Efficacy and safety of pioglitazone and vitamin E in non-alcoholic steatohepatitis. An overview of systematic reviews.

A thesis submitted in fulfilment of the requirements for the degree of Master of Science in Medical Research Methodology

By

Nikolaos Giotis

Thessaloniki, May 2020
Contents
Abstract ................................................................................................................................. 4
BACKGROUND .......................................................................................................................... 5
Description of the condition .................................................................................................. 6
Description of the interventions and how they might work .................................................. 7
Pioglitazone .............................................................................................................................. 7
Vitamin E ................................................................................................................................ 7
Why it is important to perform this overview ....................................................................... 7
OBJECTIVES .......................................................................................................................... 8
METHODS .................................................................................................................................. 8
Protocol .................................................................................................................................... 8
Eligibility criteria .................................................................................................................... 8
Search methods for the identification of systematic reviews ................................................ 9
Search strategy for the MEDLINE via PubMed .................................................................... 10
Study selection and data extraction ...................................................................................... 10
Quality appraisal and assessment of the evidence ............................................................... 10
Data synthesis ........................................................................................................................ 11
Patient and public involvement ............................................................................................ 12
RESULTS ................................................................................................................................... 12
Literature search .................................................................................................................... 12
Characteristics of included studies case by case .................................................................. 16
Characteristics of excluded studies ...................................................................................... 21
Quality appraisal .................................................................................................................... 22
Patient characteristics ......................................................................................................... 23
Interventions and comparators ............................................................................................ 23
Outcomes .................................................................................................................................. 23
DISCUSSION ............................................................................................................................... 27
Strengths and limitations ....................................................................................................... 30
Conclusion .............................................................................................................................. 30
Implications for future research .......................................................................................... 30
References ................................................................................................................................ 31
Appendix A. PRIO harms checklist ....................................................................................... 38
Appendix B. Search strategy for MEDLINE via PubMed ...................................................... 42
Appendix C. Data extraction list ............................................................................................ 43
Appendix D. Quality appraisal per AMSTAR 2 ...................................................................... 44
Efficacy and safety of pioglitazone and vitamin E in non-alcoholic steatohepatitis. An overview of systematic reviews.

An MSc thesis submitted in fulfillment
of the requirements for the degree of
Master of Science in Medical Research Methodology

At

The Faculty of Health Sciences
School of Medicine
Aristotle University of Thessaloniki

By

Nikolaos Giotis

Thessaloniki May 2020

Anna-Bettina Haidich

9435 words
Commonly appearing abbreviations

NAFLD: non-alcoholic fatty liver disease
NASH: non-alcoholic steatohepatitis
LSI: Lifestyle intervention
SR(s): Systematic review(s)
RCT(s): Randomized controlled trial(s)

All studies are identified by the name of the first author and the publication year
Abstract

Background

This overview aimed to assess the evidence regarding the clinical effectiveness and safety of pioglitazone and vitamin E as adjuncts to standard lifestyle intervention in adult patients with non-alcoholic steatohepatitis and to determine if research or guideline adaptation is needed.

Methods

PubMed, The Cochrane Library, and DARE were searched for systematic reviews of randomized controlled trials from inception until January 21, 2020, along with relevant grey literature sites. Following two levels of screening, data extraction, quality appraisal, and assessment of overlapping, the results for all-cause mortality, liver-related morbidity, histological improvement of hepatic fibrosis or inflammation features, and serious adverse events were displayed in tables.

Results

A total of 14 systematic reviews were included after screening 1945 titles and abstracts and 22 full-text articles. Three of the included systematic reviews were rated as high quality using the Assessment of Multiple Systematic Reviews 2 (AMSTAR2) tool, one as low and ten as critically low quality. No events were reported in the short term (18 months) regarding the clinical outcomes of mortality, liver morbidity, and serious adverse events, making it difficult to assess the risk-benefit ratio for either intervention. Data for improvement of inflammatory features were consistent for pioglitazone and non-existent for vitamin E. Data for reversal of fibrosis was inconsistent for pioglitazone and non-existent for vitamin E.

Conclusions

There is no clinical evidence to support the addition of pioglitazone or vitamin E to the standard lifestyle intervention in adult patients with non-alcoholic steatohepatitis.

Keywords

Overview of systematic reviews, non-alcoholic fatty liver disease, non-alcoholic steatohepatitis, pioglitazone, vitamin E.
BACKGROUND

Non-alcoholic fatty liver disease (NAFLD) comprises two pathologic entities: non-alcoholic fatty liver (NAFL) and non-alcoholic steatohepatitis (NASH). In NAFL, hepatic steatosis affects > 5% of hepatocytes with no or mild inflammation present. The defining feature of NASH, besides hepatic steatosis and inflammation, is histologic evidence of hepatocyte injury manifested as ballooning hepatocyte degeneration.

The distinction between NAFL and NASH is of paramount importance because the two entities carry different prognoses. While NAFL runs a relatively benign course, NASH, on the other hand, can progress to hepatic fibrosis and cirrhosis, liver failure, and hepatocellular carcinoma. Therefore, it is necessary to identify and treat those patients with NASH.

The most prominent scientific societies in the field agree that weight loss is the most effective treatment for NAFLD (1–5). A bodyweight reduction of 7 – 10%, through energy restriction (i.e., hypocaloric diet) or regular physical exercise, improves or even resolves steatosis and inflammation and regresses fibrosis in patients with NASH (6,7). Regarding pharmacotherapy, there are no officially approved drugs for the treatment of NASH yet. All societies agree that any clinical use of any medication in this context is considered off-label. However, when it comes to appreciating the usefulness of certain medications, significant inconsistencies emerge regarding two agents: pioglitazone and vitamin E.

On the one hand, the American Association for the Study of Liver Diseases (AASLD) and the National Institute for Health and Care Excellence (NICE) favor the use of pioglitazone in patients with (biopsy-proven) NASH regardless of type 2 Diabetes Mellitus (T2DM) status, following either a discussion with the patient over the risks and benefits (AASLD) or informed consent (NICE) (1,4). On the other hand, both the European Association for the Study of the Liver (EASL) and the Asian – Pacific Party (AAP) do not recommend pioglitazone for general use in NASH, except for selected patients with T2DM (2,5). Finally, the Italian Association for the Study of the Liver (AISF) acknowledges the existing evidence over pioglitazone but avoids making any recommendation at all (3).

Regarding vitamin E, while the AASLD considers its use strictly in nondiabetic adults with biopsy-proven NASH, following discussion with the patient (1), the NICE regards its administration in advanced liver fibrosis irrespective of T2DM status (4). The EASL, AISF, and AAP societies do not make any firm recommendation; however, the former allows for its use in nondiabetic and non-cirrhotic NASH patients (2,3,5) keeping more in line with the AASLD.

Both pioglitazone and vitamin E have received extensive research over the last years investigating their role in NASH, with several relevant randomized controlled trials and systematic reviews having been published. However, their results appear not to be decisive enough to prompt the issue of
consistent guidance by the involved scientific societies. Therefore, to facilitate the achievement of a broader consensus over the role of those two agents in NASH, we performed an overview of systematic reviews trying to draw together information from systematic reviews of those two pharmacologic interventions to treat NASH, attempting to answer the following PICO question: among patients with NASH (population), does the combination of lifestyle intervention and either pioglitazone or vitamin E (intervention), as compared to lifestyle intervention alone (comparator), result in i) improved all-cause mortality, ii) reduced liver-related morbidity (namely, liver cirrhosis, liver failure, and hepatocellular carcinoma), iii) increased risk of serious adverse events, and iv) histopathologic improvement or reversal of NASH (outcomes)?

Description of the condition

NASH development is perceived as a two-hit model, with the first hit being a metabolic disturbance resulting in increased inflow of fatty acids to the liver as well as increased de novo lipogenesis in the liver, eventually culminating to hepatic steatosis. The second hit involves increased oxidative stress and induction of proinflammatory cytokines leading to inflammation and fibrosis (8).

NAFL results when caloric intake exceeds caloric expenditure creating an energy surplus that is being deposited in the form of non-esterified fatty acids (NEFAs) in sites outside the adipose tissue, such as the liver (9). Under these conditions, the rate of triglyceride inflow to and synthesis within the liver exceeds the rate of hepatic triglyceride oxidation and VLDL secretion (9). Apart from dietary fats and sugars, insulin resistance (IR) is also a significant contributing factor to the hepatic fat content through the failure to suppress, at the adipose tissue, the peripheral lipolysis of triglycerides (10).

Hepatic steatosis, especially in genetically predisposed individuals (11), may induce hepatocyte injury and death through increased oxidative stress and mitochondrial dysfunction (12–14). In its turn, hepatocellular death triggers inflammation and fibrosis (15,16), further resulting in the progression of NAFL to NASH at first and cirrhosis finally. The process is slow: it takes 57 years for NAFL and 24 years for NASH to progress to cirrhosis (17).

The prevalence of NAFLD in the USA is 24% (18). The prevalence of NASH in the general population is unknown; however, the prevalence of NASH among NAFLD patients in the USA is 21% (18), suggesting that the prevalence of NASH in the general population in the USA is 3-4% (18). It becomes evident that most patients with NAFLD have simple steatosis rather than NASH. Morbidities that frequently co-exist with NASH are obesity (82%), hyperlipidemia (82%), metabolic syndrome (76%), hypertension (70%) and T2DM (48%) (18).

