicine* **13**, 8 (2014).
191. Benn, M., Nordestgaard, B. G., Frikke-Schmidt, R. & Tybjaerg-Hansen, A. Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. *BMJ (Clinical research ed.)* **357**, j1648 (2017).
192. Bruce, D. G., Davis, W. A. & Davis, T. M. E. Low serum HDL-cholesterol concentrations in mid-life predict late-life cognitive impairment in type 2 diabetes: The Fremantle diabetes study. *Journal of Diabetes & its Complications* **31**, 945–947 (2017).
193. Chiang, C. J. *et al.* Midlife risk factors for subtypes of dementia: A nested case-control study in Taiwan. *American Journal of Geriatric Psychiatry* **15**, 762–771 (2007).
194. Dodge, H. H., Chang, C. C. H., Kamboh, I. M. & Ganguli, M. Risk of Alzheimer's disease incidence attributable to vascular disease in the population. *Alzheimer's and Dementia* **7**, 356–360 (2011).
195. Forti, P. *et al.* Metabolic syndrome and risk of dementia in older adults. *Journal of the American Geriatrics Society* **58**, 487–492 (2010).

196. Gottesman, R. F. *et al.* Associations between midlife vascular risk factors and 25-Year incident dementia in the atherosclerosis risk in communities (ARIC) cohort. *JAMA Neurology* **74**, 1246–1254 (2017).
197. Gustafson, D. R. *et al.* Epidemiological evidence for lipid-based dementia prevention: The Lipididiet approach. *Journal of Molecular Neuroscience* **1**, S45 (2012).
198. Hayden, K. M. *et al.* Vascular risk factors for incident Alzheimer disease and vascular dementia - The Cache County study. *Alzheimer Disease & Associated Disorders* **20**, 93–100 (2006).
199. Kimm, H. *et al.* Mid-life and late-life vascular risk factors and dementia in Korean men and women. *Archives of Gerontology & Geriatrics* **52**, e117–122 (2011).
200. Kivipelto, M. *et al.* Midlife vascular risk factors and Alzheimer's disease in later life: Longitudinal, population based study. *BMJ (Clinical research ed.)* **322**, 1447–1451 (2001).
201. 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).
202. Kuo, S. C. *et al.* Association between comorbidities and dementia in diabetes mellitus patients: Population-based retrospective cohort study. *Journal of Diabetes & its Complications* **29**, 1071–1076 (2015).
203. Mainous, 3rd., 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).
204. Mielke, M. M. *et al.* High total cholesterol levels in late life associated with a reduced risk of dementia. *Neurology* **64**, 1689–1695 (2005).
205. Mielke, M. M. *et al.* The 32-year relationship between cholesterol and dementia from midlife to late life. *Neurology* **75**, 1888–1895 (2010).
206. Mielke, M. *et al.* Serum ceramides predict incident alzheimer's disease: The women's health and aging study (whas) II. *Alzheimer's and Dementia* **1**, S607 (2011).
207. Muller, M. *et al.* Metabolic syndrome and dementia risk in a multiethnic elderly cohort. *Dementia & Geriatric Cognitive Disorders* **24**, 185–192 (2007).
208. Noale, M. *et al.* Incidence of dementia: Evidence for an effect modification by gender. The ILSA Study. *International Psychogeriatrics* **25**, 1867–1876 (2013).
209. Notkola, I. L. *et al.* Serum total cholesterol, apolipoprotein E epsilon 4 allele, and Alzheimer's disease. *Neuroepidemiology* **17**, 14–20 (1998).
210. Peters, R. *et al.* Cardiovascular and biochemical risk factors for incident dementia in the Hypertension in the Very Elderly Trial. *Journal of Hypertension* **27**, 2055–2062 (2009).

211. Raffaitin, C. *et al.* Metabolic syndrome and risk for incident Alzheimer's disease or vascular dementia: The Three-City Study. *Diabetes Care* **32**, 169–174 (2009).
212. Rantanen, K. *et al.* Cardiovascular risk factors and glucose tolerance in midlife and risk of cognitive disorders in old age up to a 49-year follow-up of the Helsinki businessmen study. *Annals of Medicine* **49**, 462–469 (2017).
213. Ronnemaa, E., Zethelius, B., Lannfelt, L. & Kilander, L. Vascular risk factors and dementia: 40-year follow-up of a population-based cohort. *Dementia & Geriatric Cognitive Disorders* **31**, 460–466 (2011).
214. 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 / Public Library of Science* **14**, e1002265 (2017).
215. 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 & Geriatric Cognitive Disorders* **28**, 75–80 (2009).
216. Solomon, A. *et al.* Lipid-lowering treatment is related to decreased risk of dementia: A population-based study (FINRISK). *Neurodegenerative Diseases* **7**, 180–182 (2010).
217. Strand, B. H. *et al.* Midlife vascular risk factors and their association with dementia deaths: Results from a Norwegian prospective study followed up for 35 years. *Journal of the Neurological Sciences* **324**, 124–130 (2013).
218. Su, B. *et al.* [P3]: EXPLORING LATE-LIFE RISK FACTORS OF ALZHEIMER's DISEASE AND OTHER AGE-RELATED DEMENTIAS IN CPRD. *Alzheimer's & Dementia* **13**, P1178–P1179 (2017).
219. Svensson, T. *et al.* The association between midlife serum high-density lipoprotein and mild cognitive impairment and dementia after 19 years of follow-up. *Transl Psychiatry Psychiatry* **9**, 26 (2019).
220. Tynkkynen, J. *et al.* Apolipoproteins and HDL cholesterol do not associate with the risk of future dementia and Alzheimer's disease: The National Finnish population study (FINRISK). *Age (Melbourne, Vic.)* **38**, 465–473 (2016).
221. Tynkkynen, J. *et al.* Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts. *Alzheimer's & Dementia* **14**, 723–733 (2018).
222. Wang, K.-C. *et al.* Risk of Alzheimer's disease in relation to diabetes: A population-based cohort study. *Neuroepidemiology* **38**, 237–244 (2012).
223. 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).

224. Yamada, M. *et al.* Incidence and risks of dementia in Japanese women: Radiation Effects Research Foundation Adult Health Study. *Journal of the Neurological Sciences* **283**, 57–61 (2009).
225. Yoshitake, T. *et al.* Incidence and risk factors of vascular dementia and Alzheimer's disease in a defined elderly Japanese population: The Hisayama Study. *Neurology* **45**, 1161–1168 (1995).
226. Zimetbaum, P. *et al.* Plasma lipids and lipoproteins and the incidence of cardiovascular disease in the very elderly. The Bronx Aging Study. *Arteriosclerosis & Thrombosis* **12**, 416–423 (1992).
227. Andrews, S. J. *et al.* Causal Associations Between Modifiable Risk Factors and the Alzheimer's Phenome. *Annals of Neurology* **89**, 54–65 (2021).
228. Burgess, S. & Davey Smith, G. Mendelian randomization implicates high-density lipoprotein cholesterol-associated mechanisms in etiology of age-related macular degeneration. *Ophthalmology* **124**, 1165–1174 (2017).
229. Larsson, S. C. *et al.* Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* **359**, j5375 (2017).
230. Mukherjee, S. *et al.* Vascular disease, vascular risk factors and risk of late-onset Alzheimer's disease: Mendelian randomization analyses in the combined adgc dataset. *Alzheimer's and Dementia* **1**, P694 (2013).
231. Ostergaard, S. D. *et al.* Associations between potentially modifiable risk factors and Alzheimer disease: A Mendelian randomization study. *European Neuropsychopharmacology* **27** (Supplement 2), S166–S167 (2017).
232. So, H.-C., Chau, C. K. & Zhao, K. Exploring repositioning opportunities and side-effects of statins: A Mendelian randomization study of HMG-CoA reductase inhibition with 55 complex traits. 170241 (2017). doi:10.1101/170241
233. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nature Communications* **9**, 224 (2018).
234. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *bioRxiv* (2017). doi:10.1101/168674
235. Andrews, S. J., Marcora, E. & Goate, A. Causal associations between potentially modifiable risk factors and the Alzheimer's phenotype: A Mendelian randomization study. 689752 (2019). doi:10.1101/689752
236. Sterne, J. A. C. *et al.* Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* **343**, d4002 (2011).
237. Shepherd, J. *et al.* Pravastatin in elderly individuals at risk of vascular disease (PROSPER): A randomised controlled trial. *The Lancet* **360**, 1623–1630 (2002).
238. 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

239. Yamada, M. *et al.* Incidence and risks of dementia in Japanese women: Radiation Effects Research Foundation Adult Health Study. *Journal of the Neurological Sciences* **283**, 57–61 (2009).
240. Yamada, M. *et al.* Incidence and risks of dementia in Japanese women: The adult health study. *Journal of the Neurological Sciences* **283**, 306 (2009).
241. 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).
242. 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
243. 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).
244. 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).
245. Power, M. C., Weuve, J., Sharrett, A. R., Blacker, D. & Gottesman, R. F. Statins, cognition, and dementiasystematic review and methodological commentary. *Nature reviews. Neurology* **11**, 220–229 (2015).
246. Aromataris, E. *et al.* Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare* **13**, 132–140 (2015).
247. Smith, V., Devane, D., Begley, C. M. & Clarke, M. Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology* **11**, 15 (2011).
248. Sohani, 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).
249. Larsson, S. C. *et al.* Modifiable pathways in Alzheimer's disease: Mendelian randomisation analysis. *BMJ* **359**, j5375 (2017).
250. 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).
251. 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).
252. Benn, M. Re: Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. *BMJ* (2017).

253. Andrews, S. J. *et al.* Causal associations between potentially modifiable risk factors and the Alzheimer's disease phenotype: A Mendelian randomization study. *bioRxiv* (2019). doi:10.1101/689752
254. 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).
255. 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).
256. 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
257. Hsieh, C.-Y. *et al.* Taiwan's National Health Insurance Research Database: Past and future. *Clinical Epidemiology* **11**, 349–358 (2019).
258. 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).
259. Wilkinson, T. *et al.* Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimer's & Dementia* **14**, 1038–1051 (2018).
260. Jeyaraman, M. M. *et al.* Methodologically rigorous risk of bias tools for nonrandomized studies had low reliability and high evaluator burden. *Journal of Clinical Epidemiology* **128**, 140–147 (2020).
261. Dorn, H. F. Philosophy of Inferences from Retrospective Studies. *American Journal of Public Health and the Nations Health* **43**, 677–683 (1953).
262. Suissa, S. Immortal Time Bias in Pharmacoepidemiology. *American Journal of Epidemiology* **167**, 492–499 (2008).
263. 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 Alzheimer's and other neurodegenerative diseases? Protocol for an observational cohort study in the UK Clinical Practice Research Datalink. *BMJ Open* **6**, (2016).
264. 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).
265. Herrett, E. *et al.* Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology* **44**, 827–836 (2015).

