Lipids and dementia
An investigation of their relationship

Luke A McGuinness

University of Bristol

A thesis submitted for the degree of
Doctor of Philosophy in Population Health Sciences

TBC

Abstract

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

Lipids and dementia
An investigation of their relationship

Luke A McGuinness

University of Bristol

A thesis submitted for the degree of

Doctor of Philosophy in Population Health Sciences

TBC

For Brendan McHugh

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: .....
Luke McGuinness
Canynge Hall, Bristol
1 December 2021

Acknowledgements

Additionally, I would like to thank Sharen Oâ€™Keefe.
Thanks to the creator of the oxfordown which was used
There are so many people without whom this would not have been possible:
[**Check all name spellings**]
Sharen O'Keffe Neal Haddaway Martin Westgate
robvis contributors
Statistical support:
All authors who volunteered time and effort

Canynge Hall, Bristol
1 December 2021

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.

Contents

List of Figures

List of Tables

List of Abbreviations

DOI . . . . . Digitial object identifier

HDL . . . . . High Density Lipoprotein

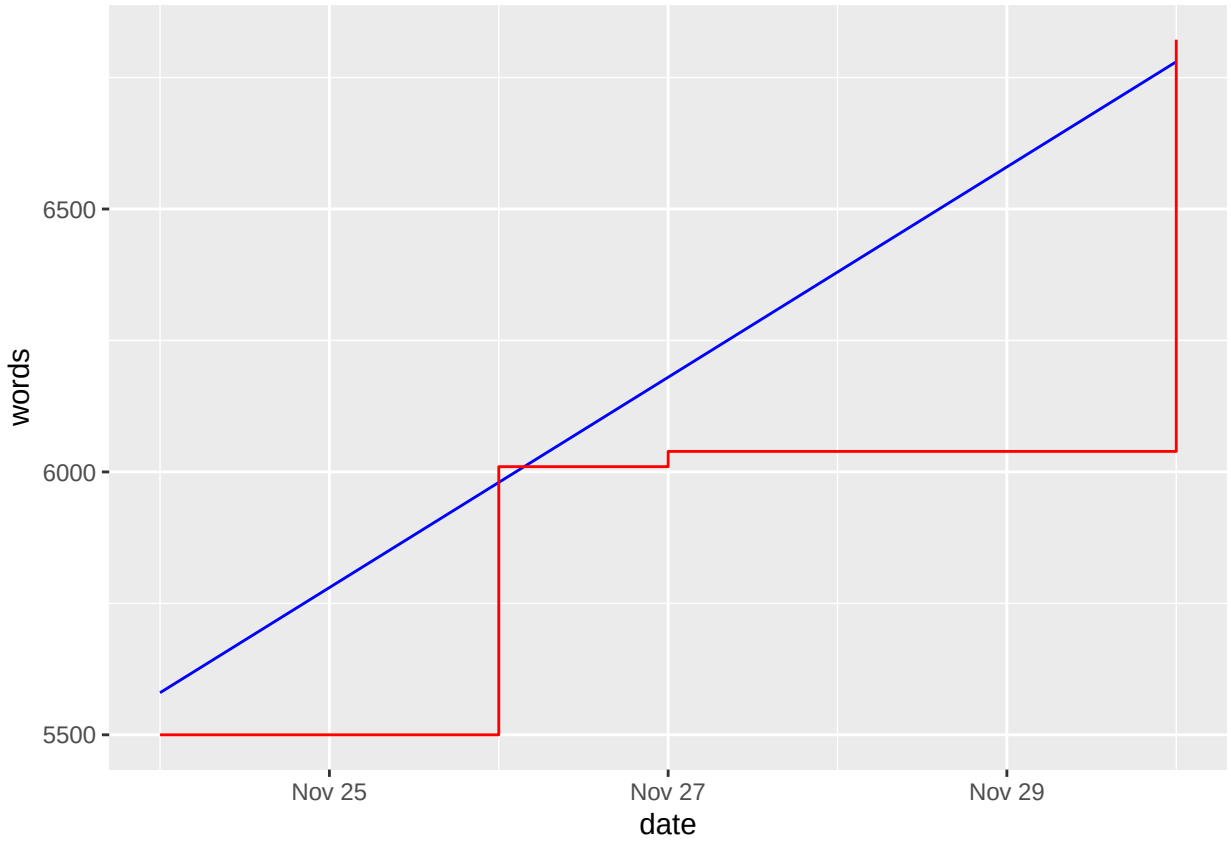
MR . . . . . Mendelian randomization

Covering material

Word count: 6822

Days: 6

Words behind: 42



1 Introduction

Contents

1.1 Aims and Objectives of the thesis

1.1.1 Aim

1.1.2 Objectives

- To review the published literature with respect to the effec tof lipids and lipid regulating agents

- 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

- To meta-analyze the

1.2 Chapter Overview

- **Chapter 2:** Background information on dementia and blood lipid levels. This chapter provides an introduction to the topics covered in this thesis to non-subject area experts, and discusses the motivation for the remainder of the thesis.

- **Chapter 3** This Chapter describes a comprehensive systematic review and meta-analysis of all available evidence on the relationship between blood lipids and lipid

- **Chapter 4:** 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, based in England.

- **Chapter @ref():** This Chapter seeks

- **Appendix 5** This Chapter introduces new evidence synthesis two tools built in R: `robvis` and `medrixvr`. The motivation for, development and impact of, and future plans for these tools are outlined.

1.3 Thesis Output

The outputs of this thesis are detailed below, and include published peer review papers, presentations at conferences, open source evidence synthesis tools, and . . .

1.3.1 Peer reviewed papers

Include other papers I have been involved in?

1.3.2 Papers under review

- McGuinness et al.