The most common cause of death in NAFLD patients is cardiovascular disease (19). Malignancy is the second most common cause of death (20), with NAFLD being the second most common cause of Hepatocellular Carcinoma (HCC) in patients listed for liver transplantation (21). Of note, 50% of NAFLD
related HCC occurs in patients without cirrhosis, resulting in its late diagnosis (22). Liver-related mortality ranks third as a frequent cause of death in patients with NAFLD (23).

Description of the interventions and how they might work

Pioglitazone
Pioglitazone is an insulin-sensitizer approved for the treatment of T2DM. NAFLD is closely associated with insulin resistance and T2DM (18); therefore, it is rational to anticipate a positive role for pioglitazone in the treatment of NAFLD. At a molecular level, however, things are not that straightforward. Pioglitazone is recognized as an agonist of the Peroxisome Proliferator-Activated Receptor gamma (PPARγ), a nuclear receptor that modulates lipid and glucose metabolism through regulation of gene expression. It is not fully understood how pioglitazone exerts its hypoglycemic and insulin-sensitizing effects. It seems that it acts on the adipose tissue, where PPARγ promotes adipocyte differentiation and facilitates fatty acid storage in adipocytes, thus reducing free fatty acid secretion in the bloodstream and diminishing hepatic influx of fatty acids (24). Since the increased fatty acid delivery to the liver is a contributing factor to hepatic steatosis (8), pioglitazone could, at least theoretically, reverse this process and have a beneficial effect on hepatic steatosis. In the liver, PPARγ is expressed mainly on stellate and Kupfer cells, and its activation results in their inhibition, promoting anti-inflammatory action (25–30); thus, it has the potential to protect NAFL from progressing to NASH.

Vitamin E
Fat accumulation in the liver, as a result of increased free fatty acid inflow and increased de novo lipogenesis, results in increased fatty acid β-oxidation in the mitochondria, which translates into excessive electron flux in the respiratory chain and overproduction of reactive oxygen species (ROS). The excess of ROS, on the one hand, damages the macromolecules of the respiratory chain as well as the mitochondrial DNA, resulting in mitochondrial dysfunction and further production of ROS. On the other hand, the excess ROS causes the peroxidation of polyunsaturated free fatty acids, generating metabolites that induce inflammatory cytokines that flare up the inflammation and fibrinogenesis processes, resulting in the long run in end-stage liver disease (31). Vitamin E intercepts the fatty acid peroxidation (32), and thus can, at least theoretically, improve inflammation and fibrinogenesis, protecting against cirrhosis.

Why it is important to perform this overview
Allowing for regional variations, almost one in four people have NAFLD, a rather high prevalence that parallels the obesity and T2DM epidemics. While there isn’t any approved drug treatment for the condition, the keen clinician gets conflicting guidance by the respective scientific societies over the role of pioglitazone and vitamin E, in a way that the following question arises: does the NASH patient stand to gain any clinical benefit from pioglitazone or vitamin E when either drug is added to the
current standard of treatment, i.e., weight loss through lifestyle intervention? By clinical benefit, we mean a longer lifespan or reduced risk of hepatic morbidity like end-stage liver disease.

The existing body of evidence over pioglitazone and vitamin E in NASH comprises a perplexing landscape where relevant systematic reviews match or even outnumber the available randomized clinical trials; narrative reviews are countless, and the most prestigious scientific societies issue discordant guidance. We believe that an overview of systematic reviews under the angle of our PICO question will do some tidy-up of the evidence and aid in a broader consensus over the role of those two agents in NASH.

OBJECTIVES

To carry out an overview synthesizing systematic reviews over pioglitazone and vitamin E for the treatment of patients with NASH. We aimed to:

- Search systematically for and gather all the relevant evidence from corresponding systematic reviews
- Appraise the quality of the evidence by standard tools
- Present an overview table juxtaposing the results
- Answer the PICO question, stating explicitly whether the evidence suggests that pioglitazone or vitamin E offer any clinical gain, cause any harm, or provide only histopathological improvement.

METHODS

Protocol

This overview was commissioned in fulfillment of the requirements for the degree of Master of Science in Medical Research Methodology. A protocol for the overview of systematic reviews was written a priori by the MSc student in consultation with the supervisor and the two members of the advisory committee. Results were reported using the Preferred Reporting Items for Overviews of Systematic Reviews Including Harms (PRIO-harms) checklist (33) (Appendix A).

Eligibility criteria

Eligibility criteria for the overview were tailored to the Population, Intervention, Comparator, Outcome, and Study design framework to include the following:

- Patients: adult participants ≥ 18 years of age, irrespective of diabetic status, with a histologic diagnosis of NASH and no other concurrent liver disease or exposure to hepatotoxic substances (e.g., alcohol) or drugs (e.g., methotrexate). Statins were excluded.
Interventions: Pioglitazone and lifestyle intervention or vitamin E and lifestyle intervention. Lifestyle intervention (LSI) is defined as any advice on diet modification or regular aerobic exercise, aiming at weight reduction.

Comparators: lifestyle intervention alone.

Effectiveness Outcomes: Primary ones comprised i) reduction in all-cause mortality and ii) decrease in liver-related morbidity, defined as hepatic cirrhosis, end-stage liver disease, and hepatocellular carcinoma. Secondary outcomes comprised i) histopathological reversal of fibrosis and ii) histopathological improvement in NASH, which is defined as the reversal of balloon degeneration and hepatic inflammation. Reversal of fibrosis was chosen because it is meant to serve as a surrogate marker for all-cause and liver-related mortality as well as for liver-related morbidity (34,35). Histopathological reversal of NASH was added because it is the gold standard of diagnosis and disease activity compared with biochemical markers such as aminotransferases and imaging studies such as hepatic ultrasonography (2).

Harms Outcomes: We focused on the serious adverse events related to the long term use of pioglitazone and vitamin E, which have the potential of precluding their use in this clinical context: Pioglitazone-related: congestive heart failure (36), bone fractures (37,38) and bladder cancer (39). Vitamin E related: prostate cancer (40).

Study design: systematic reviews (SR) of randomized controlled trials (RCTs) irrespective of meta-analysis (MA) conduction. An eligible study had to demonstrate a systematic search of at least two databases to identify all primary studies that met a pre-determined set of eligibility criteria. Study selection and data extraction had to be performed in duplicate by independent researchers. Publications identified themselves as rapid reviews, literature reviews, narrative reviews, or other non-systematic knowledge syntheses were excluded from the overview.

Other: Both published and unpublished systematic reviews were eligible for inclusion, as well as publications in any language.

Search methods for the identification of systematic reviews

We identified published literature by searching the following databases: MEDLINE, through the PubMed interface, the Cochrane Database of Systematic Reviews, and the Database of Reviews of Effects (DARE) of the Centre of Reviews and Dissemination of the University of York, from inception until January 21, 2020. We scanned the references of the retrieved systematic reviews to identify additional ones. We checked for ongoing systematic reviews by searching the PROSPERO database for relevant protocols. We tried to identify unpublished (grey) literature by searching the websites of the EASL, the AASLD and the Asian Pacific Association for the study of the Liver, as well as the website of the Takeda pharmaceutical company, which manufactures the brand-name drug under the trademark ACTOS®. Below, we present the search strategy for PubMed; we applied the same one with the
appropriate amendments to the Cochrane Database of Systematic Reviews. We performed a broader search in the DARE database aiming at retrieving all reports containing the variations of the terms related to the NAFLD due to the overall small number of registered studies meeting the strict criteria of the Centre for Reviews and Dissemination.

Search strategy for the MEDLINE via PubMed
The search strategy contained both free-text and MeSH terms for the key concepts of NAFLD, NASH, pioglitazone, and vitamin E, without restriction on the language or publication date. Additional filters for humans and systematic reviews were applied at the end. The search strategy was reviewed against the PRESS 2015 evidence-based checklist (41). A detailed syntax tailored to the PubMed interface is presented in Appendix B.

Study selection and data extraction
The records retrieved from each database were entered into Mendeley to remove the duplicate ones. At the first level of the selection process, the grossly irrelevant records were excluded on account of their titles and abstracts. The full texts of the remaining publications were obtained and screened against the eligibility criteria.

A data extraction form was created in Excel format. Abstracted items included:

- Review characteristics, such as year of publication, date range of literature search, number of databases searched, designs of primary studies, and date range of primary studies.
- Patient characteristics, such as number, age group (adults or children), country of origin, and means of NAFLD or NASH diagnosis (histology, liver chemistry, or imaging).
- Interventions, especially whether pioglitazone or vitamin E, was combined with any advice of lifestyle modification.
- Comparators used, such as placebo or lifestyle modification or another active comparator.
- Outcomes examined, such as the name, definition, and measure of the outcome, method of analysis/synthesis, and statistical significance.
- Other, such as funding sources and conflicts of interest. The complete extraction list is presented in Appendix C.

Quality appraisal and assessment of the evidence
Assessing the quality of the primary studies included in each systematic review was out of the scope of this overview. However, data regarding the tool used to evaluate the methodological quality of the primary studies, as well as the respective result was extracted as reported in each eligible systematic review.

The methodological quality of each systematic review was appraised by applying the Assessing the Methodological Quality of Systematic Reviews tool version 2 (AMSTAR2) (42).
The GRADE approach was meant to rate the strength of evidence for each outcome (43). To minimize the subjectivity of this method, a GRADE algorithm developed for Cochrane overviews of reviews was used instead (44). In this algorithm, each review is assumed to have an initial ranking of high certainty, and it is subsequently downgraded one or more levels for serious methodological concerns regarding the following well-defined domains.