266. 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).
267. Wood, L. & Coulson, R. Revitalizing the general practice research database: Plans, challenges, and opportunities. *Pharmacoepidemiology and drug safety* **10**, 379–383 (2001).
268. 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).
269. Booth, N. What are the Read Codes? *Health Libraries Review* **11**, 177–182 (1994).
270. Wolf, A. *et al.* Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *International Journal of Epidemiology* **48**, 1740–1740g (2019).
271. Wishart, D. S. *et al.* DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research* **46**, D1074–D1082 (2017).
272. Walker, V. M., Davies, N. M., Martin, R. M. & Kehoe, P. G. Comparison of Antihypertensive Drug Classes for Dementia Prevention. *Epidemiology* **31**, 852–859 (2020).
273. Khan, N. F., Perera, R., Harper, S. & Rose, P. W. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. *BMC Family Practice* **11**, 1 (2010).
274. Taylor, G. M. J. *et al.* Effectiveness of varenicline versus nicotine replacement therapy on long-term smoking cessation in primary care: A prospective, cohort study of electronic medical records. *The Lancet* **388**, S107 (2016).
275. Wright, A. K. *et al.* Life Expectancy and Cause-Specific Mortality in Type 2 Diabetes: A Population-Based Cohort Study Quantifying Relationships in Ethnic Subgroups. *Diabetes Care* **40**, 338–345 (2017).
276. 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).
277. 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).
278. Sterne, J. A. C. *et al.* Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ* **338**, b2393 (2009).
279. 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).

280. 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).
281. 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).
282. 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).
283. 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).
284. Pigott, T. D. A review of methods for missing data. *Educational research and evaluation* **7**, 353–383 (2001).
285. 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).
286. 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).
287. 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).
288. Herrett, E. *et al.* Statin treatment and muscle symptoms: Series of randomised, placebo controlled n-of-1 trials. *BMJ* **372**, n135 (2021).
289. Macedo, A. F., Douglas, I., Smeeth, L., Forbes, H. & Ebrahim, S. Statins and the risk of type 2 diabetes mellitus: Cohort study using the UK clinical practice pesearch datalink. *BMC Cardiovascular Disorders* **14**, 85 (2014).
290. Smit, R. A. J. *et al.* Statin-induced LDL cholesterol response and type 2 diabetes: A bidirectional two-sample Mendelian randomization study. *The Pharmacogenomics Journal* **20**, 462–470 (2020).
291. 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).
292. 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).

293. Carbonari, D. M. *et al.* Use of demographic and pharmacy data to identify patients included within both the Clinical Practice Research Datalink (CPRD) and The Health Improvement Network (THIN). *Pharmacoepidemiology and Drug Safety* **24**, 999–1003 (2015).
294. 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).
295. 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).
296. 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).
297. Taylor, F. *et al.* Statins for the primary prevention of cardiovascular disease. *The Cochrane Database of Systematic Reviews* **2013**, CD004816 (2013).
298. Newman, C. B. *et al.* Statin Safety and Associated Adverse Events: A Scientific Statement From the American Heart Association. *Arteriosclerosis, Thrombosis, and Vascular Biology* **39**, e38–e81 (2019).
299. Shepardson, N. E. Cholesterol Level and Statin Use in Alzheimer Disease: II. Review of Human Trials and Recommendations. *Archives of Neurology* **68**, 1385 (2011).
300. Porta, M. *A dictionary of epidemiology*. (Oxford University Press, 2014).
301. Bennet, A. M. *et al.* Association of apolipoprotein E genotypes with lipid levels and coronary risk. *JAMA* **298**, 1300–1311 (2007).
302. Goldacre, B., Morton, C. E. & DeVito, N. J. Why researchers should share their analytic code. *BMJ* **367**, l6365 (2019).
303. Levis, B., Hattle, M. & Riley, R. D. PRIME-IPD SERIES Part 2. Retrieving, checking, and harmonizing data are underappreciated challenges in individual participant data meta-analyses. *Journal of Clinical Epidemiology* **136**, 221–223 (2021).
304. Wang, L. *et al.* Assessing robustness of hazard ratio estimates to outcome misclassification in longitudinal panel studies with application to Alzheimer's disease. *PLOS ONE* **12**, e0190107 (2017).
305. Abo-Zaid, G. *et al.* Individual participant data meta-analyses should not ignore clustering. *Journal of Clinical Epidemiology* **66**, 865–873.e4 (2013).
306. McGuinness, L. A. & Sheppard, A. L. A descriptive analysis of the data availability statements accompanying medRxiv preprints and a comparison with their published counterparts. *MetaArXiv* (2020). doi:10.31222/osf.io/p75xe

307. Riboli, E. The EPIC Project: Rationale and study design. European Prospective Investigation into Cancer and Nutrition. *International Journal of Epidemiology* **26**, 6S–14 (1997).
308. Riboli, E. *et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): Study populations and data collection. *Public Health Nutrition* **5**, 1113–1124 (2002).
309. 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).
310. Krawczyk, M. & Reuben, E. (Un)Available upon Request: Field Experiment on Researchers' Willingness to Share Supplementary Materials. *Accountability in Research* **19**, 175–186 (2012).
311. Miyakawa, T. No raw data, no science: Another possible source of the reproducibility crisis. *Molecular Brain* **13**, 24 (2020).
312. Naudet, F. *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* k400 (2018). doi:10.1136/bmj.k400
313. Jaspers, G. J. & Degraeuwe, P. L. A failed attempt to conduct an individual patient data meta-analysis. *Systematic Reviews* **3**, 97 (2014).
314. Hill, A. B. The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* **58**, 295–300 (1965).
315. Turner, R. M., Spiegelhalter, D. J., Smith, G. C. S. & Thompson, S. G. Bias modelling in evidence synthesis. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)* **172**, 21–47 (2009).
316. Chao, T.-F. *et al.* Statins and the risk of dementia in patients with atrial fibrillation: A nationwide population-based cohort study. *International Journal of Cardiology* **196**, 91–97 (2015).
317. Williams, D. M., Finan, C., Schmidt, A. F., Burgess, S. & Hingorani, A. D. Lipid lowering and Alzheimer disease risk: A mendelian randomization study. *Annals of Neurology* **87**, 30–39 (2020).
318. Stone, J. C., Glass, K., Munn, Z., Tugwell, P. & Doi, S. A. R. Comparison of bias adjustment methods in meta-analysis suggests that quality effects modeling may have less limitations than other approaches. *Journal of Clinical Epidemiology* **117**, 36–45 (2020).
319. Amer, M. A. *et al.* A meta-epidemiological study of bias in randomized clinical trials of open and laparoscopic surgery. *British Journal of Surgery* **108**, 477–483 (2021).
320. Page, M. J. *et al.* Empirical Evidence of Study Design Biases in Randomized Trials: Systematic Review of Meta-Epidemiological Studies. *PLOS ONE* **11**, e0159267 (2016).

321. Rhodes, K. M. *et al.* Adjusting trial results for biases in meta-analysis: Combining data-based evidence on bias with detailed trial assessment. *Journal of the Royal Statistical Society. Series A, (Statistics in Society)* **183**, 193–209 (2020).
322. VanderWeele, T. J. & Ding, P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of Internal Medicine* **167**, 268–274 (2017).
323. Yusuf, S., Collins, R. & Peto, R. Why do we need some large, simple randomized trials? *Statistics in Medicine* **3**, 409–420 (1984).
324. Wilson, G. *et al.* Best Practices for Scientific Computing. *PLOS Biology* **12**, e1001745 (2014).
325. Wilson, G. *et al.* Good enough practices in scientific computing. *PLOS Computational Biology* **13**, e1005510 (2017).
326. 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).
327. 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 1

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 2

A.2.1 Code for publication rate analysis

```
med_res <-
  # Use snapshot of published status from July 2021
  medrxivr::mx_search(
    medrxivr::mx_snapshot("ccedfb8a44304b9fba4e3ba518a8ce4ed2294770"),
    query = "*",
    from_date = "2019-07-01",
    to_date = "2019-08-01"
  ) %>%
  # Create indicator to show which records have been published
  mutate(pub_ind = ifelse(published == "NA", 0, 1)) %>%
  # Group by indicator variable and count
  group_by(pub_ind) %>%
  count()
```

A.3 Chapter ??

A.3.1 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	or/1-17	407352
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

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

#	Search term	Hits
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	("fibrac 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
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

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

#	Search term	Hits
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
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 subsitut* 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

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

#	Search term	Hits
100	98 or 99	1738
101	RANDOMIZED CONTROLLED TRIALS AS TOPIC/	124147
102	(RCT? or (randomized 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
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.3.2 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.3.3 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.3.4 Calculating Gwet's AC1

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:

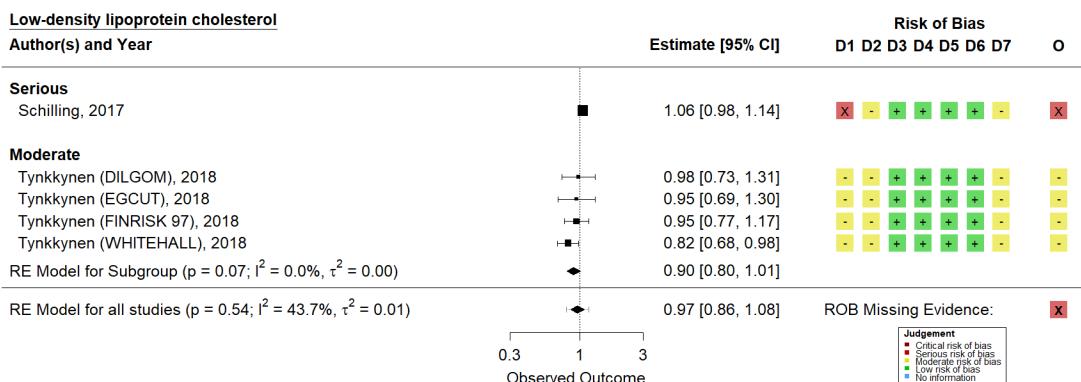
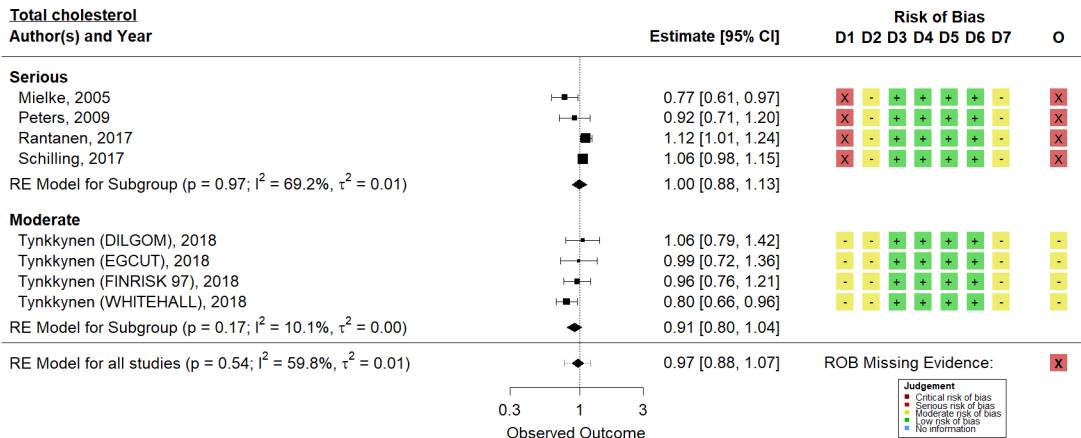
$$AC1 = \frac{\frac{A+D}{N} - e(\gamma)}{1 - e(\gamma)} \quad (\text{A.1})$$