1.3.3 Software

A lot of packages were used in this project[1–16]

- `robvis`: An R package and associated `shiny` web application that allows users to easily visualise the results of risk of bias graphs.
- `medrxvir`: An R package and associated `shiny` web application that allows users to easily search and retrieve bibliographic data from the medRxiv[1] preprint repository.

1.3.4 Talks

- Talk at Cochrane on medrxivr
- Presentation to the
- Webinar to Evidence Synthesis Ireland on Risk of Bias

1.4 Summary

-
-
-

[1] `https://www.medrxiv.org/`

Science knows it doesn't know everything; otherwise, it'd stop

— Dara O'Briain

2 Background

Contents

2.1 Dementia

2.1.1 Definition and Subtypes

Dementia is major neurocognitive disorder, with symptoms including impairment of executive cognitive functions such as speech, judgement and memory. The most common causes of dementia are Alzheimer's disease and vascular dementia, accounting for ~60-80% and ~10% of cases respectively. The remaining 10-30% of cases are caused other dementia subtypes (e.g. Lewy Body dementia) or by progression of other neurological diseases (e.g. Parkinson's disease).

2.1.2 Diagnosis

Dementia is difficult to diagnose, primarily due to the absence of a gold standard test for the condition. Information from multiple diagnostic tools are utilised, from medical history examination, through assessment of patients mental ability (e.g. the Mini-Mental State Examination), to clinical tests (e.g. Magnetic Resonance Imaging (MRI) scans).

2.1.3 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,6 an ageing population looks set to create a dementia epidemic, particularly in Westernised countries. In the UK, there are estimated to be 800,000 people currently living with dementia, with this figure expected to double by 2040.7 Globally, the prevalence of dementia is expected to reach 75 million by 2030.8 Despite the number of dementia cases and decades of research, there remains much unknown about the pathogenesis and progression of the disease.

2.1.4 Challenges in the study of dementia using traditional study designs

Speak to the

Historically, studies on dementia have faced a range of challenges. As a range of determinants (genetic, environmental and lifestyle) are thought to jointly influence the risk, progression and outcomes of dementia, any individual association is likely to have only a small effect. Therefore, the statistical power need to detect these associations will often require sample sizes that are unfeasible using primary data collection. This fact is further complicated by the long latency of dementia, which necessitates a costly long-term approach to patient follow-up. Furthermore, dementia studies often have limited generalisability to the target population, as certain subgroups (e.g. the very old) are frequently under-represented in study cohorts due to difficulties associated with their recruitment.9 Finally, studies employing a case-control design are further limited by the high potential for differential recall bias between those who have and have not developed dementia.10

Fortunately, these limitations may potentially be addressed through the use of routinely collected health data, and the electronic health record databases in which they are stored.

Include CIND definition here

A further clinical subtype that is of particular import the

And MCI

And different scales

2.1.5 History

> Two major clinical trials are often cited as providing evidence that statins do not have an effect on the incidence ofdementia: the Prospective Study ofPravastatin in the Elderly (PROSPER) and the Medical Research Council/British Heart Foundation Heart Protection Study; however, because of methodologic limitations in relation to dementia outcomes in these two trials, the results of these trials are difficult to evaluate. Dementia incidence or cognitive outcomes were not preplanned endpoints in either of them, neither study included a clinical cognitive

> evaluation, and numbers of patients with follow-up information for cognitive evaluations were not reported in either study manuscript. In PROSPER a post hoc analysis compared changes in cognitive scores over a 3-year period between statintreated and placebo patients and found no significant differences. In the MRC/BHF HPS trial,27 similar percentages of participants (0.3% in each—statin vs placebo—group) developed dementia during the 5-year follow-up period. The report did not state how the outcome of dementia was determined (e.g., reported as an adverse event or by follow-up phone interview). *Taken from Cramer 2008 - included in review*

There exist only two large scale randomized controlled trials - in fact these studies

are the only two trials included in a 2016 review of statins for the prevention of

dementia produced by the Cochrane.[17] Both showed no effect of

Queries around how the data were collected, along

2.1.6 Economic impact

2.1.7 Risk factors

2.1.8 Treatments

2.1.9 Preventative measures

2.2 Serum lipids

2.2.1 Range of lipids

2.2.2 HDL-c

2.2.3 LDL-c

2.2.4 Triglycerides

2.2.5 Total cholesterol

2.3 Serum lipid interventions

2.3.1 Statins

Lipophilic

Hydrophilic

2.4 Methods used in the thesis

2.5 Evidence synthesis

2.5.1 Triangulation

2.5.2 Individual patient data meta-analyis

2.6 Summary

— Dara O'Briain — Dara O'Briain

Science knows it doesn't know everything; otherwise, it'd stop

— Dara O'Briain

3 Systematic Review

Contents

3.1 Additional ideas

- Evidence map - show the distribution of the different studies population across the world.
- Living systematic review approach - update weekly based on medRxiv

3.2 Aims

The aim of this chapter is to systematically review all available literature on the association between blood levels of total cholesterol and it's constituent parts (HDL-c,LDL-c and triglycerides) on the subsequent risk of dementia.

Based on the review of prev, no previous

Literature con

3.3 Methods

3.3.1 Search strategy

We will systematically search electronic bibliographic databases to identify potentially relevant records. The search strategy used in each database will be developed in an iterative manner using a combination of free text and controlled vocabulary (MeSH/EMTREE) terms to identify studies which have examined the relationship between blood lipids levels and dementia, incorporating input from an information specialist. The strategy will include terms related to lipids, lipid modifying treatments, and dementia and its sub-types, and will be designed for MEDLINE before being adapted for use in the other bibliography databases listed. An outline of the general strategy is presented in the Table 3.2 below and the full draft search strategies for each database are attached to this protocol. To ensure that the study design filters are not excluding potentially relevant records, a random sample of 500 records identified by the main search but excluded by the filters (defined as Line 7 NOT Line 13 in Table 3.2) will be screened. If any potentially relevant studies are identified, their titles and abstracts will be searched for key terms that can be incorporated into the filters to improve search sensitivity.

