

“It is surely a great criticism of our profession that we have not organised a critical summary by speciality or sub-speciality, up-dated periodically, of all relevant RCTS.”

1

Systematic review of all available evidence

1.1 To Do

- Examine how many times a primary study was included in a review
- Look at whether searching preprints made any difference to my results
- Could apply ROBIS to previous reviews to see how they stack-up?

1.2 Lay summary

1.3 Additional ideas

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

1.4 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

1.5 Methods

1.5.1 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).

1.5.2 Study selection

Records will be imported into Endnote and deduplicated using the method outlined in Bramer et al. (2016).[@bramer2016] Screening (both title/abstract and full-text) will be performed using a combination of Endnote and Rayyan, a web based screening application.[@ouzzani2016] 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.[@zotero-766]

Inclusion criteria

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 [petersen1999]) 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 criteria

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

1.5.3 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, categorizing lipid levels into high, middle, and low groups according to study-defined criteria, and simply treating the exposure as a continuous variable.

1.5.4 Risk of bias assessment

Risk of bias assessment was performed using the domain-based risk-of-bias assessment tool appropriate to the study design. Randomized controlled trials were assessed using the RoB2 tool,[@sterne2019], non-randomized studies of interventions were assessed using the ROBINS-I tool,[@sterne2016] and non-randomized studies of exposures were assessed using the ROBINS-E tool.[CITE]

At present, no risk of bias assessment tool for Mendelian randomization studies is available. Bias in these studies was assessed with the help of an expert panel drawn from my supervisors and other external invited researchers.[CITE]

Data will be visualized using a paired forest and risk of bias plot, created using the `robvis` tool. This tool was developed as part of this thesis to aid in creating standard risk-of-bias figures, and is summarized in Appendix ??.[@mcguinness2019]

1.5.5 Patient and public involvement

1.6 Results

1.6.1 Previous reviews

Will need to include something in the methods section on this

As part of this analysis, we examined the overlap between different published reviews on this topic.

Several primary studies were captured in multiple systematic reviews. However, some studies were not captured by any previous review,

Put upSet plot here.

An upset plot shows the total size of each set (in this case, the set of included studies in each review) in the bottom left hand bar plot. For example, the XXXX review includes XXXX studies.

Where a line joins two points in the matrix, the main barchart shows the number of records shared by these reviews. So for example, XXXX records were captured both by the XXXX review and the XXXX review.

Where the matrix has just a single point highlighted, this shows the records unique to that review.

The reviews are ordered by date, to attempt to show the increasing overlap and total number of studies

Our review is presented as the last point on this plot, and captured XXXX records
The alternative is to have a barplot showing the total cumulative number of records over the years, using our set as the master, with the bars coloured by the number of records included in 1/2/3/4 reviews. The information below the bar axis could

then show the number of reviews published in that year, or alternative, you could have an upsidetdown bar to show cumulative number of reviews on this topic.

1.6.2 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 be 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 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- and intra-rater reliability was assessed for a 10% subsample of records at the title and abstract screening stage. Intra-rater reliability involved a single reviewer applying the inclusion criteria to the same set of records while blinded to their previous decisions, while inter-rater reliability involved two reviewers independently screening the same set of records.

Rater reliability was assessed using Gwet's agreement coefficient (AC1).[@gwet2008]
This measure of inter-rater reliability was chosen over other methods of assessing

inter-rater reliability such as percent agreement (number of agreements divided by total number of assessments) as 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. [Cohen1960; Gwet2008; Wongpakaran2013]

How to interpret agreement co-efficients is widely debated. Here we use guidelines based on a stricter interpretation of the Cohen’s Kappa coefficient,[McHugh2012] presented in Table 1.1.