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

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

A.3.5 MR risk of bias tool

A.3.6 Forest plots for lipid fractions



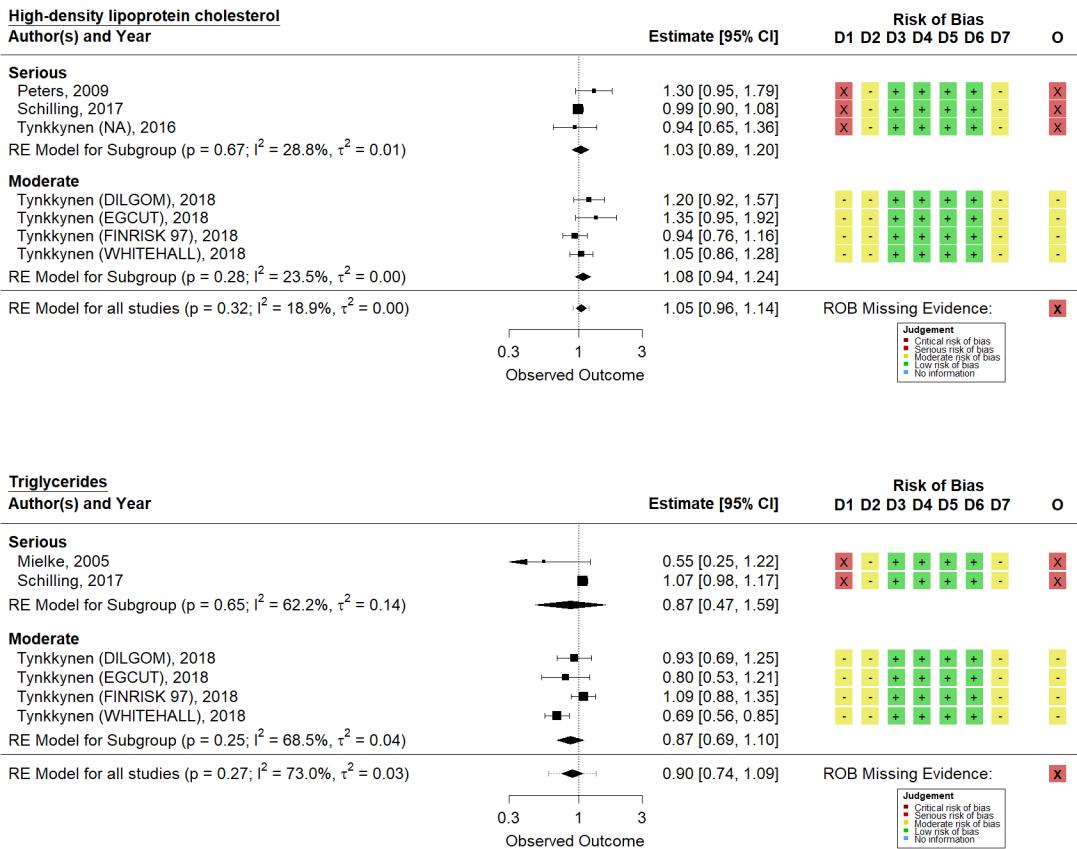


Table A.3: Tool used to assess risk of bias in Mendelian randomisation studies, adapted from that developed by Mamluk et al.¹⁴⁹

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	F»10

Table A.3: Tool used to assess risk of bias in Mendelian randomisation studies, adapted from that developed by Mamluk et al.¹⁴⁹ (*continued*)

Bias domain	Question	High	Moderate	Low
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
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

Table A.3: Tool used to assess risk of bias in Mendelian randomisation studies, adapted from that developed by Mamluk et al.¹⁴⁹ (*continued*)

Bias domain	Question	High	Moderate	Low
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.4 Chapter 5

A.4.1 Amendments to protocol

A.4.2 Code lists

A.5 Chapter 6

A.5.1 Email and documents sent to potential collaborators

A.5.2 Section 2

A.6 Chapter 8

A.6.1 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-DiagrammeR?, R-dplyr?, R-flextable?, R-ggdag?, R-ggplot2?

All projects in these thesis attempt to conform to minimal best practices for research computing.^{324,325}

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,¹⁴² 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.^{326,327,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*

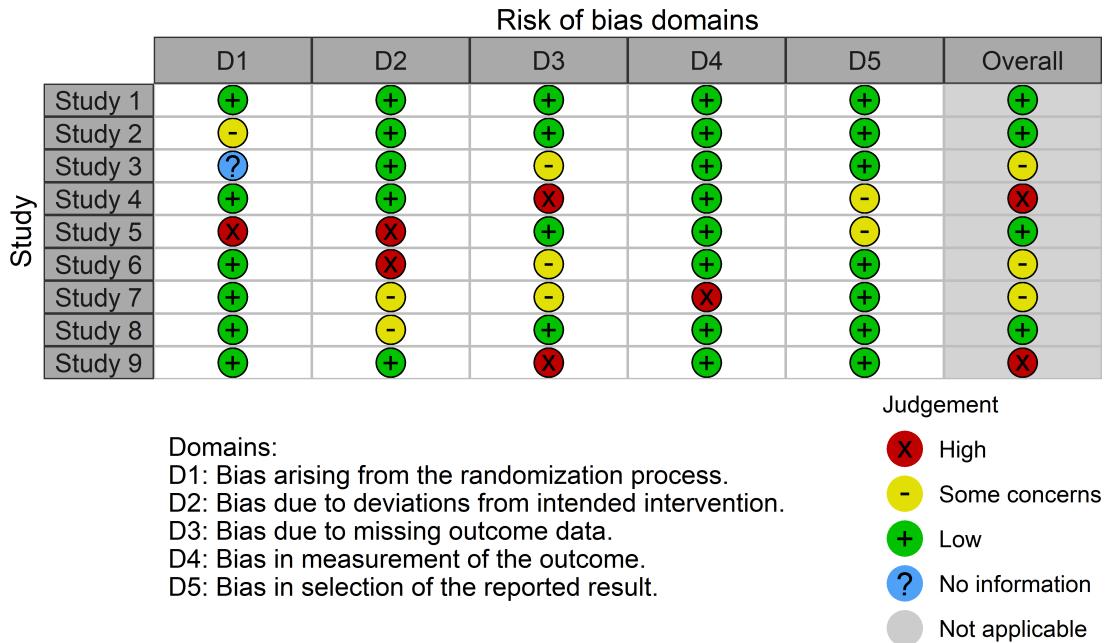


Figure B.1: Example risk of bias traffic light plot created using ‘robvis’

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",
                  psize = 15)
```

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

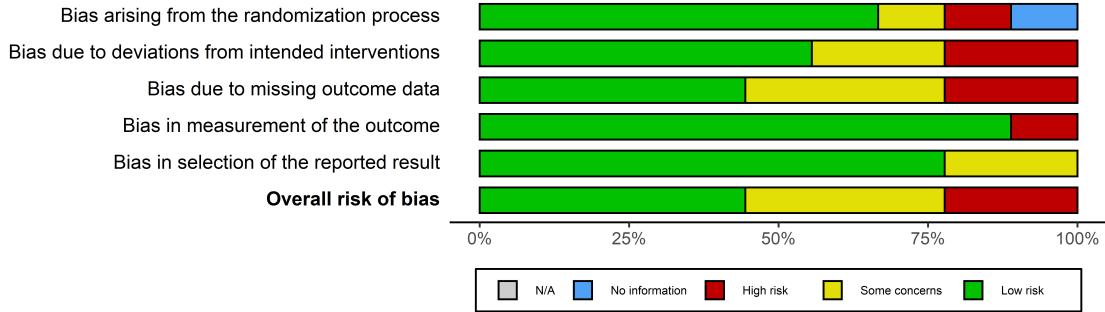


Figure B.2: Example risk of bias summary plot created using ‘robvis’ and the example ROB2 dataset

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"')

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
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")‘
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 15700 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

B.4 Copies of papers arising from this thesis

The following pages contain copies of papers arising from work performed as part of 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 than 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)
```

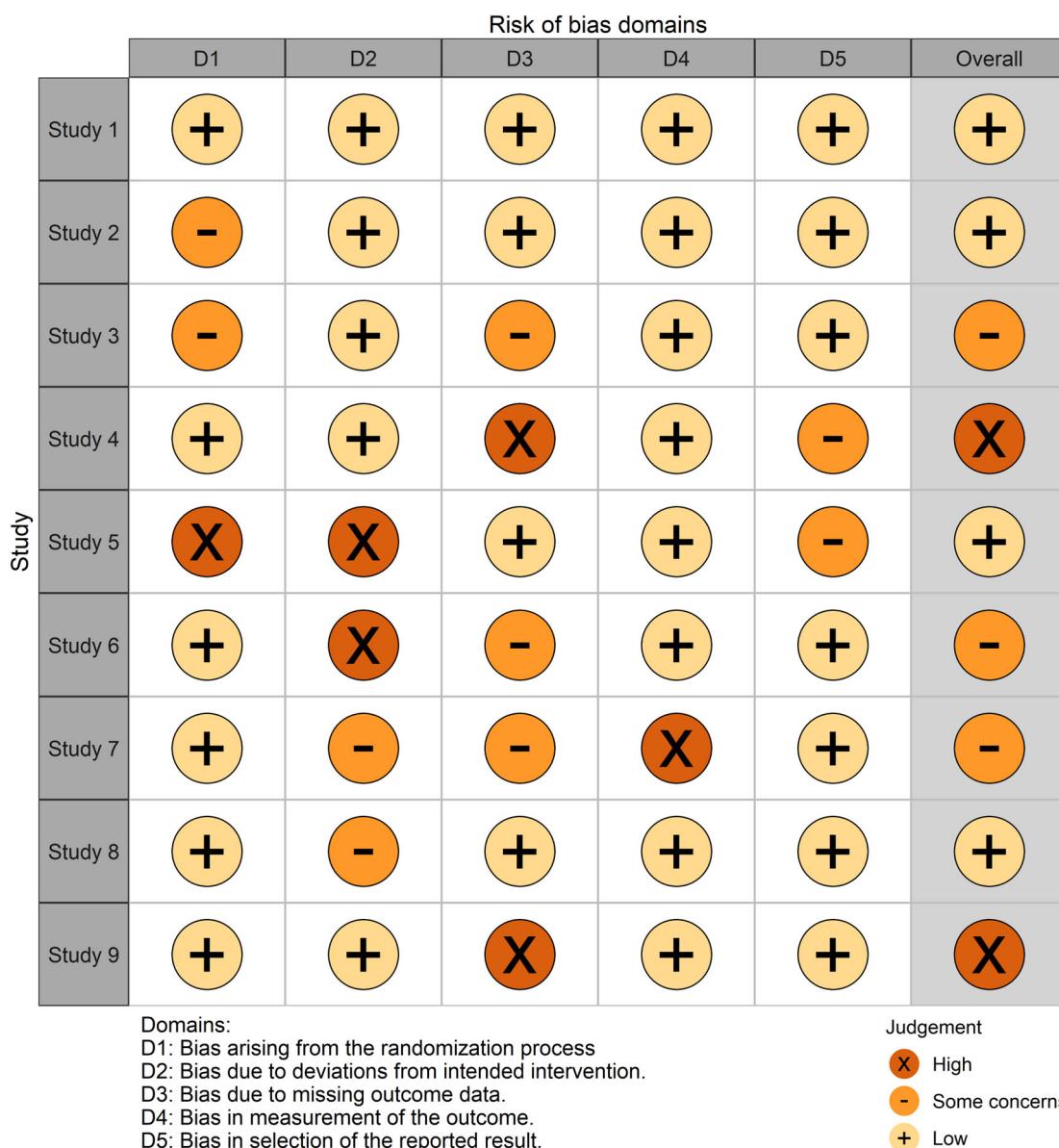


FIGURE 1 Example risk of bias traffic light plot of ROB2 assessments created using **robvis** and the colourblind palette [Colour figure can be viewed at wileyonlinelibrary.com]