The following databases will be searched from inception onwards: Medline, EMBASE, Psychinfo, Cochrane Central Register of Controlled Trials (CENTRAL), and Web of Science Core Collection. We will also search clinical trial registries, for example ClinicalTrials.gov, to identify relevant randomized controlled trials.

The abstracts list of relevant conferences (e.g. the proceedings of the Alzheimer's Association International Conference, published in the journal Alzheimer's & Dementia) will be searched. Grey literature will also be searched via ProQuest, OpenGrey and Web of Science Conference Proceedings Citation Index, while theses will be accessed using the Open Access Theses and Dissertations portal. We will also search bioRxiv and medRxiv, preprint repositories using a tool built as part of this thesis, to identify potentially relevant studies. Finally, the reference lists of included studies will be searched by hand while studies citing included studies will be examined using Google Scholar (forward and reverse citation searching).

3.3.2 Study selection

Records will be imported into Endnote and deduplicated using the method outlined in Bramer et al. (2016).[18] Screening (both title/abstract and full-text) will be performed using a combination of Endnote and Rayyan, a web based screening application.[19] Title and abstract screening to remove obviously irrelevant records will be performed by the primary author, with a random selection of excluded records being screened in duplicate to ensure consistency with the inclusion criteria. If this demonstrates a significant level of erroneous exclusion by the primary author a larger proportion will be dual-screened. Full-text screening will also be completed in full by the primary author. A second reviewer will screen a random sample of included and excluded records, in addition to any records identified by the first reviewer as being difficult to assess against the inclusion criteria. Reasons for exclusion at this stage will be recorded. Disagreements occurring during either stage of the screening process will be resolved through discussion with a senior colleague. A PRIMSA flow diagram will be produced to document how records moved through the review.[20]

Inclusion

We will seek studies that examine the relationship between blood lipid levels (or any specific lipid fraction, including total cholesterol, HDL, LDL, and triglycerides) and risk of incident dementia/MCI. Eligible study designs include randomized controlled

trials and non-randomized observational studies of lipid modifying treatments, longitudinal studies examining the effect of increased/decreased blood lipid levels, and genetic instrumental variable (Mendelian randomization) studies examining the effect of genetically increased/decreased blood lipid levels.

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

Eligible studies will define dementia according to recognised criteria, for example the National Institute of Neurological Disorders and Stroke Association-Internationale pour la Recherche en l'Enseignement en Neurosciences (NINDS-AIREN), International Classification of Diseases (ICD), or Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria. For MCI, eligible studies are those that attempted state a definition for diagnoses of MCI (e.g. an adapted version of the Petersen criteria[21]) and create ordinal groups of patients (e.g. no dementia or dementia/MCI/dementia) based on this definition.

No limitations will be imposed on publication status, publication date, venue or language, although we will require sufficiently detailed reports of the studies to be able to examine their methods. Preprints and unpublished reports will be eligible for inclusion if relevant. Multiple publications resulting from the analysis of the same data will be included and grouped.

Exclusion

Case-control studies, cross-sectional studies, qualitative studies, case reports/series and narrative reviews will be excluded. Studies which present no evidence of attempting to exclude prevalent cases from their analyses will also be excluded. Studies that measure change in continuous cognitive measures (e.g. MoCA score) without attempt to map these scores to ordinal groups (e.g. no dementia/MCI/dementia) will be excluded. Conference abstracts with no corresponding full-text publication will be examined, and we will contact authors to obtain information on the study's

status. Studies that are reported in insufficient detail (e.g. only in conference abstracts, new, letters, editorials and opinion) will be excluded. Previous systematic reviews will not be eligible, but their reference lists will be screened to identify any potentially relevant articles. Studies with outcomes not directly related to the clinical syndrome of dementia (e.g., neuroimaging), studies implementing a "multi-domain intervention" where the lipid-regulating agent is included in each arms (e.g. for example, a study examining exercise + statins vs statins alone, but a study examining exercise + statins vs exercise alone would be included), and studies where there was no screening for dementia at baseline except if the sample was initially assessed in mid-life (i.e. below the age of 50) will be excluded.

Excluded studies performing autopsy unless it was done under accepted criteria Exclude studies using a dietary intervention, for example omega-3 fatty acid enriched diet, as it hard to disentangle the effect of other elements contained within te diet, vs simple tablet based supplements of

3.3.3 Data extraction

Harmonization of cholesterol measures across studies was performed, as different studies used different methods to quantify exposure, including comparing differing risks in the highest vs lowest quartiles of a lipid, using a binary classification of patients into a hypercholesterolaemia or not, categorising lipid levels into high, middle, and low groups according to study-defined criteria, and simply treating the exposure as a continuous variable.

3.3.4 Risk of bias assessment

Risk of bias assessment was performed using the domain-based risk-of-bias assessment tool appropriate to the study design. Randomised controlled trials were assessed using the RoB2 tool,[22], non-randomised studies of interventions were assessed using the ROBINS-I tool,[23] and non-randomised studies of exposures were assessed using the ROBINS-E tool.[TBC]

At present, no risk of bias assessment tool for Mendelian randomisation studies is available. Bias in thesee studies was assessed with the help of an expert panel [TBC]

Data will be visualized using a paired forest and risk of bias plot

3.3.5 Patient and public involvement

3.4 Results

3.4.1 Screening results

Following screening, XXX studies were included.

The distribution of included studies over time demonstrates that despite the conduct of several previous reviews of different types of literature surrounding this question, primary studies continue to the published as these reviews have yet to provide a definite answer.