- The number of participants within pooled analysis: a range between 100 and 199 mandates downgrading by 1 level and below 100 mandates lowering by two levels. This is a way of ascertaining the imprecision of a treatment effect.
- The proportion of participants included in the pooled analysis judged to have a low risk of bias for randomization and observer blinding: a percentage of < 75% mandates downgrading by one level. If such data were not reported in the primary studies, we assumed the worst-case scenario; that is, less than 75% of the participants had a low risk of bias.
- Heterogeneity, as assessed by the \( I^2 \) statistic: a value of >75% mandates downgrading by one level. If \( I^2 \) was not reported, we assumed the worst-case scenario, which is a value >75%, and downgraded the ranking accordingly. If only one trial contributed to the analysis, there was no downgrade. This is a way of evaluating inconsistency arising from statistical heterogeneity.
- Responses to AMSTAR2 questions regarding the a priori research design (explicitly described in a protocol), the comprehensiveness of the literature search, the study selection in duplicate, and the data extraction in duplicate). A score of three “yes” mandates downgrading by one level and a score of two “yes” by two levels. This is a way of evaluating the risk of bias due to limitations in the design or conduct of the study.

Based on the above domains, we intended to rate our assessment of the strength of evidence for each comparison according to the following scale:

- High certainty (further research is unlikely to change our confidence in the estimate of effect)
- Moderate certainty (new research may impact significantly on our confidence in the estimate of effect)
- Low certainty (further research is very likely to change the estimate of effect)
- Very low certainty (the true effect is probably markedly different from the estimated effect).

However, the GRADE assessment was not applied as planned because, in the end, the best data of interest was not derived from pooled analyses, but from individual primary studies.

**Data synthesis**

No formal statistical analysis was planned for this overview as substantial overlapping was anticipated across the included systematic reviews, and pooling the data would not have been appropriate in this situation. Instead, a list of all the primary studies included in the eligible systematic reviews was...
compiled and cross-linked with the individual reviews in a citation matrix, with the rows representing the primary studies and the columns the systematic reviews. Grey-shaded cells along every row indicated in how many systematic reviews each primary study was included. Subsequently, the overlap measures of Covered Area (CA) and Corrected Covered Area (CCA) were calculated to demonstrate the degree of overlap among the systematic reviews (45). The CCA cut-off values of 0-5, 6-10, 11-15, and >15 were used to characterize the overlap as slight, moderate, high, and very high, respectively (45).

In the case of overlapping systematic reviews, the decision of which one best addressed our PICO question rested on how well the PICO elements of the systematic review matched our PICO question. Among equivalent (in terms of PICO elements) systematic reviews, the choice rested on the following prioritizing criteria:

- The most updated systematic review, as judged by the date of the last search.
- The systematic review with the higher methodological quality, as judged by the assessment of the risk of bias.

The above selection process was performed for each comparison and outcome separately, aiming at answering our research question with the best available data.

**Patient and public involvement**

Patients or the public were not involved in the development, design, or conduct of this research.

**RESULTS**

**Literature search**

One thousand nine hundred forty-nine references were identified through electronic searches of PubMed (n = 1633), Cochrane Database of Systematic reviews (n = 309), and DARE (n = 7). After the removal of 5 duplicates, 1944 were scanned. One thousand nine hundred twenty-three clearly irrelevant ones were excluded from going through the titles and abstracts. The full texts of 22 publications were retrieved for further assessment, of which one reference was identified from scanning the reference lists of the systematic reviews. No additional records were identified through other sources. Eight studies were then excluded (see characteristics of excluded studies). In total, 14 systematic reviews met the inclusion criteria (Fig. 1), consisting of:

- 11 Systematic Reviews; 8 with meta-analyses: Amanullah 2019 (46), Musso 2017 (47), Said 2017 (48), He 2016 (49), Sato 2015 (50), Boettcher 2012 (51), Musso 2012 (52) and Mahady 2010 (53); 3 without meta-analyses: Shyangdan 2011 (54), Socha 2009 (55) and Lirussi 2007 (56)
- One attempted, and two completed Network meta-analyses (NMA): Lombardi 2017 (attempted) (56), Sawangjit 2016 (57), and Singh 2015 (58).

The included studies spanned more than a decade, from 2007 to 2019. The latest literature search extended up to July 2018 (Amanullah 2019). The first authors of the studies based in Europe (Italy, the UK, Poland) in six cases, with the rest being in the USA (four), China (one), Australia (one), Pakistan (one), and Japan (one). The number of included primary studies per systematic review ranged from two to nine.

Regarding pioglitazone, a total of 10 primary studies were cited 60 times across 11 systematic reviews, resulting in a CCA = 0.50; it indicates considerable overlap across the included reviews. Many of them included the same primary studies. A detailed citation matrix for pioglitazone is displayed in table 1.

Regarding vitamin E, a total of 16 primary studies were cited 44 times across nine systematic reviews, producing a CCA of 0.22, which indicates considerable overlap across the included reviews. A detailed citation matrix for vitamin E is given in table 2.
Table 1: Citation matrix for pioglitazone studies. Shaded cells indicate the inclusion of a primary study in the SR; the number in the bracket corresponds to the number of the quality criteria assessed; the number outside the bracket equals the criteria met, as reported in each SR; All studies are identified by the first author and the publication year. NASH: Non-alcoholic Steatohepatitis. Piogl: Pioglitazone. Rosigl: Rosiglitazone. LSI: Lifestyle Intervention. Vit E: Vitamin E. Metf: Metformin.

<table>
<thead>
<tr>
<th>Primary Studies</th>
<th>Participants (Number)</th>
<th>Intervention (Duration in months)</th>
<th>Outcome</th>
<th>Systematic Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belfort 2006</td>
<td>NASH (n=52)</td>
<td>Piogl &amp; LSI vs Placebo &amp; LSI (D=6)</td>
<td>Liver histology (Kleiner score)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Aithal 2008</td>
<td>NASH, non-diabetic (n=74)</td>
<td>Piogl &amp; LSI vs Placebo &amp; LSI (D=12)</td>
<td>Liver histology (Brunt score)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Ratzl 2008</td>
<td>NASH, non-diabetic (n=64)</td>
<td>Rosigl &amp; LSI vs Placebo &amp; LSI (D=12)</td>
<td>Liver histology (Brunt score)</td>
<td>? (6)</td>
</tr>
<tr>
<td>Idilman 2008</td>
<td>NASH, non-diabetic (n=50)</td>
<td>Rosigl &amp; LSI vs LSI (D=12)</td>
<td>Liver histology (Brunt &amp; Kleiner score)</td>
<td></td>
</tr>
<tr>
<td>Sanyal 2010</td>
<td>NASH, non-diabetic (n=163)</td>
<td>Piogl vs Placebo (D=24)</td>
<td>Liver histology</td>
<td>? (6)</td>
</tr>
<tr>
<td>Omer 2010</td>
<td>NAFLD &amp; prediabetes/ T2DM (n1=42, n2=42)</td>
<td>Rosigl &amp; LSI vs Metf &amp; LSI (1) or Rosigl &amp; Metf &amp; LSI (2) (D=12)</td>
<td>Liver histology (Kleiner score)</td>
<td></td>
</tr>
<tr>
<td>Jin 2010</td>
<td>NASH (n=120)</td>
<td>Piogl &amp; &quot;Conventional care&quot; vs &quot;Conventional care&quot; (D=6)</td>
<td>Liver chemistry (ALT, AST, γGT)</td>
<td></td>
</tr>
<tr>
<td>Yan 2015</td>
<td>NAFLD (n=184)</td>
<td>Piogl &amp; LSI vs LSI (D=4)</td>
<td>Liver imaging (hepatic fat content)</td>
<td></td>
</tr>
<tr>
<td>Cusi 2016</td>
<td>NASH / prediabetes or T2DM (n=101)</td>
<td>Piogl &amp; LSI vs LSI (D=18)</td>
<td>Liver histology</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 Citation matrix for Vitamin E studies. Shaded cells indicate the inclusion of a primary study in the SR; the number in the bracket corresponds to the number of the quality criteria assessed; the number outside the bracket equals the criteria met, as reported in each SR; All studies are identified by the first author and the publication year. HQ = High Quality; LQ = Low Quality. NASH: Non-alcoholic steatohepatitis. NAFLD: Non-alcoholic Fatty Liver Disease. IFG: Impaired Fasting Glucose. T2DM: Type 2 Diabetes Mellitus. Vit E: Vitamin E. Vit C: Vitamin C. LSI: Lifestyle Intervention. UDCA: Ursodeoxycholic Acid. Metf: Metformin