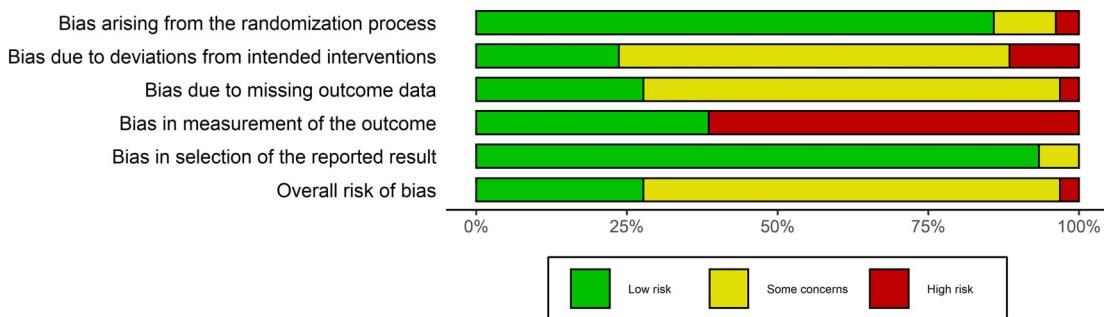


FIGURE 2 Example risk of bias summary plot of ROB2 assessments created using **robvis** and the standard Cochrane palette [Colour figure can be viewed at wileyonlinelibrary.com]

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>

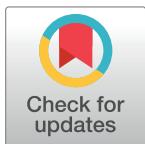
RESEARCH ARTICLE

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

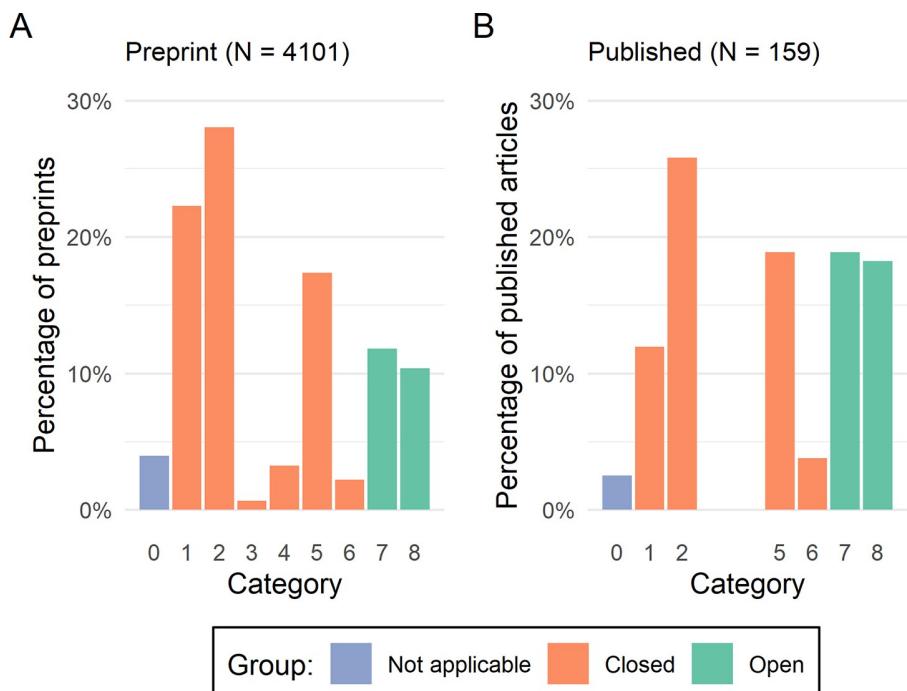


Fig 1. Distribution of the data availability statements of preprinted (Panel A) and published (Panel B) records by category from [Table 1](#).

<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



Ensuring Prevention Science Research is Synthesis-Ready for Immediate and Lasting Scientific Impact

Emily A. Hennessy¹ · Rebecca L. Acabchuk² · Pieter A. Arnold³ · Adam G. Dunn⁴ · Yong Zhi Foo⁵ ·
Blair T. Johnson⁶ · Sonya R. Geange^{3,7} · Neal R. Haddaway^{8,9,10} · Shinichi Nakagawa⁵ · Witness Mapanga¹¹ ·
Kerrie Mengersen¹² · Matthew J. Page¹³ · Alfredo Sánchez-Tójar¹⁴ · Vivian Welch¹⁵ · Luke A. McGuinness¹⁶

Accepted: 23 June 2021
© Society for Prevention Research 2021

Abstract

When seeking to inform and improve prevention efforts and policy, it is important to be able to robustly synthesize all available evidence. But evidence sources are often large and heterogeneous, so understanding what works, for whom, and in what contexts can only be achieved through a systematic and comprehensive synthesis of evidence. Many barriers impede comprehensive evidence synthesis, which leads to uncertainty about the generalizability of intervention effectiveness, including inaccurate titles/abstracts/keywords terminology (hampering literature search efforts), ambiguous reporting of study methods (resulting in inaccurate assessments of study rigor), and poorly reported participant characteristics, outcomes, and key variables (obstructing the calculation of an overall effect or the examination of effect modifiers). To address these issues and improve the reach of primary studies through their inclusion in evidence syntheses, we provide a set of practical guidelines to help prevention scientists prepare synthesis-ready research. We use a recent mindfulness trial as an empirical example to ground the discussion and demonstrate ways to ensure the following: (1) primary studies are discoverable; (2) the types of data needed for synthesis are present; and (3) these data are readily synthesizable. We highlight several tools and practices that can aid authors in these efforts, such as using a data-driven approach for crafting titles, abstracts, and keywords or by creating a repository for each project to host all study-related data files. We also provide step-by-step guidance and software suggestions for standardizing data design and public archiving to facilitate synthesis-ready research.

Keywords Evidence synthesis · Meta-data · Reporting · Reproducibility · Synthesis-ready research · Transparency

Systematic reviews are an important tool for informing policy and practice. When done well, they provide a robust synthesis of available evidence, account for risks of bias in individual studies, and address problems associated with the volume of primary research being published (Johnson & Hennessy, 2019; Wilson & Tanner-Smith, 2014).

To conduct a comprehensive evidence synthesis, such as a systematic review and meta-analysis, primary studies related to the research aim must be identifiable and report relevant study and outcome data. Those who conduct evidence syntheses — hereafter referred to as *evidence synthesists* — often need information about study processes and measured variables aside from a manuscripts' primary outcome of interest to explore why there are differences (i.e.,

statistical heterogeneity) between studies examining the same intervention. This information is often missing, which leads to uncertainty about the generalizability of intervention effectiveness across different groups (e.g., race/ethnicity, gender, age) and can make these efforts slower and less accurate (Borah et al., 2017; Haddaway & Westgate, 2019; Nakagawa et al., 2020; Nuijten et al., 2016). Even when studies report their statistical data and results in detail, there is often still important information missing about other factors that require further attention before they can be accurately synthesized (e.g., description of the intervention, population, or outcome measures). Data and meta-data may be missing from a primary study for many reasons, including article length restrictions, a lack of detailed reporting standards or awareness of the importance of open data, or a desire to retain control of analyses, amongst others.

Given the potential impact of their work on human health and behaviors, prevention scientists should be at the

✉ Emily A. Hennessy
ehennessy@mgh.harvard.edu

Extended author information available on the last page of the article

forefront of ensuring primary research is transparent, reproducible, and furthermore, we argue, synthesis-ready. *Synthesis-ready research* is our term to describe a study (e.g., randomized controlled trial (RCT), quasi-experimental study (QED), or observational study), its *meta-data* (i.e., descriptive information about the study's data, including protocols, tools, algorithms, so that one could replicate the study), and *data* (information collected while conducting the study) that can be readily and accurately used in evidence syntheses.

Resources are available to enable authors to share detailed data and meta-data associated with a study; yet, these resources are often not directly connected to scientific journal practices. Some of these options make it much easier to discover, use, and cite authors' research, substantially increasing the impact and legacy of their work. Thus, our aim was to draw on resources from across disciplines (Grant et al., 2018; Montgomery et al., 2018; Schulz et al., 2010; Wilkinson et al., 2016) to produce a guide to reporting primary research that is open, accessible, and synthesis-ready. To illustrate synthesis-ready research principles and streamline this process for prevention scientists, we briefly present an empirical example from a mindfulness meditation RCT to improve college student well-being ("Introduction to Empirical Example"), followed by a demonstration of the application of best practices for synthesis-ready research to this example ("Evidence Synthesis Process Applied to Our Empirical Example" and "How to Make Research Synthesis-Ready: Steps from Study Design to Completion").

From Primary Study Research to Evidence Synthesis

Primary study conduct and reporting guidelines now advocate for reporting enough information for research to be transparent and reproducible and also provide specific information on how to do so (Chan et al., 2013; Grant et al., 2018; Schulz et al., 2010). Similarly, efforts have focused on improving the degree of FAIRness of primary studies, by making research Findable, Accessible, Interoperable, and Reusable (Wilkinson et al., 2016).¹

Findable indicates research is identifiable to anyone searching for it; data and meta-data and their locations clearly described. Data and meta-data that are retrievable by their identifier (e.g., a digital object identifier, DOI) and are freely obtainable in full without barriers to access (e.g.,

without needing to register with an external website or contacting study authors) are considered *Accessible*. Data and meta-data that are understandable across humans or machines without the need for specialized tools or translation and include linkage to their relevant data/meta-data are *Interoperable*. Finally, *Reusable* refers to the degree to which information can be appropriately repurposed (e.g., integrated/built upon). Thus, research data and meta-data should have a data usage license, detailed records of data origins, and be clearly and richly described in their accompanying materials. By following Open Science recommendations and addressing FAIR principles (Frankenhuis & Nettle, 2018; Vicente-Saez & Martinez-Fuentes, 2018), prevention scientists can provide synthesis-ready research that improves the translation of research into practice.

Evidence syntheses rely on the data reported in research studies and by the researchers themselves. Similar to the conduct of primary research, there are best practices for conducting and reporting of evidence syntheses, including the following: problem formulation, systematically identifying and selecting studies for inclusion, coding studies for key features, calculating effect sizes (if synthesizing quantitative data, such as by conducting a meta-analysis), analyzing the data (assess quality and risk of bias, synthesize coded information from studies), and interpreting and communicating results (Online, Fig. 1). The content and the comprehensiveness of the review (if the review authors follow best practice guidelines) are a direct result of the available information reported in primary studies.

Potential Advantages of Synthesis-Ready Research

Synthesis-ready research can increase the transparency, integrity, and reproducibility of research, ultimately leading to improved scientific evidence (Beugelsdijk et al., 2020). Pre-registration and subsequent data sharing practices should reduce the prevalence of post hoc hypothesis (HARKing; Forstmeier et al., 2017; Kerr, 1998) and falsifying or misrepresenting data in results (Banks et al., 2019; Miyakawa, 2020). Doing so can reduce duplication of effort and introduce findings to the research community sooner and ensure that all findings (beneficial and adverse) are reported, results that can prevent costly mistakes and accelerate the impact of prevention scientists (Gupta et al., 2015; Pasquetto et al., 2019). Indeed, synthesis-ready data can be used repeatedly and for multiple purposes, ultimately informing more research questions: thus it is more cost-effective when compared to data that is only used by the team that collected it.

Primary researchers who make their work synthesis-ready should also benefit from these efforts. Synthesis-ready research is more likely to be identified and used in evidence synthesis (Christensen et al., 2019; Gerstner et al., 2017).