Table **??** shows the characteristics [**Include column here that says whether it was included in a systematic review - see below**]

As part of our our forward snowballing exercise (where articles citing an included study are cited), we recorded whether a study included in our review had been included in any previous evidence synthesis attempt in an attempt to qualify the added value of this analysis. Additionally, if an included article was subsequently cited by a review, all studies in that review were screened for inclusion for the sake of completeness. This analysis was performed by extracting the citing articles from Google Scholar on XXXX and screening them manually. The DOI of articles extracted from this analysis are included in the appendix, as the Google Scholar search functionality is not readily reproducible.

As a summary of the duplication of work in this area, we looked at how many reviews a single included study had previously been included in.

Inter-rater reliability

Inter- and intra-rater reliability was asssessed for a 10% subsample of records at the title and abstract screening stage. Intra-rater reliability involved a sinlge reviewer

Table 3.1: Inter-rater reliability

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

applying the inclusion criteria to the same set of records while blinded to their previous decisions, while inter-rater reliability involved two reviewers independently screening the same set of records.

Rater reliability was assessed using Gwet's agreement coefficient (AC1).[24] This measure of inter-rater reliabilty was chosen over other methods of assessing inter-rater reliability such as percent agreeement (number of agreements divided by total number of assessments) as i account for chance agreement between reviewers but does not suffer from severely imbalanced marginal totals in the same way that Cohen's kappa value does. [25:[24,26]

How to interpret agreement co-efficients is widely debated. Here we use guidelines based on a stricter interpreation of the Cohen's Kappa coefficient,[27] presented in Table **??**.

Gwet's AC1 is calculated using the following:

-> Insert formula here. Need to be sure of how to calculate.[24,28]

Test

Kappa	Interpretation
0 – 0.20	None
0.21 – 0.39	Minimal
0.40 –.59	Weak
0.60 –0.79	Moderate
0.80–0.90	Strong
> 0.90	Almost perfect

For the inter-rater reliability, percentage agreement was 97.3% (AC1 = XXXX, Table 3.2), while for the intra-rater reliability, agreement was 98.6% (AC1 = XXXX, Table 3.2).

Table 3.2: Intra-rater reliability

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

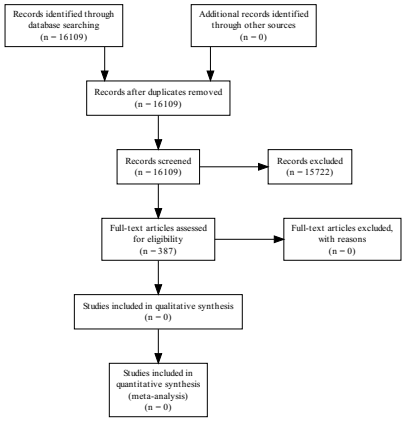


Figure 3.1: PRISMA flow diagram, detailing the movement of studies through the systematic review.

The discrepancy between the percent agreement and the associated value of AC is expected, due to the heavy imbalance in this sample towards exclusion.[29]

Those records which were excluded in the initial screening, but were included either by the same reviewer on their second viewing (n=4), or by the second reviewer (n=29), were investigated. This discrepancy between the two reviewers was explained in all cases by differing interpretations of the inclusion criteria, specifically around the definition of cognitive decline vs mild cognitive impairment and the definition of eligible lipids.

PRISMA flowchart

```
## pdf
```

Figure 3.2: PRISMA flow diagram, detailing the movement of studies through the systematic review.

```
##  2
```

Following de-duplication, the titles and abstracts of 16109 records were assessed for eligibility. 387 were deemed potentially eligible and the full text records for these were requested and screened. [**CROSS-REF to PRISMA flow here**]

A breakdown of records by

3.5 Included studies

XXXX studies met the criteria for inclusion in the review.

Include here:

- PRISMA Flowchart (use PRISMA2020 from GitHub - depending on level of involvement, could refer to the Appendix here too.)
- Summary of types of study
- Summary of locations
- Summary of diagnostic criteria used
- Summary of risk of bias
- Long table, horizontally. Newer version of flextable (GH) allows for PDF output

3.6 Publication bias

- Check if there are protocols available for any of the published reports, and whether there were
- Vascular dementia has substantially less published reports. Many (reference Smeeth et al 2010 here) simply group into AD and non-AD making comparison between published st

3.6.1 Triangualtion across evidence sources

One key question for which multiple distinct sources of evidence were available were those looking at Laz

A key limitiation for other types of dementia, in particularly vascular, is that there has yet to be a GWAS identifying relevant SNPS that could then be used in a MR study with SNPS for lipids to estimates the causal effect of lipids on vascular dementia [**Is this true?**] This rules out the use

3.7 Discussion

Of note, as part of the review, we identified serveral previous systematic reviews of this topic. However, this review is the firest to

3.8 Conclusion

"When dealing with human beings controlled experiments frequently prove to be impracticable, so for a scientific basis for our assumptions we turn to past history to reconstruct the suspected causal chain of events - and then our statistical troubles may begin"

— Harold F. Dorn, 1953[30]

4

Primary analysis of lipid regulating agents and dementia

Contents

4.1 Additional ideas

One of the problems I'll point out with the trials in the systematic review chapter (Chapter 3) is that they only included people with a high cardiovascular risk - we are kind of doing the same by using elevated test cholesterol results as the index event. Might be able to get around this by pointing out

4.2 Things we tried

- Conditioning entry to the cohort on QRISK score (as defined by codes) rather than. Need to be able to demonstrate here that not many meeting our original entry criteria actually go on to have a statin quickly. Need total number of those starting and average time to start.
- Positive controls. Need to be able to explain why these controls have such wildly increased results.
- Competing risk analysis, with death as a competing risk. Need numbers of deaths in each group, and preliminary results
- Allowing for time-varying confounders for binary covariates.
- Replicating a matched analysis a la Smeeth et al 2010 - though questions were raised as to how they had actually done the analysis (they did not match on propensity score, they simply adjusted for it - find paper that explores 4 ways of accounting for PS and shows this is the worst), they adjusted for things likely to be on the causal pathway, and they saw a huge change in direction for MI pre- vs post- adjustment.