Table 1.1: Agreement for kappa

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

In a two by two table with cells A, B, C and D, Gwet’s AC1 is calculated using the following:

$$AC1 = \frac{p - e(\gamma)}{1 - e(\gamma)}$$

Here, $p = \frac{A+D}{N}$ and $e(\gamma)$ is the chance agreement between raters, given as $2q(1 - q)$, where $q = \frac{(A+C)+(A+B)}{2N}$

-> Insert formula here. Need to be sure of how to calculate. [Gwet2008; Sim2005]

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

The discrepancy between the percent agreement and the associated value of AC is expected, due to the heavy imbalance in this sample towards exclusion.[Feinstein1990]

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

Figure 1.1: PRISMA Flowchart: Overview of how records moved through the systematic review process

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.

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. 1.1

1.6.3 Characteristics of 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

1.6.4 Risk of bias

1.6.5 Sources of heterogeneity

Detail that some of these are exploratory, in particular the effect of different scales on the association between the groups. [CROSS]

1.6.6 Publication bias

- Check if there are protocols available for any of the published reports (unlikely for non-randomised controlled trials), 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

1.7 Triangulation across evidence sources

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

A key limitation 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. In addition, there was limited literature available - as discussed in CHapter ??

Consideration of the potential impact of the magnitude and direct of residual confounders/bias is not a major stretch from what is already happening in the assessment of the quality of evidence (GRADE) framework. Within GRADE, the overall quality of evidence can be upgraded when there is deemed to be unmeasured or residual confounding variables which reduce the. For example, if the propensity to treatment is related to comorbidity burden, but those on treatment still have better outcomes then those on control, it is likely that the true effect of the intervention is being underestimated. [guyatt2011]

Without our framework and as part of the risk of bias assessments reported in 1.6.4, we test test

1.8 Discussion

Major limitation is that several included studies used data from electronic health databases, which come with serious concerns regarding validity [hsieh2019] [ALSO CITE WILKINSON AND MCGUINNESS HERE].

In addition, the a

1.8.1 Comments on the process

Systematic reviews should not be performed as part of a thesis, without suitable support and resourcing guaranteed. Assumption that everyone does a systematic

review (without risk of bias assessments, inclusion of all literature, searching for other reviews) is foolish.

Great learning experience in terms of managing a team, but due to a rotating schedule of people, much harder to do that expected.

A final point on the progress of the review is that, due to the need for dual screening and data extraction, a number of external researchers became involved in this review. I found the people-management aspect particularly challenging and could definitely have improved the process through better communication of deadlines, but it has provided good experience of leading a review team.

1.8.2 Original contribution to research

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

- *Comprehensiveness*: While several reviews of this research topic exist,[@chu2018b; @yang2020; @muangpaisan2010; @poly2020b] 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.
- *Structured risk of bias assessment*: The majority of the highly cited reviews on this topic either do not formally consider the risk of bias in the observational studies they include or do not use an appropriate domain-based assessment tool (e.g. ROBINS-I/E). This is important area in which this thesis can add value, as based on the risk-of-bias assessments I have performed to date, several primary studies are at high risk of bias and this should be reflected in the findings of any review on this topic.
- *Inclusion of preprints*: Unlike other available reviews and enabled by the tool described in Chapter 3, this review systematically searched preprinted health-related manuscripts as a source of grey literature. As part of this

chapter, I plan to examine the extent of the additional information provided to the review by the inclusion of preprints.

1.8.3 Comparison with other reviews

Of note, as part of the review, we identified several previous systematic reviews of this topic.[CITE] However, this review is the first to use established domain based assessments tools (for example, the RoB 2 tool for randomized controlled trials)[@sterne2019] to assess the risk of bias in included studies, and explore the heterogeneity of results across different levels of risk of bias levels.

Some previous reviews did assess risk of bias, but used non-domain based assessment tools. Newcastle-Ottawa scale,

The duplication of work across reviews (including, ironically, by this review) is substantial. In Section XXXX, I demonstrated that the each primary study included in this review was also included in other review so this topic, but that not all studies were included in all reviews. However, by creating a comprehensive review that attempted to draw together all available evidence from across the full range of study types.

1.8.4 Strenghts and limitations

Test

1.9 Conclusion

Test