¹ Although FAIR principles primarily refer to enhancing the ability of machines to find and use data, these recommendations also enhance data reuse by individuals. Where appropriate, we have adapted the FAIR components to more directly apply to human-readability efforts for the purposes of this manuscript.

For example, a recent study showed that hosting data in a repository resulted in 25% more citations than merely stating data were available by request (Colavizza et al., 2020). Additionally, depending on the organization of the metadata and the data repository used, the dataset may be cited directly, as may other study materials such as the analytical code, further enhancing the profile of the work and those who conducted it (Crosas, 2013).

Because synthesis-ready research is also more identifiable, there will be increased opportunities for future collaborations (Popkin, 2019), which potentially improves future resource-sharing and reduces costs. Producing synthesis-ready research can also increase the recognition of an individual study team to an international audience. Many funding sources require data sharing of their grantees and necessitate the inclusion of such plans in grant proposals (e.g., Wellcome Trust, the UK Medical Research Council, and the US National Institutes of Health: Gaba et al., 2020; NIH Data Sharing Policy & Guidance, 2003; Wellcome Trust Public Health Research Data Forum, 2012). Thus, researchers who prepare and deliver grant-funded research findings with synthesis-ready considerations are considered responsive to funders while potentially attaining these other discussed benefits.

Introduction to Empirical Example

Our empirical example is a RCT ($N=140$ participants; registered in Clinical Trials, NCT03402009) conducted across two university semesters in the 2018 academic year (Acabchuk et al., 2021). The study evaluated two active treatment conditions to determine what tools best assist university students in developing a personal meditation practice to self-manage symptoms of depression, anxiety, and stress. One group was assigned to the *10% Happier* meditation phone app (labeled the “App Group”; Ten Percent Happier: Mindfulness Meditation Courses, 2020), which provides guided meditations, and the other was assigned to the *10% Happier* meditation phone app in combination with a Muse EEG neurofeedback device (labeled the “Muse Group”; Choose Muse, 2018). The study included a baseline assessment, randomization to one of two groups, an orientation session for both groups, collection of salivary biomarkers, and follow-up survey assessments on a variety of well-being and behavior outcomes throughout the study (Study Timeline: Online, Fig. 2).

Evidence Synthesis Process Applied to Our Empirical Example

Drawing from our empirical example, if one sought to conduct a systematic review of all types of mindfulness interventions for college students on health-related outcomes

(*Problem Formulation*), they would first need to be able to identify all primary research matching their inclusion criteria (*Finding and Selecting Studies*). Clearly delineating the research problem is a foundational step in a review, as it drives the search strategy and ability to accurately identify and synthesize primary study research. Research questions in an evidence synthesis should address all the main elements of the studies to be included; many draw on the PICOS framework with its focus on the population(s), intervention(s), comparator(s), outcome(s), and study design(s) elements of a research problem (Stern et al., 2014). Once the relevant literature is identified, data are extracted from the studies (*Code Studies for Relevant Features*) so that the analysis can commence by synthesizing coded data (*Analyze the Systematic Review Database*).

Within this broader process, the types of data and metadata that need to be reported in a primary study depend on the task and/or type of evidence synthesis. For example, research methods need to be fully reported to assess risk of bias. If the review aimed to measure the effectiveness of mindfulness interventions on depression symptoms using a meta-analysis, then primary studies should report the data necessary to calculate an effect size from their depressive symptom assessment for all participants (e.g., means, standard deviations, and sample sizes for each group). To address relevant potential effect modifiers (e.g., age, year in school, gender, intervention fidelity), any data collected during the primary study on participant characteristics and the study methods (e.g., description of groups, duration of intervention) needs to be transparently reported. To examine geographic and temporal effect modifiers, key information about the study location and timing should be provided (Johnson et al., 2017). Additionally, if the synthesis focuses on participants’ perspectives on why interventions are more or less effective, they would need access to any qualitative data collected as part of the study (e.g., interviews, focus groups). Given this variability in the type of questions that could be addressed in a synthesis, it is important for investigators to transparently and comprehensively report a variety of information about their research.

How to Make Research Synthesis-Ready: Steps from Study Design to Completion

Approaching every step of a research project with data sharing in mind is useful and often necessary to ensure that data sharing is possible. This approach has the added benefit of encouraging better overall project management, including clearer workflows and comprehensive documentation of the analytical process. We now illustrate each step of these suggestions from study design to sharing with examples from our mindfulness trial (Fig. 1).

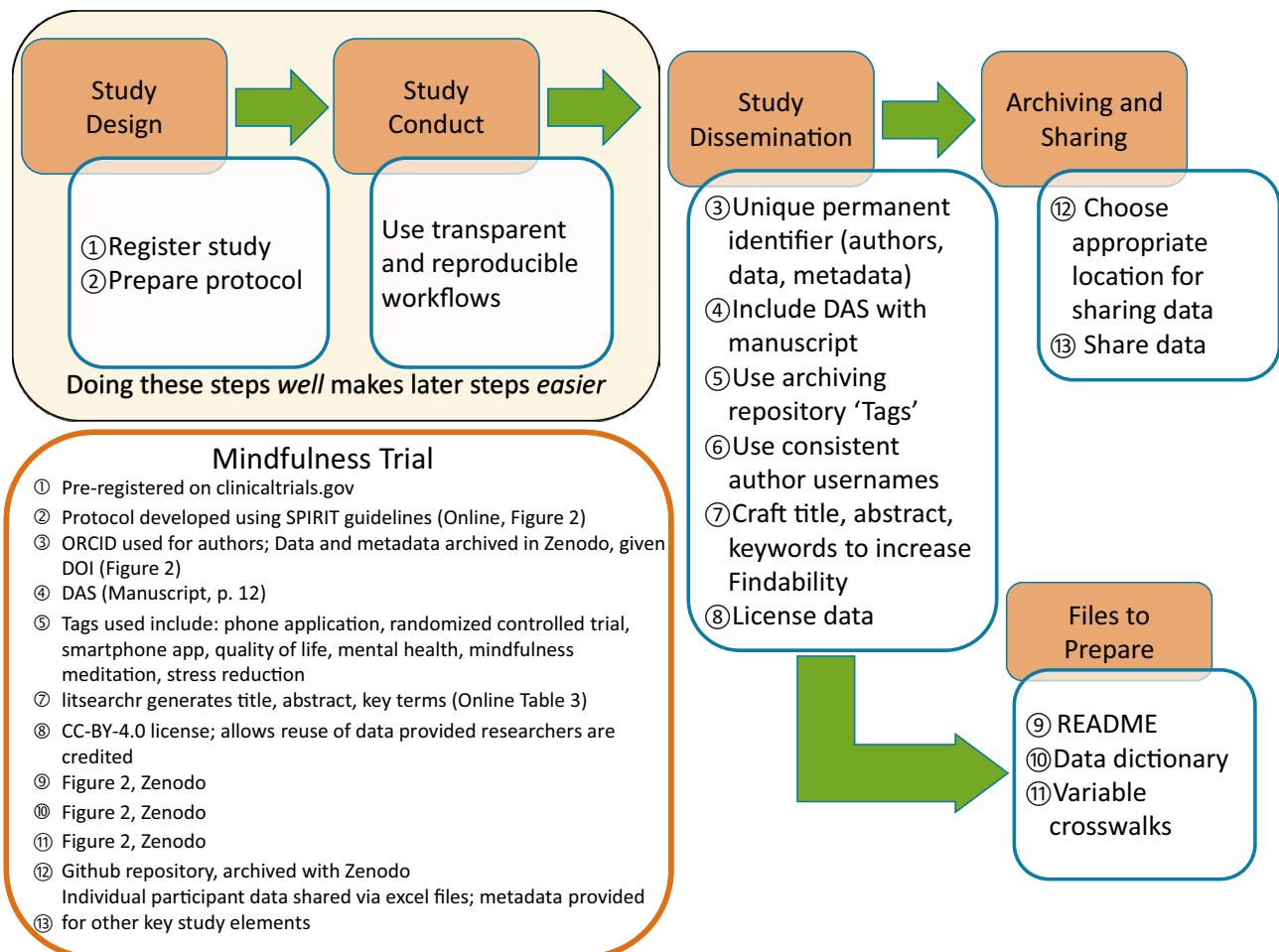


Fig. 1 Best practices in generating synthesis-ready research applied to mindfulness trial. DAS, data availability statement

Study Design

Preparing research for sharing can be a resource-intensive process, and so planning and provisioning for this task are vital to consider when submitting funding applications and considering funders' reporting requirements. When conducting research with human participants, advance planning for data sharing is essential because it ensures appropriate documentation is prepared. Any potential ethical concerns with data sharing should be addressed during the review of study materials by the governing Institutional Review Board, which ensures that participant consent documents include the correct language to ensure participants can consent to anonymized data sharing (Ohmann et al., 2017; Thorogood & Knoppers, 2017). The Inter-university Consortium for Political and Social Research (ICPSR) recommends carefully constructing these sections and provides sample text to use to ensure that participants are informed about the potential use of their anonymized data in future research (Inter-university Consortium for Political and Social Research, 2010).

The study protocol is the foundation of the study, detailing all the processes, procedures, and materials planned for a research project. Best practice for primary studies and secondary data analysis is to register a study protocol that includes primary questions and/or hypotheses, methodology, and planned analyses. With careful attention to developing a comprehensive protocol using a standard structure, this document can increase the synthesis-readiness of the research. The information provided in a study protocol includes important meta-data that may not be available in a single published manuscript but that can be used by evidence synthesists to address a variety of research questions, including using it to assess study quality and risk of bias due to selection of reported results or "outcome switching" (Falk Delgado & Falk Delgado, 2017; Kahan & Jairath, 2018). We recommend that authors review published guidelines for the best tool for the study design in question: The EQUATOR network provides an online resource to compare these tools (<https://www.equator-network.org/>). Online Tables 1 and 2 present a mapping of critical elements for a study protocol

(using the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT); Chan et al., 2013) and for a study *manuscript* (using the Consolidated Standards of Reporting Trials for social and psychological interventions (CONSORT); Grant et al., 2018)) and indicate how study information provided following these guidelines could be used in an evidence synthesis. Table 1

The protocol for our mindfulness trial uses the SPIRIT guidelines (Chan et al., 2013) and is archived in a Zenodo repository (DOI: <https://doi.org/10.5281/zenodo.4011716>).

Study Conduct

To facilitate data sharing and to distribute the workload across the projects' lifecycle, transparent and reproducible data management workflows should be employed throughout the research process, including (1) defining data file contents; (2) clearly describing variable names, definitions, units, and formats; (3) implementing consistent data organization and file structure; (4) performing basic quality assurance to ensure that values fall within expected ranges; and (5) providing documentation including the analytical steps taken (crosswalks), data dictionaries (i.e., codebooks), and metadata (Barton et al., 2010; Strasser, 2012; Yenni et al., 2019). Building in these meta-data collection steps into the analysis process makes it easier to create items like variable crosswalks and data dictionaries to accompany shared research.

Study Dissemination

The most labor-intensive stage of making data synthesis-ready within a FAIR perspective occurs during preparation for study dissemination as it involves several steps. To ensure their research is *Findable* and *Accessible*, authors should (1) ensure all authors have a unique permanent identifier such as an ORCID iD (Open Researcher and Contributor ID, 2009) and include this identifier with manuscript and data submission (because institutional emails are often impermanent); (2) include a detailed data availability statement with manuscript submission; (3) assign data and meta-data a globally unique and persistent identifier (e.g., DOI), which ensures long-term discoverability; (4) use available data ‘tags’ in the archiving repository; (5) use consistent names/usernames across the different journal and archiving platforms; and (6) craft the title and abstract and select keywords that will optimize the chances of research being identified via the most typical evidence synthesis method of searching for research (i.e., electronic database searches). To ensure their data is also *Interoperable* and *Reusable*, researchers need to attend to licensing, data file structures and formats, and detailed meta-data when preparing their data for sharing.