4.3 Things we could try

- Work through data using positive control to see if the way it is being added is what is causing the issues.
- Examine Cramer et al. thesis and see what they did
- Marginal structural models approach (a la J. Sterne)

4.4 Aims

In this

4.5 Methods

4.5.1 Estimation methods

Potential biases included time varying confounding, selection bias due to censoring on death and We performed a Cox test this si to addd some words.

4.5.2 Estimating the value of the time-varying confounders

Mean time from index event to first prescription of statins was 2.4 years. This negates the promised benefit of ruling out confounding by indication (where the test result leads to the prescription of the treatment and also increases the risk of the outcome, distorting the relationship between the two), as there is no relationship between index TC/LDL-c and eventual LRA prescription.

Additionally, the time between index event and prescription does lead to a problem in terms of time varying confounding, as an average time of 2.4 years between current measurement of the covariates and treatment switching means there is plenty of time for the value of the covariate to change. This is problematic when the descision to change treatments (in this case to move from no LRA use to LRA use) is influenced by a set of prognostic factors that in turn may have been influenced by the inital treatment decision, as is likely to be the case for a range of covariates included in the model. For example:

`_No CVD (t=0) -> No LRA (t=0) -> CVD (t=1) -> LRA (t=1) -> Dementia (t=2)_`

In this case, the decision to move to LRA use is influenced by CVD status at *Time 1*, which will not be captured by adjusting only for CVD status at *Time 0*.

In practice, this means that the value of the prognostic factor should be regularly captured

However, in electronic health records, a change in the value of the prongostic factors is only important if it is recorded in a patients record, as for it to have an impact on treatment decisions, it must be recorded.

This means we can find the most recent value of the covariate before the switch and apply a marginal structural model approach, filling all values for that variable

before the most recent measure with the baseline measurement, and all after the most recent meausure with the value of the most recent measure (on the basis that you won't go from having CVD back to not having CVD).

i.e.

Timepoint 12345678

CVD 00001111

Treatment 00000111

Split into 3 month blocks since index event and use the same approach as above to work out the values of each covariate at each time point.

Note: this will be harder for things that are not dichotmous and can go up as well as down. Examples include total cholesterol and BMI, which can go up as well as down.

4.5.3 The effect of total cholesterol or LDL-cholesterol on LRA prescription

It would be fair to assume that the baseline total cholesterol/LDL-cholesterol would at least in part predict the likelihood of someone being prescribed a statin.

However, this is not the case. Baseline cholesterol level are predicted to be a poorer instrument for than QRISK2 score,[31] which estimates a patients' 10-year risk of a cardiovascular event. Current NICE guidelines state that those with a QRISK score of 10% or higher, and in whom lifestyle modifcation is ineffective/inappropriate, should recieved a lipid regulation agent. However, this analysis could not find any effect of QRISK2 scores on statins precription levels at 6 months. [Need to cite Lauren's eventual paper here focusing on QRISK2, but also display a RD analysis of TC/LDL-c levels here on statins at 6 months. Need also to check, as Lauren mentioned she found some evidence that there is a relationship in practices that acutally did what they should.]

As expected, in a confirmatory analysis using lipid levels, there was no association between the most recent total cholesterol or LDL-cholesterol reading in the CPRD and the treatment, incdicating that adjusting for this variable was not required. *

Why are open source statistical programming languages the best?

Because they R.

— Bealy, 2013[32]

5

Creating new systematic review tools in R

Contents

5.1 Background

Developing new tools for facilitating evidence synthesis will benefit patients by speeding up the speed with which systematic reviews are produced. A recent study put the mean duration of a systematic review project, from date of registration to date of publication is 67.3 weeks.[33] A large proportion of this time will be spent on manual tasks such as title and abstract screening, - therefore, efforts to increase the efficiency of the process should focus on process that can be automated.

Below I introduce two new R packages created to faciliate some aspects of the systematic review process.

5.2 Producing risk-of-bias visualisation with `robvis`

5.2.1 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.[34] 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.[34] These tools include the RoB 2 tool for randomized trials,[35] the ROBINS-I tool for non-randomized studies of interventions,[36] the QUADAS 2 tool for test accuracy and the ROBIS tool for systematic reviews.[37] Within each 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.[35] 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.[38]

Researchers can face a number of barriers in creating these plots. While some evidence synthesis platforms, such as Cochrane's Review Manager,[39] 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,[40] 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,[41] 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.[42–44] Here, we present `robvis` (Risk Of Bias VISualiation),[45] 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`).

5.2.2 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.[46–50]