<table>
<thead>
<tr>
<th>Primary Studies</th>
<th>Population (Number)</th>
<th>Intervention (Duration in months)</th>
<th>Outcome</th>
<th>Key PICO Elements</th>
<th>Systematic Reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vajro 2004</td>
<td>Obese Children with elevated transaminases and bright liver on US</td>
<td>Vit E &amp; LSI vs Placebo &amp; LSI (D=5)</td>
<td>Liver Ultrasonography &amp; Liver chemistry</td>
<td>HQ (4)</td>
<td>5 (5)</td>
</tr>
<tr>
<td>Sanyal 2004</td>
<td>NASH, non diabetic (n=20)</td>
<td>Vit E vs Vit E &amp; Pioglu</td>
<td>Liver histology</td>
<td>LQ (4)</td>
<td>5 (5)</td>
</tr>
<tr>
<td>Bugianesi 2005</td>
<td>NAFLD adults (n=83)</td>
<td>Vit E vs Metf (D=12)</td>
<td>Liver Histology</td>
<td>HQ</td>
<td>4 (8)</td>
</tr>
<tr>
<td>Dufour 2006</td>
<td>NASH Adults (n=48)</td>
<td>UDCA &amp; Vit E vs UDCA &amp; Placebo vs Placebo (D=24)</td>
<td>Liver Histology</td>
<td>LQ (4)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Nobili 2006</td>
<td>NAFLD children (n=88)</td>
<td>Vit E &amp; vit C vs Placebo (D=12)</td>
<td>Liver chemistry</td>
<td>8 (8)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Wang 2008</td>
<td>Not Reported</td>
<td>Not Reported</td>
<td>Non histological outcomes</td>
<td>6 (7)</td>
<td>5 (5)</td>
</tr>
<tr>
<td>Nobili 2008</td>
<td>NAFLD children (n=53)</td>
<td>Vit E &amp; Vit C &amp; LSI vs Placebo &amp; LSI (D=24)</td>
<td>Liver Histology</td>
<td>8 (8)</td>
<td>6 (7)</td>
</tr>
<tr>
<td>Sanyal 2010</td>
<td>NASH without T2DM (n=167)</td>
<td>Vit E vs Placebo (D=24)</td>
<td>Liver histology</td>
<td>8 (8)</td>
<td>7 (7)</td>
</tr>
<tr>
<td>Lavine 2011</td>
<td>Children with NAFLD (n=116)</td>
<td>Vit E vs Placebo (D=24)</td>
<td>Liver Histology and chemistry</td>
<td>8 (8)</td>
<td>7 (7)</td>
</tr>
<tr>
<td>Foster 2011</td>
<td>NAFLD adults (n=80)</td>
<td>Vit E &amp; Vit C &amp; atorv &amp; asp vs Placebo (D=48)</td>
<td>Non histological outcomes</td>
<td>8 (8)</td>
<td>7 (7)</td>
</tr>
<tr>
<td>Magosso 2013</td>
<td>NAFLD (US) adults (n=87)</td>
<td>Vit E vs Placebo (D=12)</td>
<td>Liver ultrasonography</td>
<td>7 (8)</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Hoofnagle 2013</td>
<td>NAFLD adults (n=139)</td>
<td>Vit E &amp; Pio vs Placebo (D=30)</td>
<td>Liver histology</td>
<td>4 (8)</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Basu 2014</td>
<td>NAFLD/NASH (n=75)</td>
<td>Vit E vs Placebo</td>
<td>Fibrosis and steatosis markers</td>
<td>4 (8)</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Han 2014</td>
<td>NAFLD with IFG</td>
<td>Vit E &amp; LSI &amp; Metf vs Bicyclol &amp; LSI &amp; Metf</td>
<td>Liver histology</td>
<td>5 (7)</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Aller 2015</td>
<td>NAFLD adults (n=36)</td>
<td>Vit E &amp; LSI &amp; Silymarin vs Placebo (D=3)</td>
<td>Non histological outcomes</td>
<td>4 (8)</td>
<td>3 (5)</td>
</tr>
</tbody>
</table>
Characteristics of included studies case by case

In the Lirussi 2007 SR (56), the participants were adults or children with NAFLD; the intervention comprised any antioxidant supplement; the comparator was no intervention, placebo or other active intervention; the outcomes included all-cause mortality, hepatic related mortality, radiological response, biochemical response, and histological response. The search extended up to June 2006, spanned six electronic databases and identified – among a total of six – three potentially relevant RCTs, published between 2003 and 2005, that examined Vitamin E in the intervention arm. However, none of the rest comparison elements matched exactly our PICO (table 2), due to variations in the participants (pediatric in the Varjo 2004), the co-interventions (Vitamin C in Harisson 2003), the comparators (Metformin in Bugianesi 2005) and the outcomes (Liver imaging and biochemistry in Varjo 2004). Regarding quality rating, one RCT was deemed high quality (Harisson 2003) and the other two low quality (Varjo 2004, Bugianesi 2005) across four domains (generation of the allocation sequence, allocation concealment, blinding, follow-up).

In the Socha 2009 SR (55), the participants included both adults and children with NAFLD or NASH; the intervention comprised any pharmacologic agent; the comparator was any other intervention; the outcome included normalization of aminotransferases and liver histology. The search extended up to December 2006, spanned three electronic databases, and identified – among a total of fifteen – four potentially relevant RCTs that examined Vitamin E in the intervention arm and two RCTs that examined pioglitazone in the intervention arm. The four Vitamin E-related RCTs were published between 2003 and 2006 and comprised the same three RCTs captured by the Lirussi 2007 study and an additional one (Nobili 2006) that included children. Regarding quality rating, two RCTs (Varjo 2004 and Harisson 2003) scored five and two (Nobili 2006 and Bugianesi 2005) scored four across five domains (allocation concealment, method of randomization, blinding, intention-to-treat analysis, follow-up). The two pioglitazone-related RCTs were published between 2004 and 2006, and both included adults with biopsy-proven NASH as participants and histological responses as outcomes. However, variations existed over the co-interventions (Vitamin E in Sanyal 2004, LSI in Belfort 2006) and comparators (Vitamin E in Sanyal 2004, Placebo, and LSI in Belfort 2006) (table 1). Both scored 4/5 in the same domains of quality appraisal (Randomization, Allocation, Blinding, Intention-to-treat, and Follow up).

In the Mahady 2010 SR (53), the participants were adults with biopsy-proven NASH; the intervention comprised any TZD (Tro-, Rosi- or Pioglitazone); the comparator was any intervention; the outcome was the histological response. The search extended up to May 2010, spanned two electronic databases, and identified – among a total of seven – four potentially relevant RCTs that examined pioglitazone in the intervention arm. They were published between 2004 and 2010 and two of them
(Sanyal 2004 and Belfort 2006) had been captured by the Socha 2009 SR. While all RCTs included adults with NASH and reported their results in terms of liver histology, significant variations existed in the co-interventions: LSI was implemented only in two (Belfort 2006 and Aithal 2008) and vitamin E in one (Sanyal 2004). The placebo served as the comparator in three (Belfort 2006, Aithal 2008, Sanyal 2010), but it was combined with LSI only in two (Belfort 2006, Aithal 2008) (table 1). The risk of bias was assessed against six domains (sequence generation, allocation concealment, blinding, handling of outcome data, intention-to-treat analysis, sponsorship); however, it was not reported individually for the included primary studies. Four meta-analyses were performed over equal histologic outcomes combining data from three pioglitazone and two rosiglitazone studies.

In the Shyangdan 2011 SR (54) the participants included both adults and children with NAFLD; the interventions included metformin, pioglitazone and rosiglitazone alone or in combination with other treatments versus no treatment, placebo or other pharmacologic interventions; the outcomes comprised hepatic related morbidity, hepatic histological features (fibrosis and cirrhosis), adverse events as well as cardiovascular events, new diabetes and quality of life. The search extended up to June 2010, spanned six electronic databases, and identified – among a total of fifteen – four relevant RCTs that examined pioglitazone in the intervention arm. They were published between 2004 and 2010 and matched precisely the ones captured by the Mahady 2010 SR (Table 1). Despite this, important variations over quality rating and statistical handling occurred. Regarding quality rating, only one (Sanyal 2010) was judged as low risk of bias and the other three (Sanyal 2004, Belfort 2006, Aithal 2008) as high risk across eight domains (randomization, allocation concealment, blinding, intention-to-treat analysis, the percentage who completed the trial, power calculation, similarity of groups at baseline and sponsorship). Regarding statistical handling, the four RCTs were deemed too heterogeneous to be combined into a single meta-analysis, contrary to Mahady 2010 study.

In the Musso 2012 SR (52), the participants included both adults and children with NAFLD; the interventions comprised any non-surgical treatment for NAFLD versus any comparator; the outcomes of interest included liver-related morbidity (cirrhosis, liver failure, and hepatocellular carcinoma) and hepatic histology – among others. The search extended up to November 2011, spanned five electronic databases and identified – among a total of seventy-eight – the same four RCTs with pioglitazone in the intervention arm that Shyangdan 2011 SR and Mahady 2010 SR had captured (Table 1). Contrary to Shyangdan 2011 SR, three of them were deemed low risk of bias (Belfort 2006, Aithal 2008, Sanyal 2010) across the eight domains of the Cochrane Risk of Bias tool (i.e., they scored 7 or 8). Several meta-analyses were performed over various histologic features combining data from RCTs of both pioglitazone and rosiglitazone. Regarding Vitamin E, the search identified five potentially relevant RCTs published between 2003 and 2011. Two of them (Nobili 2008, Lavine 2011) were conducted on the pediatric population, with the rest (Harisson 2003, Dufour 2006, Sanyal 2010) concerning adults with biopsy-proven NASH and reporting results in terms of liver histology (Table 2). All three were
deemed low risk of bias across the domains of the Cochrane Risk of Bias Tool (i.e., all scored 8). Meta-
analyses were performed over various histologic outcomes combining data from RCTs of both vitamin E and other antioxidants, as well as both adult and pediatric populations.

In the Boettcher 2012 SR (51), the participants were patients with NASH; interventions comprised pioglitazone or rosiglitazone versus placebo; outcomes of interest included histological response. The search extended up to September 2010, spanned three electronic databases, and identified – among a total of four – three potentially relevant RCTs with pioglitazone in the intervention arm. They matched precisely the ones captured previously by the preceding SRs (Belfort 2006, Aithal 2008, Sanyal 2010). They were all deemed high quality (scoring three) across the three domains of the Jadad scale (randomization, generation of random sequence, double-blinding) (Table 1). Meta-analysis over histologic outcomes was performed, combining data from the three pioglitazone RCTs and the one rosiglitazone study. A separate meta-analysis with the three pioglitazone studies was also performed.