Data Availability Statements to Ensure Findability A data availability statement (DAS; also known as “data access statements” or “availability of materials”) provides information about how and where other researchers can access the study data and is often required when submitting a manuscript for publication (Federer et al., 2018; Graf et al., 2020). In some closed-access journals, DAS are treated as article meta-data, so they are accessible even if the article is behind a paywall, ensuring the data are findable. Thus, the DAS should be as explicit as possible, even if some information that is in the manuscript is repeated in the DAS.

The DAS for our mindfulness trial example states: “Both the full raw dataset and the finalized analytical dataset used in this analysis are available from Zenodo (DOI: <https://doi.org/10.5281/zenodo.4011716>), along with comprehensive data dictionaries and crosswalks.”

Creating Titles, Abstracts, and Keywords to Optimize Findability Carefully crafting titles and abstracts, in addition to choosing appropriate keywords, is essential to help other researchers find your research. As much as possible given word limits, these should describe the condition, exposure, and settings of a study (Grant et al., 2018; Montgomery et al., 2018). These pieces of the submission are often the last to be created but are the main pieces of information indexed and displayed in electronic databases and are used to identify relevant research for an evidence synthesis. As a result of their ability to increase the findability of research, we suggest a data-driven approach to crafting these sections. The `litsearchr` package for the R programming environment offers a systematic approach to generate a list of words likely to improve the identifiability of the study manuscript (Grames et al., 2019; R Core Team, 2019).

Using our mindfulness trial example, we began with a set of keywords we thought likely to capture a set of relevant published articles from which we could draw the most highly used keywords. We conducted a search using these keywords in PubMed on August 24, 2020 and used `litsearchr` to systematically generate a list of important terms. The results suggest that several terms could be strategically incorporated in the title, abstract, and keywords to increase the findability of the trial (Online, Table 3 provides further details using our example).

Data Licenses Assigning a license to the data being shared is important to govern how the data can be reused and, more importantly, to help ensure that the study team is credited.²

For our mindfulness trial example, we chose the CC-BY-4.0 license, which allows the reuse of the data provided the original researchers are credited with the creation of the dataset.

² Comparing the full suite of licenses available is beyond the scope of this article; online resources such as ChooseALicense (<https://choosealicense.com/non-software/>) and Creative Commons (<https://creativecommons.org/licenses/>) provide detailed descriptions of the range of licenses available for data sharing.

Table 1 Comparison of the options for sharing data

Location	Examples	Advantages	Disadvantages	Examples
Non-affiliated repositories	Zenodo; Dryad; Figshare; Open Science Framework	Free (except Dryad) Self archiving. DOI assigned. Long-term storage. Can blind repositories for double-blind peer- review (Anonymous Github; OSF “Anonymous” Links)	Some have a Data Publishing Charge (e.g., Dryad charges \$120). No review of files uploaded Available from https://doi.org/10.5281/zenodo.3529382	Repository for Abbott et al., 2019, contains raw dataset, analysis data- set, cleaning code, analysis code.
Institutional repositories	data.bris; Harvard Dataverse; Inter-university Consortium for Political and Social Research; UK Data Service	Supported by institution DOI assigned. Often can get help from staff when uploading data	Not as well known as other repositories—limits findability	Pescosolido et al., 2020, contains scripts and data to replicate all fig- ures/analyses found in the original manuscript
Supplemental files attached to the online version of the published article	Many file types (.txt,.xlsx,.doc,.pdf) with data or metadata	Likely relatively less work to prepare compared to the other options	No DOI (for supplemental files) Can be closed-access. Potential for data to be lost. Least findable of the other options	Slep et al., 2020, contains mPLUS code and data Tomiczyk et al., 2020, open access, includes data in manuscript and sup- plemental files

Data File Structures and Formats Sharing individual participant data, i.e., data on each variable collected for each participant, affords those wishing to reuse or synthesize these data a range of advantages. Access to data on all variables collected, rather than just those used in the primary study's analysis, allows synthesists to ensure that an estimate from a primary study is adjusted for a common set of variables when it is included in a meta-analysis, even if the published effect estimate for that study was not (Riley et al., 2010). Even among individual participant data, different levels of granularity exist.

It may not be possible to share each individual participant data point, especially when participant confidentiality is at risk, for example, when studying a rare disease or among participants of specialized programs. In these cases, there are several options for sharing individual participant data. The first is to generate a "synthetic" dataset that mimics the individual-level data by preserving statistical properties and the relationships between variables. Software, for example, the synthpop R package, can help users generate a synthetic dataset (Quintana, 2020, provides a tutorial on using this package). A second option is to remove some individual characteristics that might give away participants' identities, such as age, sex, and location of recruitment source, but provide the rest of the variables as individual data points. Finally, if it is not possible to share any individual participant data, then at the very minimum, summary-level data (e.g., means and standard deviations or percentages which describe the distribution of all variables between intervention/exposure groups) should be provided for all variables included in the paper and, if possible, the dataset, not solely those thought by the primary study team to be relevant. Of course, for investigators conducting years of longitudinal studies examining numerous questions for which they are not ready to share data, we advise authors to explicitly state in their DAS the reasons for not sharing data. By beginning to capture the true reasons for not sharing data, data availability statements can then begin to be used as support for introducing better protection for and citation of data producers.

To illustrate the distinction in the level of data shared, for our mindfulness trial, we shared the raw data exported from the survey instrument, which contains each participant's answer to each individual question, and the processed individual participant data, which include scale score summaries (i.e., total DASS-21 score, comprised of 21 individual questions). To explore these different datasets, see the "REDCAP_data.xlsx" and "combined_data.xlsx" files in our Github repository which is archived by Zenodo (DOI: <https://doi.org/10.5281/zenodo.4011716>).

Study data can be shared as machine-readable (e.g., unformatted Excel/CSV files and raw DNA microarray

sequences) or human-readable formats (e.g., formatted PDF tables, figures, and data contained in text such as "There was evidence of an association (Odds ratio: 0.91, 95%CI: 0.85–0.96)"). Although tools are being developed to extract data automatically from human-readable sources for evidence syntheses (Marshall et al., 2020), it is often challenging to import these data into statistical analysis software, reducing interoperability and synthesis-readiness.

For our mindfulness trial example, all study data are shared via Excel files in our Github repository which is archived by Zenodo (DOI: <https://doi.org/10.5281/zenodo.4011716>). This includes the mindfulness trial's quantitative survey data. It also includes qualitative data, which were generated from in-person, semi-structured interviews and coded to examine adherence and problems with using the tools. To demonstrate synthesis-readiness, the relevant responses for qualitative coding are included in the data file rather than the full interview notes.

Meta-data Providing a README, a plain text file that contains useful information about data or software, can help ensure that data are correctly accessed, interpreted, and reused by other researchers. The README file should include a (1) definition of the primary research aims using a relevant alternative framework that addresses the main components of the research which will help others identify and use the research (e.g., PICOTS, TOPICS + M; Haynes, 2006; Johnson & Hennessy, 2019; Stern et al., 2014); (2) a short description of all the data files included in the project repository; (3) signposts to the detailed data dictionaries for each data file which contain definitions of all variables, including coded (e.g., "CRP_t2") and interpretable (e.g., "C-Reactive Protein levels at 2 months post-baseline") variable names, the values that each variable is allowed to take, and include missing data indicators and units of measurement; (4) a summary of data processing steps that were used to produce the final analytic datasets from the raw data files; (5) a description of each data file's relationship to other data files in the repository, and relevant linkage variables, such as participant ID; (6) permanent author contact details; and (7) details of the license, including the full name of the license holder, the year, and where the full text of the license can be found.

The README for our mindfulness trial contains the seven main elements necessary for a comprehensive metadata file; see Fig. 2 for a snapshot of these pieces and visit Zenodo to view and download all files: <https://doi.org/10.5281/zenodo.4011716>.

Archiving and Sharing Research Data

Research data sharing can be done via institutional and non-institutional repositories or alongside published articles

<p>Study description</p> <p>Taken from the Study protocol.docx document, which describes the study approach in detail:</p> <p>"Young adults in college experience high levels of stress, anxiety, and depression, which can lead to maladaptive behaviors and chronic health issues including inflammation and HPA-axis dysregulation. University mental health centers are seeking effective programs that teach coping skills to self-manage stress, with low stigma, that are easy to implement in large scale. The aim of this protocol is to describe the design of a study aiming to evaluate the effectiveness of a one-month self-guided meditation program for university students comparing various tools to improve mental health, physical health and health behaviors, and explore moderators. Secondary aims include reaching a diverse audience and documenting adverse experiences."</p>	<p>PICOTS</p> <p>Population - 140 university undergraduate or postgraduate students studying at an American university.</p> <p>Intervention - 10 minutes of self-guided meditation daily for four weeks, facilitated by the 10% Happier App, in combination with neurofeedback from the Muse Neurofeedback Tool.</p> <p>Comparison - 10 minutes of self-guided meditation daily for four weeks, facilitated by the 10% Happier App.</p> <p>Outcomes - Outcome measures included those related to Personality, Stress, Distress, Resilience (coping), Mindfulness, Emotion regulation, Sleep, Self-Regulation of Eating, Motivation Survey, and Student Adherence (for more details, see Section 2.4 of the Study protocol.docx document and the Instrument List.xlsx document).</p> <p>Time - 4 week intervention followed by post-intervention surveys (1-, 6- and 12-months). See Timeline.xlsx for more details on the tasks performed and the data collected at each timepoint.</p> <p>Study Design - randomized controlled trial with two intervention groups</p>
<p>(a)</p> <p>Data source files</p> <p>All data files are available as Excel spreadsheets in the <code>data/</code> subdirectory. The relationships between each data file, and between data files and other information contained in this repository, are described below:</p> <ul style="list-style-type: none"> • <code>RECAP_data.xlsx</code> (Data dictionary: Data Dictionary for REDCAP Data.pdf) contains the raw survey data exported from the REDCAP software. • <code>External_data.xlsx</code> (Data dictionary: Data Dictionary for External Data.xlsx) contains additional variables on each participant which were not captured by the REDCAP survey. Note some variables, such as C-Reactive protein or cortisol levels, are only available for a subset of participants, based on the availability of saliva samples for that participant. See Section 2.2.1 in Study protocol.docx for more details. • <code>Combined_data.xlsx</code> was created by merging the <code>RECAP_data.xlsx</code> and <code>External_data.xlsx</code>, and is used as the complete dataset from which composite variables are created using the <code>create_composite_variables.do</code> STATA do file. The crosswalk between the individual measures and the composite variables are described in more detail in <code>Stata Variable Crosswalk.xlsx</code> • <code>Debriefing Interview - Wave 1.xlsx</code> : contains qualitative responses from "Wave 1" participants to the debriefing interview questions. Note: the interview was only done for Wave 1 participants. <p>(c)</p>	<p>(b)</p> <p>Permanent contact about this data</p> <p>Rebeca L. Acabchuk, University of Connecticut, Institute for Collaboration on Health, Intervention, & Policy, Department of Psychological Sciences, United States. ORCID: 0000-0001-8254-2396</p> <p>› Licence</p> <p>The materials in this repository are licensed under the Creative Commons Attribution 4.0 (CC-BY-4.0) International. The full license file can be found here.</p> <p>(d)</p>

Fig. 2 Mindfulness trial metadata included in the README file. Panels **a** and **b** provide the study description according to the PICOTS framework; **c** provides data source files, locations, and rela-

tionships between the files; **d** provides permanent author information and licensing information

(Table 1). We have ordered these sharing options from best practice (most likely to be synthesis-ready) to good practice (still synthesis-ready but may require additional effort to prepare for synthesis).