5.2.3 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")
```

5.2.4 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 5.1

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 5.1 is created using:

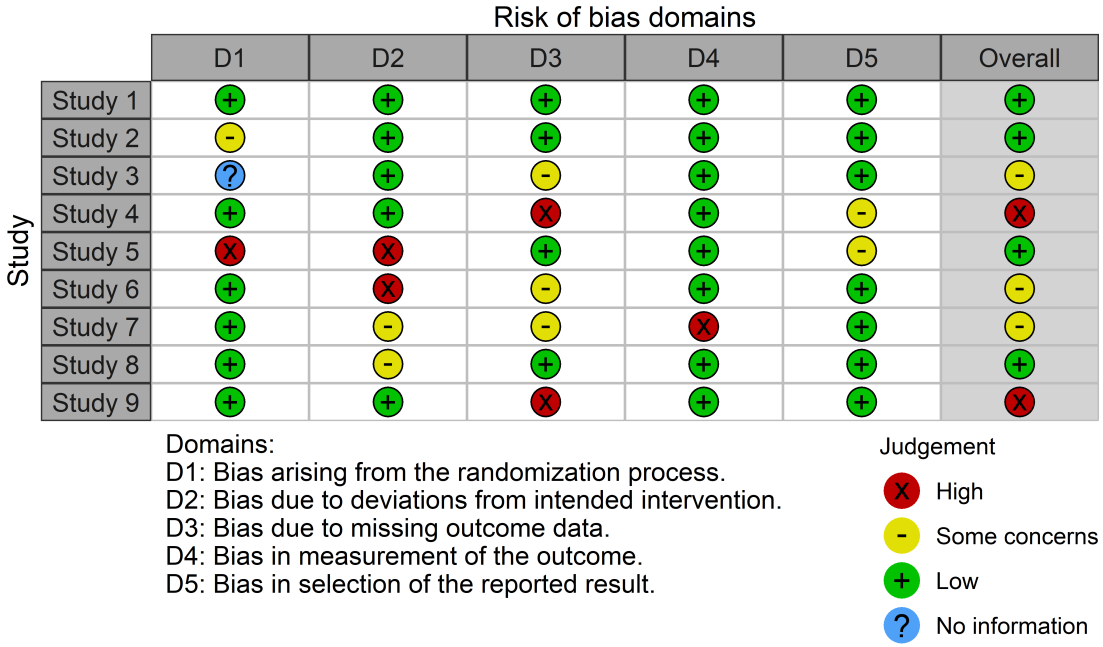


Figure 5.1: Example risk of bias traffic light plot created using 'robvis'

-> NEED TO ADD DATA HERE!

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

Similary, using the same data set, the summary barplot shown in Figure 5.2 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 5.1

Table 5.1: Description of the arguments available in the two main 'robvis' functions. 'X' indicates that the option is available for the respective function. *(continued)*

Argument	rob_traffic_light()	rob_summary()	Description

Table 5.1: Description of the arguments available in the two main 'robvis' functions. 'X' indicates that the option is available for the respective function.

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

5.2.5 Reception and Future Plans

As of November 2020, `robvis` has been downloaded more than 8700 times. It has been well recieved but the systematic review community, and has been cited

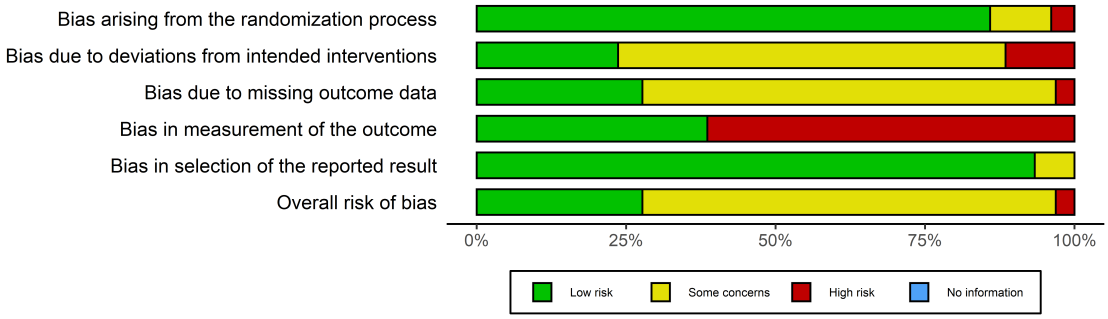


Figure 5.2: Example risk of bias summary plot created using 'robvis' and the example ROB2 dataset

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 in to the "Doing Meta-Analysis in R" online textbook.[51]

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 developement.[52] 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.[34] 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.

5.3 Searching for health-related preprints with `medrxivr`

5.3.1 Introduction

Searching pre-print repositories is becoming an increasingly important part of a systematic review. Preprints - unpublished versions of manuscripts, frequently uploaded to a repository at the same time they are submitted to a journal for peer review - represent an important source of grey literature.

However, while several of the main repositories (e.g. arXiv, bioRxiv) have existing methods by which records can be searched and downloaded en masse,[1] this is not true for medRxiv, a repository launched in June 2019 to host preprints in the medical, clinical, and related health sciences.

At present, medRxiv allows only simple search queries, as opposed to the often complex Boolean logic that information specialists use to query other major databases. Additionally, record metadata (titles/abstracts/author lists) must be accessed individually, rather than in batches, meaning that downloading relevant records for title and abstract screening a time consuming task.

This source of information is particularly important to this thesis, in that one of the primary types of evidence that I planned to use in the triangulation exercise (Mendelian randomization (MR)) is still developing and many MR studies that look at the clinical question that this project aims to investigate are posted as preprints on the bioRxiv, and more recently medRxiv, preprint servers. Being able to systematically search these records for the purpose of the systematic review described in Chapter @(ref:sys-rev-heading) was a necessity. In light of the above, the `medrxivr` R package was created in order to facilitate the searching of this data source.

5.3.2 Development

Early versions of the tool

[1] `https://arxiv.org/help/api`

5.3.3 Usage

To install `medrxivr` from CRAN:

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

To install the development version from GitHub:

```
devtools::install_github("ropensci/medrxivr")
```

5.3.4 Methods

The `medrxivr` project is split into two parts:

- A webscraper that runs once a day which collects and cleans new records uploaded to medRxiv, and adds them to a machine readable database.
- A lightweight R package that provides an interface for this database, allowing users to search for relevant records, and easily retieve the associated full tex PDFs.

The webscraper is a straightforward R script built in R using `rvest` and a range of text processing packages[53–55]. To quickly illustrate the process, the code to retrieve the total number of records on medRxiv is included below:

```
# Perform a search for "*" (defined by %252A), which acts as a wildcard
# This approach will capture all records
url <- "https://www.medrxiv.org/search/%252A"

# Read the HTML from this link
page <- read_html(url)

# Extract the text (html_text()) value of the CSS node of interest, in this
# case, the "#page-title" node

results <- page %>%
  html_nodes("#page-title") %>%
```

```
  html_text()