In the Singh 2015 SR (58), the participants had biopsy-proven NASH; the interventions comprised vitamin E, TZDs, pentoxifylline, or Obeticholic Acid or a combination of these for at least one year; the comparator was another active agent or placebo; the outcomes were all histological: improvement in fibrosis, ballooning degeneration, steatosis, or lobular inflammation. The search extended up to November 2014, spanned six electronic databases and identified – among a total of nine – three potentially relevant RCTs with pioglitazone in the intervention arm. They matched precisely the ones captured by the previous systematic reviews (Belfort 2006, Aithal 2008, Sanyal 2010). All three were judged as high-quality across the seven domains of the Cochrane Risk of Bias tool, achieving maximum scores (Table 1). Direct meta-analyses over histological outcomes were performed pooling data from four RCTs (i.e., the three pioglitazone studies and the one rosiglitazone study). Regarding vitamin E, the search identified three potentially relevant RCTs (Harisson 2003, Sanyal 2010, Lavine 2011) – which had previously been captured by the Musso 2012 SR. They were all deemed high quality across the seven domains of the Cochrane Risk of Bias tool (all achieving maximum scores), and they were all combined into a direct meta-analysis over various histological outcomes, although one of them (Lavine 2011) had a pediatric population (Table 2).

In the Sato 2015 SR (50), the participants were patients with NAFLD or NASH; the intervention was vitamin E alone or combined with other treatment; the comparator was not specified; the outcomes included biochemical and histological responses. The search extended up to March 2014, spanned four databases, and identified five relevant RCTs, four of which (Varjo 2004, Dufour 2006, Sanyal 2010, Lavine 2011) had already been captured by previous SRs (Table 2). The fifth one (Wang 2008) examined non-histological outcomes. Meta-analyses were performed over inflammation, ballooning, and fibrosis pooling data from three RCTs (Dufour 2006, Sanyal 2010, Lavine 2011). Quality assessment was not carried out.
In the Sawangjit 2016 SR (57), the participants included patients of any age with biopsy-proven NAFLD; the interventions consisted of any type of pharmacological or not treatments, single or combined, tested against the placebo or other active comparator; the outcomes comprised all-cause mortality, cirrhosis, histological improvement of fibrosis, histological improvement of hepatic inflammation and adverse events. The search extended up to November 2015, spanned eight electronic databases and identified – among a total of forty-four – three potentially relevant RCTs with pioglitazone in the intervention arm (Belfort 2006, Aithal 2008, Sanyal 2010), all matching exactly the ones captured previously by the preceding SRs (table 1). They were all deemed low-risk of bias across the seven domains of the Cochrane Risk of Bias tool (scoring six or seven). Direct meta-analyses over various histological outcomes were performed pooling data from four RCTs, i.e., the three pioglitazone studies and one rosiglitazone study. Regarding Vitamin E, the search identified seven potentially relevant RCTs (Harisson 2003, Sanyal 2004, Dufour 2006, Nobili 2008, Sanyal 2010, Lavine 2011, Han 2014) with vitamin E in the intervention arm (table 2). Except for one (Han 2014), all scored six or seven across the seven domains of the Cochrane Risk of Bias tool. Two of them (Sanyal 2010, Lavine 2011) contributed data to direct meta-analyses over various histological outcomes, even though one (Lavine 2011) included pediatric participants.

In the He 2016 SR (49), the participants were adults with biopsy-proven NASH; the intervention comprised a TZD as drug monotherapy tested against the placebo or other comparator; the outcomes included histological response (grade of steatosis, lobular inflammation score) as well as various biochemical markers. The search extended up to 2015 (no further details), spanned six electronic databases and identified – among a total of five – three potentially relevant RCTs with pioglitazone in the intervention arm (Belfort 2006, Aithal 2008, Sanyal 2010), all matching the ones captured previously by the preceding SRs (table 1). Their quality was judged against three domains (randomization, allocation concealment, double-blinding), resulting in a score of two for all. The meta-analyses across the histological outcomes of hepatic fibrosis and lobular inflammation pooled data from five RCTs, i.e., the three pioglitazone and the two rosiglitazone studies. Two subgroup analyses were also performed: one with the two TZD studies that also applied LSI as co-interventions, combining data from one pioglitazone (Belfort 2006) and one rosiglitazone (Idilman 2008) study; another one with the rest three studies (Aithal 2008, Ratziu 2008, Sanyal 2010) with TZD alone. However, the inclusion of Aithal 2008 RCT into the second group (where the TZD was not combined with LSI) was questionable, as its report clearly stated that standard diet and exercise served as co-interventions (59).

In the Lombardi 2017 SR (60), the participants were patients with NAFLD irrespective of the method of diagnosis or diabetic status; the interventions comprised any pharmacologic agent either alone or in combination tested against the placebo, no intervention or other comparator; the outcomes included mortality at maximal follow-up, mortality at one year, mortality at one to five years, adverse
events up to three months after treatment cessation (classified as serious and non-serious), health-related quality of life, liver transplantation at maximal follow-up, decompensated liver disease at maximal follow-up, cirrhosis at maximal follow-up and resolution of fatty liver disease at maximal follow-up. The histological outcomes of fibrosis and NAFLD activity score were also included as unvalidated, potential surrogate markers. The search extended up to August 2016, spanned five electronic databases and identified – among a total of forty-one – six potentially relevant RCTs with pioglitazone in the treatment arm. Four of them (Sanyal 2004, Belfort 2006, Aithal 2008, Sanyal 2010) matched exactly the ones captured previously by the preceding SRs; the other two compared pioglitazone and LSI versus LSI (Yan 2015) or pioglitazone and “conventional treatment” versus “conventional treatment” (Jin 2010) over non-histological outcomes, such as liver biochemistry (Jin 2010) or hepatic fat content determined by imaging (Yan 2015) (Table 1). Quality was assessed across eight domains (random sequence generation, allocation concealment, double-blinding, incomplete outcome data, selective reporting, for-profit bias, other bias). Except for one RCT (Aithal 2008), which scored seven, all the other scored poorly. Regarding the vitamin E, the search identified five potentially relevant studies with vitamin E in the intervention arm. Three of them (Harisson 2003, Dufour 2006, Sanyal 2010) had already been captured by previous SRs; the other two (Magosso 2013, Basu 2014) compared vitamin E against the placebo over non-histological outcomes, such as hepatic fat content by ultrasonography (Magosso 2013) or fibrosis markers (Basu 2014) (table 2). Regarding quality assessment, all scored poorly across the eight domains, apart from one (Magosso 2013), which scored seven. Although the researchers intended to perform a network meta-analysis and to generate rankings of the available pharmacological interventions according to their safety and efficacy, they didn’t do so because it was judged that they couldn’t assess whether the potential effect modifiers were similar across different comparisons.

In the Said 2017 SR (48), the participants were adults with biopsy-proven NASH; the interventions included metformin, TZDs, and vitamin E compared against the placebo; the outcomes included biochemical and histological changes. The search extended up to December 2014, spanned four electronic databases and identified – among a total of nine – three potentially relevant RCTs with pioglitazone in the intervention arm (Belfort 2006, Aithal 2008, Sanyal 2010), all matching the ones previously captured by the preceding SRs. All studies were deemed high quality against the Cochrane Risk of Bias Tool and the Jadad 3-point scale without reporting individual scores (table 2). The meta-analyses that were performed over various histological outcomes pooled data from the three pioglitazone studies and the one rosiglitazone study. Regarding vitamin E, the search identified three potentially relevant RCTs with vitamin E in the intervention arm (Harisson 2003, Dufour 2006, Sanyal 2010) (table 2). They all had been previously captured by preceding SRs, were deemed high quality, and were combined into meta-analyses over various histological outcomes.
In the Musso 2017 SR (61), the participants were diagnosed with NAFLD or NASH by imaging or histological means; the interventions comprised TZDs; the comparators were not specified; the outcomes were defined as improvement in advanced fibrosis and improvement in fibrosis of any stage. The search extended up to August 2016, spanned seven electronic databases and identified – among a total of eight – five potentially relevant RCTs with pioglitazone in the intervention arm. Four of them (Sanyal 2004, Belfort 2006, Aithal 2008, Sanyal 2010) had previously been captured by the preceding SRs; the fifth one (Cusi 2016) was the most recent RCT ever to be included into a SR (Table 1). The quality assessment was based on the eight domains of the Cochrane Risk of Bias tool and resulted in scores of eight and seven for all. The meta-analyses over the histological outcomes of advanced fibrosis and fibrosis pooled data from all five pioglitazone studies.

In the Amanullah 2019 SR (46), the participants were patients with NAFLD; the intervention was vitamin E; the comparator was not defined; the outcomes included biochemical and histological responses. The search spanned 12 databases, extended up to July 2018, and identified nine relevant RCTs, six of which (Harisson 2003, Varjo 2004, Nobili 2006, Nobili 2008, Sanyal 2010, Lavine 2011) had already been identified by previous SRs. The other three (Foster 2011, Hoofnagle 2013, Aller 2015) examined combinations of vitamin E with other agents versus placebo and over non-histological outcomes, apart from one (Hoofnagle 2013) that also reported on hepatic inflammation and ballooning (table 2). Two meta-analyses over ballooning and fibrosis pooled data from two studies (Nobili 2008, Lavine 2011) with pediatric participants. Quality appraisal was based on the Jadad score (Randomisation mentioned, concealment of randomization, blinding, Appropriate blinding method, report of withdrawals) and yielded variable scores (Table 2).