Non-affiliated Repositories Non-affiliated repositories are commercial or charity repositories that allow researchers to self-archive, and assign a DOI to, their research data. Examples include Dryad, Open Science Framework, and Zenodo (Dryad, 2020; Open Science Framework, 2020; Zenodo—Research Shared, 2020). The main limitation of these repositories is that there are often conditions on how much data can be stored; some charge fees for storage.

For our mindfulness trial example, the repository was created in GitHub and archived using Zenodo, as GitHub deposition is non-permanent and does not assign a DOI to the uploaded data. Repositories of this type can be embargoed prior to publication to prevent “scooping,” and some repositories also provide services to help prevent scooping. For example, projects hosted in the Open Science Framework can have “anonymous” links to remove identifying meta-data as part of double-blind peer-review requirements.

Institutional Repositories An institutional repository is one hosted by an academic institution, such as the ICPSR (ICSPR, 2021). An advantage of affiliated repositories is that some provide staff to guide researchers who are new to data sharing. These repositories are less well-known, which may limit the findability of data when compared to other options. The terms of deposition should be reviewed to ensure that authors retain data ownership and copyright over the deposited materials.

Supplemental Files A common option for sharing research data is as a supplementary file to the published manuscript. However, this method has several major limitations. This approach reduces data findability because datasets shared as supplementary material are grouped under the publication’s DOI and are not uniquely identifiable. There is also the potential for supplementary files to become disassociated with their respective publications, for example, during a website redesign. If the published article is behind a paywall, then data attached to it as supplementary files may also be blocked and so data shared as supplemental files may not be fully accessible. Sharing data as supplementary files to an open-access publication should be considered the minimum level of data sharing needed to make research synthesis-ready.

Potential Challenges and Future Considerations

In this article, we have outlined the benefits of synthesis-ready research and provided guidelines on how to follow best practice and produce synthesis-ready research. Yet, we appreciate that there are many perceived and actual challenges to doing so, including costs to data sharing, lack of incentives, lack of awareness or training, not wanting to “give away” data or get “scooped” on findings, and the necessity of ensuring participant confidentiality (Evans, 2016; Tennant et al., 2019; Tenopir et al., 2011, 2015; van Panhuis et al., 2014; Walters, 2020).

To address the lack of knowledge and training, we have provided information, workflows, and a full demonstration using an empirical example, in addition to signposting to further literature, to help researchers plan for the resources needed for making research synthesis-ready at project conceptualization. Many funders have recognized the value of data sharing and will provide funds to do so based on investigator budgets which can offset the tangible (platform sharing fees) and intangible (time spent preparing data) costs of this process.

Researchers have valid concerns about having their research scooped, and yet, several measures can help prevent scooping including pre-registering study analysis plans and depositing preprints while awaiting publication, all core practices of open science. An additional solution is to license data at the point of publication or archival (“[How to Make Research Synthesis-Ready: Steps from Study Design to Completion](#)”). Finally, some journals have implemented “scoop-protection” policies for research, a system-level solution we hope more journals will implement to address this challenge (Transpose, 2020). Ensuring that all types of research data are reused appropriately is a warranted concern (Bishop, 2009; Yardley et al., 2014); yet, evidence suggests that individuals reusing data seek to collaborate with the original data creators to harness their knowledge about the data, rather than using it to scoop findings or to disprove the original publication’s findings, resulting in improved and more accurate data reuse and increased opportunities for publication (Pasquetto et al., 2019). As well, providing clearly detailed documentation in the data and meta-data files (“[How to Make Research Synthesis-Ready: Steps from Study Design to Completion](#)” and Fig. 2) should reduce inappropriate use of shared data. Finally, we have provided several different options for maintaining the confidentiality of participants while attending to synthesis-readiness (“[How to Make Research Synthesis-Ready: Steps from Study Design to Completion](#)”).

Conclusions

Despite the importance of making research available and synthesis-ready, not all scientific disciplines have embraced data/meta-data sharing and transparency at the same rate (Graf et al., 2020; Hardwicke et al., 2020; Sholler et al., 2019; Vasilevsky et al., 2017). If prevention scientists share research in synthesis-ready forms, the field would see improvements, especially in terms of impact on policy and practice. We presented methods for managing data sharing processes that should make it easier to produce synthesis-ready research. Changes in research practice that focus on synthesizability will reduce the delay between implementing effective primary study approaches (Nakagawa et al., 2020) and should have a positive impact on translation of prevention science into policy and practice.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11121-021-01279-8>.

Acknowledgements The paper was first developed during the 2019 Evidence Synthesis Hackathon.

Author Contribution Project administration (EAH; LAM), roles/writing — original draft (EAH; LAM), visualization (EAH; LAM); data curation (RLA; LAM); resources (LAM); funding acquisition (NRH); investigation (EAH; LAM; RLA), methodology (EAH; LAM); software (EAH; LAM); conceptualization (All authors), writing — review and editing (all authors).

Funding The Fenner School of Environment & Society (ANU), the University of New South Wales, and the Faculty of Humanities at the University of Johannesburg funded the Evidence Synthesis Hackathon 2019. EAH has support from NIAAA (K01 AA028536-01). MJP has support from an Australian Research Council Discovery Early Career Researcher Award (DE200101618). LAM has support by an NIH Doctoral Research Fellowship (DRF-2018-11-ST2-048).

Declarations

Conflict of Interest The authors declare no competing interests.

Disclaimer The views expressed in this article are those of the authors and do not necessarily represent those of the NIAAA, NHS, the NIHR, or the Department of Health and Social Care.

References

- Abbott, S., Christensen, H., Welton, N. J., & Brooks-Pollock, E. (2019). Estimating the effect of the 2005 change in BCG policy in England: A retrospective cohort study, 2000 to 2015. *Eurosurveillance*, 24(49), 1,900,220. <https://doi.org/10/gg9nqk>
- Acabchuk, R. L., Simon, M. A., Low, S., Brisson, J. M., & Johnson, B. T. (2021). Measuring meditation progress with a consumer-grade EEG device: Caution from a randomized controlled trial. *Mindfulness*, 12, 68–81. <https://doi.org/10.1007/s12671-020-01497-1>

- Banks, G. C., Field, J. G., Oswald, F. L., O'Boyle, E. H., Landis, R. S., Rupp, D. E., & Rogelberg, S. G. (2019). Answers to 18 questions about open science practices. *Journal of Business and Psychology*, 34, 257–270. <https://doi.org/10.1007/s10869-018-9547-8>
- Barton, C., Smith, R., & Weaver, R. (2010). Data practices, policy, and rewards in the information era demand a new paradigm. *Data Science Journal*, 9, IGY95–IGY99. https://doi.org/10.2481/dsj-SS_IGY-003
- Beugelsdijk, S., van Witteloostijnen, A., & Meyer, K. E. (2020). A new approach to data access and research transparency (DART). *Journal of International Business Studies*, 51, 887–905. <https://doi.org/10.1057/s41267-020-00323-z>
- Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of Social Issues*, 44(3), 255–272.
- Borah, R., Brown, A. W., Capers, P. L., & Kaiser, K. A. (2017). Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*, 7, e012545. <https://doi.org/10.1136/bmjopen-2016-012545>
- Chan, A. W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P. C., Krleža-Jerić, K., & Doré, C. J. (2013). *SPIRIT 2013 statement: Defining standard protocol items for clinical trials*. Annals of Internal Medicine; Ann Intern Med. <https://doi.org/10.7326/0003-4819-158-3-201302050-00583>
- Choose Muse. (2018, June 25). *A deep dive into brainwaves: Brainwave frequencies explained*. Muse. <https://choosemuse.com/blog/a-deep-dive-into-brainwaves-brainwave-frequencies-explained-2/>
- Christensen, G., Dafoe, A., Miguel, E., Moore, D. A., & Rose, A. K. (2019). A study of the impact of data sharing on article citations using journal policies as a natural experiment. *PLoS One*, 14, e0225883. <https://doi.org/10.1371/journal.pone.0225883>
- Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., & McGillivray, B. (2020). The citation advantage of linking publications to research data. *PLoS One*, 15, 023–0416. <https://doi.org/10.1371/journal.pone.0230416>
- Crosas, M. (2013, October 30). *Joint Declaration of Data Citation Principles—FINAL*. FORCE11. <https://www.force11.org/datacitationprinciples>
- Dryad. (2020). https://datadryad.org/stash/our_mission
- Evans, S. R. (2016). Gauging the purported costs of public data archiving for long-term population studies. *PLoS Biology*, 14, e1002432. <https://doi.org/10.1371/journal.pbio.1002432>
- Falk Delgado, A., & Falk Delgado, A. (2017). Outcome switching in randomized controlled oncology trials reporting on surrogate endpoints: A cross-sectional analysis. *Scientific Reports*, 7. <https://doi.org/10.1038/s41598-017-09553-y>
- Federer, L. M., Belter, C. W., Joubert, D. J., Livinski, A., Lu, Y. L., Snyders, L. N., & Thompson, H. (2018). Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLoS One*, 13, e0194768. <https://doi.org/10/gdhb7j>
- Forstmeier, W., Wagenmakers, E. J., & Parker, T. H. (2017). Detecting and avoiding likely false-positive findings – a practical guide. *Biological Reviews*, 92, 1941–1968. <https://doi.org/10.1111/brv.12315>
- Frankenhuis, W. E., & Nettle, D. (2018). Open science is liberating and can foster creativity. *Perspectives on Psychological Science*, 13, 439–447. <https://doi.org/10.1177/1745691618767878>
- Gaba, J. F., Siebert, M., Dupuy, A., Moher, D., & Naudet, F. (2020). Funders' data-sharing policies in therapeutic research: A survey of commercial and non-commercial funders. *PLoS One*, 15, e0237464. <https://doi.org/10.1371/journal.pone.0237464>
- Gerstner, K., Moreno-Mateos, D., Gurevitch, J., Beckmann, M., Kambach, S., Jones, H. P., & Seppelt, R. (2017). Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods in Ecology and Evolution*, 8, 777–784. <https://doi.org/10.1111/2041-210X.12758>
- Graf, C., Flanagan, D., Wylie, L., & Silver, D. (2020). The open data challenge: An analysis of 124,000 data availability statements and an ironic lesson about data management plans. *Data Intelligence*, 1–15. https://doi.org/10.1162/dint_a_00061
- Grames, E. M., Stillman, A. N., Tingley, M. W., & Elphick, C. S. (2019). An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution*, 10, 1645–1654.
- Grant, S., Mayo-Wilson, E., Montgomery, P., Macdonald, G., Michie, S., Hopewell, S., Moher, D., Aber, J. L., Altman, D., Bhui, K., Booth, A., Clark, D., Craig, P., Eisner, M., Fraser, M. W., Gardner, F., Grant, S., Hedges, L., Hollon, S., & CONSORT-SPI Group. (2018). CONSORT-SPI 2018 explanation and elaboration: Guidance for reporting social and psychological intervention trials. *Trials*, 19, 406. <https://doi.org/10.1186/s13063-018-2735-z>
- Gupta, Y. K., Meenu, M., & Mohan, P. (2015). The Tamiflu fiasco and lessons learnt. *Indian Journal of Pharmacology*, 47, 11–16. <https://doi.org/10.4103/0253-7613.150308>
- Haddaway, N. R., & Westgate, M. J. (2019). Predicting the time needed for environmental systematic reviews and systematic maps. *Conservation Biology*, 33(2), 434–443. <https://doi.org/10.1111/cobi.13231>
- Hardwicke, T. E., Wallach, J. D., Kidwell, M. C., Bendixen, T., Crüwell, S., & Ioannidis, J. P. A. (2020). An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014–2017). *Royal Society Open Science*, 7, 1908706. <https://doi.org/10.1098/rsos.190806>
- Haynes, R. B. (2006). Forming research questions. *Journal of Clinical Epidemiology*, 59, 881–886. <https://doi.org/10.1016/j.jclinepi.2006.06.006>
- ICPSR. (2021). Retrieved February 24, 2021, from <https://www.icpsr.umich.edu/web/pages/>
- Inter-university Consortium for Political and Social Research (2010). Guide to social science data preparation and archiving best practice throughout the data life cycle. *ICPSR Institute for Social Research University of Michigan*.
- Johnson, B. T., & Hennessy, E. A. (2019). Systematic reviews and meta-analyses in the health sciences: Best practice methods for research syntheses. *Social Science & Medicine*, 233, 237–251. <https://doi.org/10.1016/j.socscimed.2019.05.035>
- Johnson, B. T., Cromley, E. K., & Marrouch, N. (2017). Spatiotemporal meta-analysis: Reviewing health psychology phenomena over space and time. *Health Psychology Review*, 11, 280–291.
- Kahan, B. C., & Jairath, V. (2018). Outcome pre-specification requires sufficient detail to guard against outcome switching in clinical trials: A case study. *Trials*, 19, 265. <https://doi.org/10.1186/s13063-018-2654-z>
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2, 196–217.
- Marshall, I. J., Johnson, B. T., Wang, Z., Rajasekaran, S., & Wallace, B. C. (2020). Semi-Automated evidence synthesis in health psychology: Current methods and future prospects. *Health Psychology Review*, 14, 145–158. <https://doi.org/10.1080/17437199.2020.1716198>
- Miyakawa, T. (2020). No raw data, no science: Another possible source of the reproducibility crisis. *Molecular Brain*, 13, 24. <https://doi.org/10.1186/s13041-020-0552-2>
- Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., Moher, D., Lawrence Aber, J., Altman, D., Bhui, K., Booth, A., Clark, D., Craig, P., Eisner, M., Fraser, M. W., Gardner, F., Grant, S., Hedges, L., Hollon, S., & CONSORT-SPI Group. (2018). Reporting randomised trials of social and psychological interventions: The CONSORT-SPI 2018 Extension. *Trials*, 19, 407–407. <https://doi.org/10.1186/s13063-018-2733-1>
- Nakagawa, S., Dunn, A. G., Lagisz, M., Bannach-Brown, A., Grames, E. M., Sánchez-Tójar, A., O'Dea, R. E., Noble, D. W. A., Westgate,