# Remove any commas from the results (e.g. 1,500 become 1500)
results <- gsub(",","",results)

# Convert it from a character string to a numeric
results <- as.numeric(word(results))
```

Using Windows Task Scheduler,[56] the script is scheduled to automatically run every morning, adding new records uploaded to medRxiv since the last run of the webscraper to a local comma-seperated-values (CSV) file. Following a quality control pipeline, which checks the completeness of the new records and ensures that all records have been captured, the updated dataset is automatically uploaded to a cloud server and immediately becomes available to the `medrxivr` package functions. The functions in the `medrxivr` package then facilitate users in working with this dataset. There are two main functions and a helper function:

- `mx_search()` [main]: Enables users to search the medrxivr data dump, using regular expressions and boolean logic.

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

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

results <- mx_search(myquery)
```

- `mx_download()` [main]: Takes the output from `mx_search()` and retrieves the full text PDF for each record, saving it to a folder specified by the user.

```
mx_download(results,        # Object returned by mx_search
            "pdf/")         # Directory to save PDFs to
```

- `mx_crosscheck()` [helper]: Provides information on the version of the data dump that the user is searching, and checks whether any new records have been uploaded to medRxiv since the last run of the webscraper.

```
mx_crosscheck()


## Using medRxiv snapshot - 2020-03-04 10:55
## No new records added to medRxiv since last snapshot.
```

5.4 Export

One of the key

5.5 Data visualisation and reporter functions

TBC

5.5.1 Reception and Future Plans

The tool has be well received by the community, and several use cases have been reported. However, several other use cases have emerged for the tool. It has been used to visalise the growing number of preprints related to the 2019 coronavirus outbreak,[2] etc, etc

Following peer-review, it has been onboarded into the rOpenSci suite of packages, a collection of "carefully vetted, staff- and community-contributed R software tools that lower barriers to working with scientific data sources on the web", and an associated article published in the Journal of Open Source Software.

A short paper describing the tool has been published in the Journal of Open Source Software following formal peer-review of the package and associated documentation and tutorials as part of the rOpenSci onboarding process. The entire review discussion can be viewed online.[3]

[2] https://twitter.com/L_Brierley/status/1233109086444695553
[3] https://github.com/ropensci/software-review/issues/380

The tool has been used reasonably widely, included in the systematic reviews, but also to extract data to facilitate the compare the behavior of researchers under two different journal policies by comparing the preprint paper with the published journal paper.

Feedback from the systematic review community has been positive.

5.5.2 Package infrastructure

The `medrxvir` package was written in R using RStudio, and followed development best practices, including complete and information documentation, a robust unit testing framework (99% of all code lines within the package are formally tested under this framework) across multiple platforms including Windows, MacOS, and Linux, and in-depth code review by two experienced reviewers.

The medRxiv snapshot is taken every morning using GitHub Actions, an automated system fo repetitive tasks.

5.6 Shiny app

Part of the key theme of accessibility meant creating a web-application that l

5.7 Dicussion

Packaging and sharing R scripts should be a fundamental part of evidence synthesis process. [EXPAND]

As

By implementing the tools described above as both as an R package and a `Shiny` web app, the functionality is available to evidence synthesists with varying levels of ability in R. These tools serve as an example of the advantages of "packaging" the R scripts that evidence synthesists often create for personal use.[57] In the case of `robvis`, it is likely that several other evidence synthesists have written scripts to produce similar risk-of-bias plots to those presented here - in fact, in the course of its development, I identified 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 timeframe, 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.[58] 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,[59] of which `robvis` and `medrivr` are 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.[60–63]

5.8 Summary

- In this Chapter, I have introduced two new tools for facilitating evidence syntheses: `robvis`, a tool for producing publication quality risk of bias assessments, and `medrxivr`, a tool for performing complex searches in the medRxiv preprint repository.

- I have outlined the motivation for developing these tools in relation to this thesis - more specifically, that they were used extensively to the conduct the comprehensive systematic review described in Chapter 3

- The impact of these packages to date, their place in the broader evidence synthesis in R ecosystem, and a roadmap for their future development has been discussed.

— Dara O'Briain — Dara O'Briain

Science knows it doesn't know everything; otherwise, it'd stop

— Dara O'Briain

6 Discussion

6.1 Summary of findings (and implications for policy makers)

6.2 Strengths and Limitations

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

6.3 Reproducible research

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

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

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

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). However, access is dependency on an ISAC application to the managing body of the CPRD.

6.4 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

6.5 Future work

6.6 Conclusions

7
Bibliography

1. R Core Team. *R: A language and environment for statistical computing*. (R Foundation for Statistical Computing, 2019).

2. Csárdi, G. *Cranlogs: Download logs from the 'rstudio' 'cran' mirror*. (2019).

3. Wickham, H., François, R., Henry, L. & Müller, K. *Dplyr: A grammar of data manipulation*. (2020).

4. Gohel, D. *Flextable: Functions for tabular reporting*. (2020).

5. Wickham, H. *Ggplot2: Elegant graphics for data analysis*. (Springer-Verlag New York, 2016).

6. Hester, J. *Glue: Interpreted string literals*. (2020).

7. Müller, K. *Here: A simpler way to find your files*. (2017).

8. Zhu, H. *KableExtra: Construct complex table with 'kable' and pipe syntax*. (2020).

9. Xie, Y. *Knitr: A general-purpose package for dynamic report generation in r*. (2020).

10. it, W. wrote. *Mal: What the package does (title case)*.

11. McGuinness, L. A. & Schmidt, L. Medrxivr: Accessing and searching medRxiv and bioRxiv preprint data in r. *Journal of Open Source Software* **5**, 2651 (2020).

12. Wickham, H. The split-apply-combine strategy for data analysis. *Journal of Statistical Software* **40**, 1–29 (2011).