**Characteristics of excluded studies**

Eight full-text studies were excluded for the following reasons:

- Included primary studies not being RCTs (n=1): Chavez – Tapia 2006 SR (62) included three non-randomized single-arm studies.
- Methodology not adhering to systematic review standards (n=2): The studies of Younossi 2014 (63) and Rinella 2015 (64) identified themselves as Systematic Reviews. However, the applied methodology was far from it for the following reasons: eligibility criteria weren’t pre-specified, and neither literature search nor study selection or data extraction was carried out by independent researchers.
- Outcomes examined not being relevant (n=3): the study of Tang 2016 (65) restricted itself in reporting the effects of pioglitazone on hepatic fat content and various biochemical parameters. The study of Ji 2016 (66) reported only effects on liver biochemistry. The study of Sridharan 2018 (67) reported effects on liver biochemistry and hepatic fat content.
• The publication being unobtainable (n=2): the study of Zeng 2012 (68) was available only in Chinese. The study of Angelico 2007 (69) has been retracted as outdated.

**Quality appraisal**

Only 3 out of the 14 SRs were rated as high confidence using the AMSTAR 2 tool (Lombardi 2017, Shyangdan 2011, Lirussi 2007); one was rated as low (Socha 2009), and ten were rated as critically low confidence (Amanullah 2019, Musso 2017, Said 2017, He 2016, Sawangjit 2016, Sato 2015, Singh 2015, Boettcher 2012, Musso 2012, Mahady 2011). The detailed appraisal is available in Appendix D; the graphical representation of the results is depicted in figure 2.

![AMSTAR2 Collective Results](image)

**Figure 2: Results of quality appraisal per AMSTAR2**

The AMSTAR 2 appraisal mandates degrading the level of confidence to low if one critical flaw exists and to critically low if two or more critical flaws exist in the systematic review (42). The critical domains of existence of protocol (item 2), appropriateness of meta-analysis (item 11), and provision of a detailed list of excluded studies (item 7) were failed by the vast majority, resulting in downgrading them to the levels of low and critically low quality. Justification of study type selection (item 3) and report of funding of primary studies were also overlooked by the vast majority. Roughly half had issues with literature search (item 4), description of primary studies (item 8), assessment of the risk of bias in primary studies (item 9), exploration of heterogeneity (item 14), and investigation of publication bias (15). Most reviews performed well in setting the PICO elements clearly (item 1), carrying out the study selection and data extraction in duplicate (items 5 and 6), and disclosing their funding sources (item 16). The domains regarding the risk of bias of primary studies (items 12 and 13) were performed adequately by the majority mainly because the respective studies were deemed low risk.
Patient characteristics
The level of agreement between the PICO elements of the overview protocol and the included SRs is depicted in figure 3. Some SRs had set NAFLD and some NASH as the diagnosis of interest. Nonetheless, all reviews retrieved primary studies that recruited patients with biopsy-proven NASH, following the requirements of our protocol.

Interventions and comparators
Almost none of the included reviews focused exclusively on the combination of pioglitazone and LSI versus LSI alone. The three primary studies (Belfort 2006, Aithal 2008, Cusi 2016) that tackled exactly this comparison have never appeared as a stand-alone triad in a meta-analysis. Instead, the SRs that performed meta-analyses combined data from primary studies of both pioglitazone and rosiglitazone, apparently treating them equally as TZD drugs. LSI as a co-intervention was overlooked in all SRs, resulting in meta-analyses that equalized pioglitazone monotherapy with the combination of pioglitazone and LSI or pioglitazone and vitamin E. The only exception was the He 2016 SR, where two subgroup analyses were performed distinguishing between TZD without LSI and the combination of TZD and LSI. The Shyangdan 2011 SR differentiated itself from the rest ones by acknowledging that the heterogeneity across the (same) primary studies was too high to combine them into a single meta-analysis.

None of the included SRs focused on the combination of vitamin E and LSI versus LSI alone. Apart from two (Lirussi 2007, Socha 2009), all the other performed meta-analyses, combining data from primary studies where vitamin E was combined with other agents into antioxidant intervention. By looking at the sixteen primary studies, only one (Varjo 2004) investigated the comparison of vitamin E and LSI versus LSI, albeit in the pediatric population and over non-histological outcomes (liver ultrasonography and biochemistry).

Outcomes
The primary outcome of all-cause mortality was assessed by three systematic reviews (Lirussi 2007, Sawangjit 2016, Lombardi 2017) (Figure 3). In Lombardi 2017 study, mortality at maximal follow-up was reported for the comparison of TZD versus no intervention, with data having been derived from a single primary study of 74 patients (Aithal 2008) where the combination of pioglitazone and LSI was tested against LSI alone for 12 months. No events were reported within a follow-up period of 12 months, and, therefore, the odds ratio was not assessed. In the same study, mortality at maximal follow-up for the comparison of antioxidants versus no intervention derived data from a single study with 87 patients (Magosso 2013) where vitamin E was tested against the placebo for 12 months. No events were reported within a follow-up period of 12 months, and, therefore, the odds ratio was not estimated. In Sawangjit 2016 study, the outcome of mortality was assessed by deriving data from two primary
Figure 3: Comparison of PICO elements between the overview protocol and the captured systematic reviews to reveal the level of agreement.
studies, where vitamin E was tested against the placebo for 24 months in adults with NASH (Sanyal 2010, n=167 participants) and in children with NAFLD (Lavine 2011, n = 116). Two events were reported among 142 Vitamin E users within a follow-up period of over a year without assessing odds ratios. In Lirussi 2007 study, no fatalities were reported in the included primary studies, where combinations of vitamin E and variable co-interventions were tested against active comparators.

The primary outcome of liver morbidity was investigated by four systematic reviews (Lombardi 2017, Sawangjit 2016, Musso 2012, Shyangdan 2011) (Figure 3). In Lombardi 2017 study, two relevant outcomes were assessed: the decompensated liver disease was evaluated by retrieving data from two primary studies; one (Aithal 2008 RCT) which assessed pioglitazone and LSI versus LSI for 12 months and one (Magosso 2013 RCT) which investigated vitamin E versus placebo for 12 months. No events were reported in either study, and odds ratios were not assessed. Risk of cirrhosis was evaluated by combining data from two primary studies (Belfort 2006 RCT, Aithal 2008 RCT) which both investigated the combination of pioglitazone and LSI versus LSI alone in 126 adult participants with NASH; the calculated odds ratio [95% Confidence Interval] was 5.99 [0.71, 50.28] in favor of the control. However, the researchers, in this case, equalized fibrosis (the outcome in those two primary studies) to cirrhosis (the outcome in the SR); this was not accurate because cirrhosis requires the presence of well-defined symptoms and signs of chronic liver disease, contrary to fibrosis. Regarding cirrhosis in vitamin E users, data came from one study (Magosso 2013 RCT), where no events were reported, and therefore odds ratios could not be calculated. Although in the rest systematic reviews (Sawangjit 2016, Musso 2012, Shyangdan 2011) liver-related morbidity (in the form of cirrhosis, liver failure, hepatocellular carcinoma, variceal bleeding) had been declared as a primary outcome in the methods section, no mention about it existed in the results section.

Reversal of hepatic fibrosis and hepatic inflammation had been declared as outcomes in all systematic reviews. In ten of them (Mahady 2010, Musso 2012, Boettcher 2012, Singh 2015, Sato 2015, Sawangjit 2016, He 2016, Said 2017, Musso 2017, Amanullah 2019) meta-analyses were performed which combined data from a variety of heterogeneous primary studies, treating pioglitazone and rosiglitazone as TZDs or vitamin E and other antioxidants as all being antioxidants. In four systematic reviews (Lirussi 2007, Socha 2009, Shyangdan 2011, Lombardi 2017), such meta-analyses were deemed inappropriate due to marked heterogeneity in interventions, comparators, and measures of outcomes. By looking at the primary studies (Belfort 2006, Aithal 2008 and Cusi 2016) that investigated exactly the comparison between the combination of pioglitazone and LSI versus LSI, the outcome of reversal of fibrosis was statistically significant (at p = 0.05) only in one (Aithal 2008) (table 3). The reversal of NASH was not reported unanimously but in the form of various histological features. Statistically significant improvement was noticed in almost all cases (table 3). Regarding vitamin E, there was no relevant data available, since the only RCT that tested the combination of vitamin E and LSI against LSI (Varjo 2004) did so in pediatric participants and over non-histological outcomes.
Table 3 Histological outcomes in the primary studies that investigated the combination of pioglitazone and lifestyle intervention versus lifestyle intervention. T2D: Type 2 Diabetes. D: Duration of treatment. FU: Duration of follow-up. Pio: pioglitazone. LSI: Lifestyle intervention. Pcb: placebo.