- M. J., Arnold, P. A., Barrow, S., Bethel, A., Cooper, E., Foo, Y. Z., Geange, S. R., Hennessy, E., Mapanga, W., Mengersen, K., Munera, C., & Hackathon, E. S. (2020). A new ecosystem for evidence synthesis. *Nature Ecology & Evolution*, 4, 498–501. <https://doi.org/10.1038/s41559-020-1153-2>
- NIH data sharing policy and guidance. (2003). National Institutes of Health. https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm
- Nuijten, M. B., Hartgerink, C. H. J., van Assen, M. A. L. M., Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, 48, 1205–1226. <https://doi.org/10.3758/s13428-015-0664-2>
- Ohmann, C., Banzi, R., Canham, S., Battaglia, S., Matei, M., Ariyo, C., Beclen, L., Bierer, B., Bowers, S., Clivio, L., Dias, M., Druml, C., Faure, H., Fenner, M., Galvez, J., Ghersi, D., Gluud, C., Groves, T., Houston, P., & Demotes-Mainard, J. (2017). Sharing and reuse of individual participant data from clinical trials: Principles and recommendations. *BMJ Open*, 7, e018647. <https://doi.org/10.1136/bmjopen-2017-018647>
- Open Science Framework. (2020). <https://osf.io/>
- ORCID iD. (2009). "ORCID website". orcid.org. Retrieved 18 April 2021.
- Pasquetto, I. V., Borgman, C. L., & Wofford, M. F. (2019). Uses and reuses of scientific data: The data creators' advantage. *Harvard Data Science Review*, 1. <https://doi.org/10.1162/99608f92.fc14bf2d>
- Pescosolido, B., Lee, B., & Kafadar, K. (2020). Replication data for: Socio-demographic similarity at the contextual level alters individual risk for completed suicide [Data set]. Harvard Dataverse. <https://doi.org/10.7910/DVN/35IV23>
- Popkin, G. (2019). Data sharing and how it can benefit your scientific career. *Nature*, 569, 445–447.
- Quintana, D. S. (2020). A synthetic dataset primer for the biobehavioral sciences to promote reproducibility and hypothesis generation. *Elife*, 9, e53275.
- R Core Team. (2019). *R: A language and environment for statistical computing*. <https://www.R-project.org/>
- Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *BMJ*, 340. <https://doi.org/10.1136/bmj.c221>
- Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ (clinical Research Ed.)*, 340, c332. <https://doi.org/10.1136/bmj.c332>
- Sholler, D., Ram, K., Boettiger, C., & Katz, D. S. (2019). Enforcing public data archiving policies in academic publishing: A study of ecology journals. *Big Data & Society*, 6, epub. <https://doi.org/10.1177/2053951719836258>
- Slep, A. M. S., Heyman, R. E., Lorber, M. F., Baucom, K. J. W., & Linkh, D. J. (2020). Evaluating the Effectiveness of NORTH STAR: A Community-Based Framework to Reduce Adult Substance Misuse, Intimate Partner Violence, Child Abuse, Suicidality, and Cumulative Risk. *Prevention Science*. <https://doi.org/10.gg898m>
- Stern, C., Jordan, Z., & McArthur, A. (2014). Developing the review question and inclusion criteria. *AJN the American Journal of Nursing*, 114, 53–56.
- Strasser, C. (2012). *Primer on data management: What you always wanted to know*.
- Ten Percent Happier: Mindfulness Meditation Courses. (2020). Ten percent happier. <https://www.tenpercent.com>
- Tennant, J. P., Crane, H., Crick, T., Davila, J., Enkhbayar, A., Havemann, J., Kramer, B., Martin, R., Masuzzo, P., Nobes, A., Rice, C., Rivera-López, B., Ross-Hellauer, T., Sattler, S., Thacker, P. D., & Vanholsbeeck, M. (2019). Ten hot topics around scholarly publishing. *Publications*, 7, 34. <https://doi.org/10.3390/publications7020034>
- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS One*, 6, e21101. <https://doi.org/10.1371/journal.pone.0021101>
- Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D., & Dorsett, K. (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS One*, 10, e0134826. <https://doi.org/10.1371/journal.pone.0134826>
- Thorogood, A., & Knoppers, B. M. (2017). Can research ethics committees enable clinical trial data sharing? *Ethics, Medicine and Public Health*, 3, 56–63.
- Tomczyk, S., Schomerus, G., Stolzenburg, S., Muehlan, H., & Schmidt, S. (2020). Ready, willing and able? An investigation of the theory of planned behaviour in help-seeking for a community sample with current untreated depressive symptoms. *Prevention Science*, 21(6), 749–760. <https://doi.org/10.gg898n>
- Transpose. (2020). *Transpose Database: A database of journal policies on peer review, co-reviewing, and preprinting*. <https://transpose-publishing.github.io/#/>
- van Panhuis, W. G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., Heymann, D., & Burke, D. S. (2014). A systematic review of barriers to data sharing in public health. *BMC Public Health*, 14, 1144. <https://doi.org/10.1186/1471-2458-14-1144>
- Vasilevsky, N. A., Minnier, J., Haendel, M. A., & Champieux, R. E. (2017). Reproducible and reusable research: Are journal data sharing policies meeting the mark? *PeerJ*, 5.
- Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, 88, 428–436. <https://doi.org/10.1016/j.jbusres.2017.12.043>
- Walters, W. H. (2020). Data journals: Incentivizing data access and documentation within the scholarly communication system. *Insights*, 33, 18. <https://doi.org/10.1629/uksg.510>
- Wellcome Trust Public Health Research Data Forum. (2012). *Summary of Funders' data sharing policies*. <https://wellcome.ac.uk/sites/default/files/summary-of-phrdrf-funder-data-sharing-policies.pdf>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., & Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wilson, S. J., & Tanner-Smith, E. E. (2014). Meta-analysis in prevention science. In Z. Sloboda & H. Petras (Eds.), *Defining prevention science* (pp. 431–452). Springer US. https://doi.org/10.1007/978-1-4614-7424-2_19
- Yardley, S. J., Watts, K. M., Pearson, J., & Richardson, J. C. (2014). Ethical issues in the reuse of qualitative data: Perspectives from literature, practice, and participants. *Qualitative Health Research*, 24(1), 102–113.
- Yenni, G. M., Christensen, E. M., Bledsoe, E. K., Supp, S. R., Diaz, R. M., White, E. P., & Ernest, S. K. M. (2019). Developing a modern data workflow for regularly updated data. *PLoS Biology*, 17, e3000125. <https://doi.org/10.1371/journal.pbio.3000125>
- Zenodo—Research Shared. (2020). <https://about.zenodo.org/>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Emily A. Hennessy¹  · Rebecca L. Acabchuk²  · Pieter A. Arnold³  · Adam G. Dunn⁴  · Yong Zhi Foo⁵  ·
Blair T. Johnson⁶  · Sonya R. Geange^{3,7}  · Neal R. Haddaway^{8,9,10}  · Shinichi Nakagawa⁵  · Witness Mapanga¹¹  ·
Kerrie Mengersen¹²  · Matthew J. Page¹³  · Alfredo Sánchez-Tójar¹⁴  · Vivian Welch¹⁵  · Luke A. McGuinness¹⁶ 

¹ Center for Addiction Medicine, Recovery Research Institute, Massachusetts General Hospital & Harvard Medical School, Boston, MA, USA

² Institute for Collaboration On Health, Intervention, Department of Psychological Sciences, University of Connecticut & Policy, Mansfield, USA

³ Research School of Biology, The Australian National University, Canberra, Australia

⁴ Biomedical Informatics and Digital Health, School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

⁵ Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

⁶ Department of Psychological Sciences, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020, USA

⁷ Department of Biological Sciences, University of Bergen, Bergen, Norway

⁸ Stockholm Environment Institute, Stockholm, Sweden

⁹ Mercator Research Institute On Global Commons and Climate Change, Berlin, Germany

¹⁰ Africa Centre for Evidence, Johannesburg, South Africa

¹¹ Faculty of Health Sciences, Non-Communicable Diseases Research Division of the Wits Health Consortium, University of the Witwatersrand, Johannesburg, South Africa

¹² School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia

¹³ School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

¹⁴ Department of Evolutionary Biology, Bielefeld University, Bielefeld, Germany

¹⁵ School of Epidemiology and Public Health, Bruyere Research Institute, University of Ottawa, Ottawa, Canada

¹⁶ Department of Population Health Sciences, University of Bristol, Bristol, UK