13. Jack O. Wasey. *PRISMAstatement: Plot flow charts according to the "prisma" statement.* (2019).

14. McGuinness, A, L., Higgins & PT, J. Risk-of-bias visualization (robvis): An r package and shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods* (2020). doi:10.1002/jrsm.1411

15. Wickham, H. *Stringr: Simple, consistent wrappers for common string operations.* (2019).

16. Marwick, B. *Wordcountaddin: Word counts and readability statistics in r markdown documents.* (2020).

17. McGuinness, B., Craig, D., Bullock, R. & Passmore, P. Statins for the prevention of dementia. *Cochrane Database of Systematic Reviews* (2016). doi:10.1002/14651858.CD003160.pub3

18. 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).

19. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews* **5**, 210 (2016).

20. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration | The BMJ.

21. Petersen, R. C. *et al.* Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology* **56**, 303–308 (1999).

22. Sterne, J. A. C. *et al.* RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, (2019).

23. Sterne, J. A. *et al.* ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355**, (2016).

24. 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).

25. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* **20**, 37–46 (1960).

26. 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).

27. McHugh, M. L. Interrater reliability: The kappa statistic. *Biochemia Medica* **22**, 276–282 (2012).

28. Sim, J. & Wright, C. C. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy* **85**, 257–268 (2005).

29. 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).

30. Dorn, H. F. Philosophy of Inferences from Retrospective Studies. *American Journal of Public Health and the Nations Health* **43**, 677–683 (1953).

31. Hippisley-Cox, J. *et al.* Predicting cardiovascular risk in England and Wales: Prospective derivation and validation of QRISK2. *BMJ* **336**, 1475–1482 (2008).

32. Bealy, C. R Programming Humour. *Stack Exchange* (2013).

33. Borah, R., Brown, A. W., Capers, P. L. & Kaiser, K. A. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* **7**, e012545 (2017).

34. Boutron, I. *et al.* Chapter 7: Considering bias and conflicts of interest among the included studies. in *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated August 2019)* (eds. Higgins, J. P. T. et al.) (Cochrane, 2019).

35. Sterne, J. A. C. *et al.* RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, l4898 (2019).

36. Sterne, J. A. C. *et al.* ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355**, i4919 (2016).

37. Whiting, P. F. *et al.* QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* **155**, 529–536 (2011).

38. Higgins, J. P. T. & Altman, D. G. Chapter 8: Assessing risk of bias in included studies. in *Cochrane Handbook for Systematic Reviews of Interventions* (eds. Higgins, J. P. T. & Green, S.) 187–241 (Wiley Online Library, 2008).

39. Cochrane Collaboration. Review Manager (RevMan) [Computer program]. (2014).

40. Marshall, C. & Brereton, P. Systematic review toolbox: A catalogue of tools to support systematic reviews. in *Proceedings of the 19th international conference on evaluation and assessment in software engineering* Article 23; 1–6 (Association for Computing Machinery, 2015). doi:10.1145/2745802.2745824

41. Harrison, H., Griffin, S. J., Kuhn, I. & Usher-Smith, J. A. Software tools to support title and abstract screening for systematic reviews in healthcare: An evaluation. *BMC Medical Research Methodology* **20**, 7 (2020).

42. R Core Team. *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing, 2019).

43. RStudio Team. *RStudio: Integrated development environment for R.* (RStudio, Inc., 2015).

44. Chang, W., Cheng, J., Allaire, J., Xie, Y. & McPherson, J. *Shiny: Web application framework for R.* (2019).

45. McGuinness, L. A. Robvis: An R package for visualising risk-of-bias assessments. [Computer Software - v0.3.0]. (2019). doi:10.5281/zenodo.3552342

46. Gibb, K. *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* 19005165 (2019). doi:10.1101/19005165

47. Habadi, M. I. *et al.* Prevalence of panic disorders in the primary health care setting: A systematic review and meta-analysis. *EC Microbiology* **16**, 01–09 (2019).

48. Veloso, A., Vicente, S. G. & Filipe, M. G. Effectiveness of cognitive training for school-aged children and adolescents with attention Deficit/Hyperactivity disorder: A systematic review. *Frontiers in Psychology* **10**, 2983 (2020).

49. 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).

50. 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).

51. Harrer, M., Cuijpers, P. & Ebert, D. *Doing Meta-Analysis in R: A Hands-on Guide.* (Zenodo, 2019).

52. Whiting, P. *et al.* ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology* **69**, 225–234 (2016).

53. Wickham, H. Rvest: Easily Harvest (Scrape) Web Pages. (2019).

54. Wickham, H. & RStudio. Stringr: Simple, Consistent Wrappers for Common String Operations. (2019).

55. Wickham, H., François, R., Henry, L., Müller, K. & RStudio. Dplyr: A Grammar of Data Manipulation. (2020).

56. Task Scheduler for developers - Win32 apps.

57. Wickham, H. *R packages: Organize, test, document, and share your code.* (O'Reilly Media, Inc., 2015).

58. Xie, Y., Allaire, J. J. & Grolemund, G. *R markdown: The definitive guide.* (Chapman and Hall/CRC, 2018).

59. Various Authors. Metaverse: An R ecosystem for meta-research. (2020).

60. Grames, E. M., Stillman, A. N., Tingley, M. W. & Elphick, C. S. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. *Methods in Ecology and Evolution* **10**, 1645–1654 (2019).

61. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 1–48 (2010).

62. Pick, J. L., Nakagawa, S. & Noble, D. W. A. Reproducible, flexible and high-throughput data extraction from primary literature: The metaDigitise R package. *Methods in Ecology and Evolution* **10**, 426–431 (2018).

63. Westgate, M. J. Revtools: An R package to support article screening for evidence synthesis. *Research Synthesis Methods* **10**, 606–614 (2019).

Appendices

A
By Chapter

A.1 Chapter 1

A.2 Chapter 2

B

Other Appendix