<table>
<thead>
<tr>
<th>Histological Features</th>
<th>Belfort 2006 (participants without T2D n = 52; Kleiner histological score; D = 6 months; FU 6 months)</th>
<th>Aithal 2008 (participants without T2D n = 74; Brunt histological score; D = 12 months; FU 12 months)</th>
<th>Cusi 2016 (participants with prediabetes or T2D n = 101; Kleiner histological criteria; D = 18 months; FU 18 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibrosis</td>
<td>Improvement from baseline: Pio+LSI 46%; Pcb+LSI 33% [p = 0.08] Reduction in score of ≥2 from baseline: Pio+LSI 5/12 (42%); Pcb+LSI 1/6 (17%) [p = 0.31]</td>
<td>Decrease from baseline: Pio+LSI 9/31 (29%); Pcb+LSI 6/30 (20%) [p = 0.05]</td>
<td>Improvement from baseline ≥1-point: Pio+LSI 20/50 (39%); Pcb+LSI 13/51 (25%) [p = 0.13]</td>
</tr>
<tr>
<td>Lobular inflammation</td>
<td>Improvement from baseline: Pio+LSI 65%; Pcb+LSI 29% [p = 0.008]</td>
<td>Decrease from baseline: Pio+LSI 14/31 (45%); Pcb+LSI 8/30 (27%) [p = 0.25]</td>
<td>Improvement from baseline ≥1-point: Pio+LSI 25/50 (51%); Pcb+LSI 12/51 (24%) [p = 0.004]</td>
</tr>
<tr>
<td>Ballooning necrosis</td>
<td>Improvement from baseline: Pio+LSI 54%; Pcb+LSI 24% [p = 0.02]</td>
<td>NA</td>
<td>Improvement from baseline ≥1-point: Pio+LSI 25/50 (51%); Pcb+LSI 12/51 (24%) [p = 0.004]</td>
</tr>
<tr>
<td>Combined necroinflammation</td>
<td>Improvement from baseline: Pio+LSI 85%; Pcb+LSI 38% [p = 0.001]</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Portal inflammation</td>
<td>NA</td>
<td>Decrease from baseline: Pio+LSI 8/31 (26%); Pcb+LSI 7/30 (23%) [p = 0.67%]</td>
<td>NA</td>
</tr>
<tr>
<td>Hepatocellular injury</td>
<td>NA</td>
<td>Decrease from baseline: Pio+LSI 10/31 (32%); Pcb+LSI 3/30 (10%) [p = 0.005]</td>
<td>NA</td>
</tr>
<tr>
<td>Mallory bodies</td>
<td>NA</td>
<td>Decrease from baseline: Pio+LSI 8/31 (26%); Pcb+LSI 1/30 (3%) [p = 0.004]</td>
<td>NA</td>
</tr>
<tr>
<td>Inflammation</td>
<td>NA</td>
<td>NA</td>
<td>Improvement from baseline ≥1-point: Pio+LSI 25/50 (49%); Pcb+LSI 11/51 (22%) [p = 0.004]</td>
</tr>
<tr>
<td>NASH resolution</td>
<td>NA</td>
<td>NA</td>
<td>Pio+LSI: 26/50 (51%); Pcb+LSI 10/51 (19%) [p &lt; 0.001]</td>
</tr>
<tr>
<td>≥2-point reduction in NAS</td>
<td>NA</td>
<td>NA</td>
<td>Pio+LSI: 29/50 (58%); Pcb+LSI 9/51 (17%) [p &lt; 0.001]</td>
</tr>
</tbody>
</table>
Assessment of adverse events had been declared as an outcome in three SRs (Lombardi 2017, Sawangjit 2016, Shyangdan 2011) (Figure 3). None of them reported any events of congestive heart failure, bone fractures, bladder cancer (pioglitazone related), or prostate cancer (vitamin E related).

An overview of the results is presented in tables 4 and 5.

Table 4 Outcomes of pioglitazone and lifestyle intervention versus lifestyle intervention. LSI: Lifestyle intervention

<table>
<thead>
<tr>
<th>All-cause mortality</th>
<th>Liver related morbidity</th>
<th>Histological outcomes</th>
<th>Adverse events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Decompensated liver disease</td>
<td>Cirrhosis</td>
<td>NASH improvement</td>
</tr>
<tr>
<td>No events</td>
<td>No events</td>
<td>No events</td>
<td>All three relevant RCTs reported statistically significant improvement in various features of hepatic inflammation</td>
</tr>
</tbody>
</table>

Table 5: Outcomes of vitamin E and lifestyle intervention versus lifestyle intervention. LSI: Lifestyle intervention

<table>
<thead>
<tr>
<th>All-cause mortality</th>
<th>Liver related morbidity</th>
<th>Histological outcomes</th>
<th>Adverse events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Decompensated liver disease</td>
<td>Cirrhosis</td>
<td>NASH improvement</td>
</tr>
<tr>
<td>2 events among 343 vitamin E users</td>
<td>No events</td>
<td>No events</td>
<td>No available data</td>
</tr>
</tbody>
</table>

DISCUSSION

In this overview of systematic reviews, we primarily tried to figure out if pioglitazone or vitamin E can affect the mortality and the liver-related morbidity favorably when they are added to LSI in patients with NASH and at what cost in terms of serious, well-recognized adverse events, such as congestive heart failure, bone fractures, bladder cancer, and prostate cancer. We also investigated their effects on specific histological outcomes such as improvement of hepatic fibrosis and NASH reversal.

We did not find any evidence that pioglitazone reduces mortality or liver-related morbidity in NASH patients since no primary study has ever had sufficient duration to address this question. The follow-up period of the three relevant RCTs ranged from 6 to 18 months, a span too short to allow for mortality assessment in the context of such a slowly-progressive condition like NAFLD (17). Expectedly, we did not find any evidence of the serious pioglitazone-related adverse events within the short duration of these trials.

The best available data about fibrosis in NASH patients treated with pioglitazone on top of LSI did not come from any meta-analysis but three primary studies (Belfort 2006, Aithal 2008, Cusi 2016); however, they were never combined as a stand-alone triad in a meta-analysis, although the Musso 2017 SR included them all. Therefore, we presented the results as reported in the primary studies, following the example of the Shyangdan 2011 SR. Improvement of fibrosis was statistically significant
(at \( p = 0.05 \)) in only one RCT (Aithal 2008) with 74 participants; the other two (Belfort 2006, Cusi 2016), which recruited a total of 153 participants, did not confirm that result (Table 3).

Regarding the improvement in NASH, we also resorted to the results of the primary studies, adopting the approach in the Shyangdan 2011 SR. Marked variability was noticed in the applied histological criteria and – consequently – in the way the results were reported. Literally, NASH resolution was reported only by one primary study (Cusi 2016) with 101 prediabetic or diabetic participants; it was also found to be statistically significant at \( p < 0.001 \). All other results were reported as specific histological features of NASH, varying in such a way that they could hardly be combined into a single figure. The typical pattern was statistically significant (at \( p < 0.05 \)) improvement of almost all histological features that were assessed, as presented in Table 3.

We did not find any evidence that vitamin E when added to LSI in patients with NASH, improves either mortality or liver-related morbidity. The same was true for the outcomes of reversal of fibrosis and improvement in NASH. The reason was that none of the nine included SRs has ever captured any RCT, which had studied the combination of vitamin E and LSI against LSI in adults across the clinical or histological outcomes of interest.

Does a patient with NASH stand a chance to live longer with pioglitazone or vitamin E? Does a patient with NASH have fewer chances to progress to terminal liver disease with pioglitazone or vitamin E? How is this potential clinical benefit counterbalanced by serious and well-documented adverse events associated with long term use of these agents, such as congestive heart failure, bone fractures, bladder cancer or prostate cancer? Our overview of the best available evidence suggests the answer is that we don’t know. We could have never known because there is a lack of primary studies designed to answer these questions. Instead, the available primary studies have adopted histological outcomes in the short run as surrogate markers for the disease progression, but how reliable are they?

Liver fibrosis, in contrast to other histological features of NASH, has been associated with long term mortality, liver-related morbidity, and liver transplantation in patients with NAFLD (34,35). For this reason, it has been adopted by regulatory agencies to accelerate drug-approval processes for NAFLD treatment (70). Does this positive association from retrospective, observational studies suffice to qualify fibrosis as a surrogate marker for NASH? Definitely not. This is just the first step in the validation process of a putative surrogate outcome, but correlation alone is not sufficient to qualify an outcome as a surrogate (71). The second step requires the undertaking of randomized clinical trials to demonstrate, by appropriate statistical methods, that the intervention affects both the intermediate (surrogate) and the clinical outcome significantly and that the intervention’s effect on the intermediate outcome can predict the intervention’s effect on the clinical outcome (72–76). Such a task has never been undertaken in NAFLD. Therefore hepatic fibrosis is far from being a reliable surrogate outcome by sound scientific methodology, no matter how “reasonable” it appears to be.
Needless to remind the hasty approval of antiarrhythmics for the suppression of premature ventricular beats (PVCs) after myocardial infarction, on the basis that PVCs correlated with poor clinical outcomes. The CAST trial that followed (77) demonstrated that mortality was significantly higher in the patients receiving antiarrhythmics compared to placebo, thus rendering the suppression of PVCs (an otherwise “reasonable” strategy) an invalid surrogate outcome for the prediction of death.

Several issues have emerged in this overview process that should not be overlooked. First of all, the considerable overlapping among the included SRs raises the question of redundant research, a concern that has already been underlined in the literature (78). Indicatively, we have identified 11 SRs that cited ten primary studies about TZDs 60 times! SRs are often depicted higher than primary studies in the pyramid of evidence because they are meant to synthesize data into a bigger and more precise picture. However, the mass production of overlapping SRs that ignore each other’s presence results in obscuring rather than clarifying a scientific matter.

Another issue that emerged regards the noticeable inconsistencies across the included SRs in the quality appraisal of the primary studies. As displayed in tables 1 and 2, a variety of different criteria have been used, producing mixed results. A primary study deemed high quality in one SR may well be judged as low quality in another. The quality appraisal tools and protocols have been developed to diminish subjectivity, apparently with poor results.

The critically low quality of the majority (ten out of fourteen) of the included SRs is striking. The commonest critical (per AMSTAR2) flaws were the absence of a well-documented protocol, the omission of a detailed list of excluded studies, and the lack of adequate justification over combining results from primary studies into a meta-analysis. Regarding the latter, it is worth mentioning that in three of the SRs, the meta-analysis was deemed inappropriate while in eleven appropriate, despite that all had included almost the same set of primary studies. Who was right then? To investigate this discrepancy, we analyzed the PICO elements of the relevant primary studies, although such a task had not been within the scope of our overview. Even by limiting our focus on the three primary studies (Belfort 2006, Aithal 2008, Cusi 2016) about the combination of pioglitazone and LSI, one can notice discrepancies across the populations (nondiabetic, prediabetic and diabetic), the interventions (different pioglitazone doses, different duration of treatments, different advice over lifestyle interventions) and the outcome measures (different histological criteria). Therefore combining the results into a meta-analysis might not be appropriate, an issue acknowledged only in one SR (Shyangdan 2011).

None of the SRs that went on to perform meta-analyses provided persuasive reasons for combining data from studies that differed in their PICO elements. For example, the comparison of pioglitazone and LSI versus LSI is not the same as the comparison of pioglitazone versus placebo, because LSI is an effective treatment for NAFLD (6,7); this detail has been overlooked by all included meta-analyses.
Combining pioglitazone and rosiglitazone needs caution, as they are two different molecules. Adding up adult and pediatric populations under the same meta-analysis is risky, especially without adequate justification. Many things must be explained before vitamin E can be combined into a meta-analysis together with UDCA or vitamin C under the tag of “antioxidants.” Heterogeneity of primary studies must be discussed thoroughly on clinical grounds and precede any meta-analysis, a task omitted universally in the included SRs, resulting (together with the absence of a protocol) into degrading them to the lowest scale of confidence (per AMSTAR2).

**Strengths and limitations**

The present overview captured SRs that dealt with the role of pioglitazone and vitamin E in NASH, juxtaposed their methodological features revealing striking discrepancies, and appraised them according to the AMSTAR2 tool. Our approach was primarily clinical, balancing the efficacy and safety of those agents per PRIO HARMS standards (33). However, important deviations from the protocol occurred, as some critical tasks (peer review of the search strategy, study selection, data extraction, quality appraisal) were not performed in duplicate; instead, they were carried out only by the MSc student for the dissertation. Nonetheless, it is a flaw that can’t be overlooked, because it increases the subjectivity of the analysis.

**Conclusion**

We attempted to find out if pioglitazone and vitamin E in patients with NASH make any difference on clinical grounds. We unveiled a landscape of a plethora of overlapping, low-quality by majority SRs that included and meta-analyzed data from heterogeneous primary studies and over histological outcomes that served as unvalidated surrogate markers in the short term. The result is that after almost two decades of research, we still don’t know if pioglitazone or vitamin E can help patients with NASH live longer or free of liver-related complications.

**Implications for future research**

A large number of novel drugs targeting different molecular mechanisms are currently being tested in clinical trials (3); almost all focus on histological endpoints. Therefore, the research must prioritize the validation of reliable surrogate outcomes for this slowly progressive condition; otherwise, we risk wasting more time, efforts, and valuable resources to no avail.

The production of SRs has skyrocketed while the quality has plummeted, compromising their credibility. Peer reviewers should identify redundant SRs more rigorously, and editors should discourage their publication. Researchers committed to SRs and MAs should minimum guide their work by a registered protocol, discuss heterogeneity on clinical grounds and explain the reasons for or against combining data from different studies into a single measure of effect meticulously.
References


12. Leamy AK, Egnatchik RA, Young JD. Molecular mechanisms and the role of saturated
fatty acids in the progression of non-alcoholic fatty liver disease. Prog Lipid Res. 2013 Jan;52(1):165–74.


24. Lalloyer F, Staels B. Fibrates, glitazones, and peroxisome proliferator-activated


47. Musso G, Cassader M, Paschetta E, Gambino R. Thiazolidinediones and Advanced Liver


56. Lirussi F, Azzalini L, Orando S, Orlando R, Angelico F. Antioxidant supplements for non-alcoholic fatty liver disease and/or steatohepatitis. Cochrane Database Syst Rev. 2007;(1).


64. Rinella ME. Nonalcoholic fatty liver disease: a systematic review. JAMA. 2015 Jun;313(22):2263–73.


### Appendix A. PRIO harms checklist

<table>
<thead>
<tr>
<th>Section/Topic</th>
<th>[Sub-] item #</th>
<th>Checklist item</th>
<th>Reported on page #</th>
</tr>
</thead>
</table>
| **TITLE**     | 1. Title      | 1a Specify the study design with terms such as “overview of (systematic) reviews,” “umbrella review,” “(systematic) review of systematic reviews,” or “(systematic) meta-review” in the title of the OoSRs.  
1b Mention “safety” or harms related terms, or the adverse event(s) of interest in the title of the OoSRs.                                                                                                                      | 2                 |
| **ABSTRACT**  | 2. Structured – like summary | 2a Provide a structured-like abstract, as applicable: background, objective, data sources, selection criteria, data extraction, review appraisal, data synthesis methods, results, limitations, conclusions  
2b Report the main findings of analysis of harms undertaken in the OoSRs or/and in the included SRs.                                                                                                                                                                                             | 4                 |
| **INTRODUCTION** | 3. Rationale | 3a Specify the rationale and the scope (wide or narrow agendas) for the overview in the context of an existing body of knowledge on the topic.  
3b Provide a balanced presentation of potential benefits and harms of the intervention(s).  
3c Define which events are considered harms according to previous literature and provide a clear rationale for the specific harms included in the OoSRs.                                                                                      | 5-6, 7-9          |
|               | 4. Objectives (PICOS) | 4 Provide an explicit statement of research question(s) that specifies PICOS:  
Participants: Yes  
Interventions: Yes  
Comparators: Yes  
Outcomes: Yes  
Study design: Yes  

**METHODS**

| 5. Protocol and registration | 5a Indicate clearly if a protocol exists or not.  
5b If registered, provide the name of the registry (such as a valid Web address, PROSPERO). | 8 | 8 |
| 6. Eligibility criteria and outcomes of interest | 6a Specify inclusion and exclusion criteria for study design, participants, interventions, and comparators in detail.  
6b List (and define whenever it is necessary) the outcomes for which data were recorded, ideally include prioritization of main and additional outcomes.  
6c Include adverse events as (primary or secondary) outcome of interest. Define them and grade their severity (such as mild, moderate, severe, fatal); severity could also be described in the appendix, if appropriate.  
6d Specify report characteristics (such as language restrictions, publication status, and years considered) used as criteria for eligibility for the OoSRs (see also item 7). | 8-9 | 9, 9, 9 |
<p>| 7a Search at least two electronic bases. | | 9 | 9 |</p>
<table>
<thead>
<tr>
<th>Section</th>
<th>Subsection</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Information sources</td>
<td>7b</td>
<td>Search supplementary sources (e.g. hand-searching, reference lists, related reviews and guidelines, protocol registries, conference abstracts, and other gray literature).</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>7c</td>
<td>Report the date last searched and/or dates of coverage for each database.</td>
<td>9</td>
</tr>
<tr>
<td>8. Search strategy</td>
<td>8a</td>
<td>Specify full electronic search strategy (algorithm) for at least one database including any limits used (e.g. language and date restrictions-see also subitems 6d and 7c) such that it could be repeated.</td>
<td>Appendix B</td>
</tr>
<tr>
<td></td>
<td>8b</td>
<td>Present any additional search process (e.g. algorithm or filter for adverse events, searches in pertinent websites) specifically to identify adverse events that have been investigated.</td>
<td>NA</td>
</tr>
<tr>
<td>9. Data management and selection process</td>
<td>9a</td>
<td>Describe the software that was used to manage records and data throughout the OoSRs.</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>9b</td>
<td>Define what is a SR and provide the process for selecting SRs and its relevant details (screening the title and abstract or full text by at least two reviewers, selection by multiple independent investigators and resolving disagreements by consensus).</td>
<td>9, 10</td>
</tr>
<tr>
<td></td>
<td>9c</td>
<td>Report any attempt to handle overlapping (include one review among multiple potential candidates by choosing for example the most updated SR, the most methodologically rigorous SR or the SR with larger number of primary studies).</td>
<td>12</td>
</tr>
<tr>
<td>10. Additional search for primary studies</td>
<td>10</td>
<td>Report additional search to identify eligible primary studies (e.g. searching in more databases or update the search) and its relevant details.</td>
<td>NA</td>
</tr>
<tr>
<td>11. Data collection process</td>
<td>11a</td>
<td>Describe the method of data extraction from included SRs (e.g. data collection form, extraction in duplicate and independently, resolving disagreements by consensus).</td>
<td>10 &amp; Appendix C</td>
</tr>
<tr>
<td></td>
<td>11b</td>
<td>Report any processes for obtaining, confirming, or updating data from investigators (e.g. contact with authors of included reviews, obtain data from primary studies of included reviews).</td>
<td>NA</td>
</tr>
<tr>
<td>12. Data items</td>
<td>12</td>
<td>List (and define whenever is necessary) the specific variables for which data were recorded (e.g. PICOS items, number of included studies and participants, dose, length of follow up, results, funding sources) and any data assumptions and simplifications made.</td>
<td>10</td>
</tr>
<tr>
<td>13. Assessment of methodological quality and quality of evidence</td>
<td>13a</td>
<td>State the evaluation of reporting or/and methodological quality (eg. using PRISMA or PRISMA-harms, AMSTAR or R-AMSTAR) of the included reviews.</td>
<td>10-11</td>
</tr>
<tr>
<td></td>
<td>13b</td>
<td>State the evaluation of quality for individual studies that were included in the SRs (inform whether tools such as Jadad or RoB of Cochrane were used by the included reviews) and for the additional primary studies.</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>13c</td>
<td>State the evaluation of quality of evidence (e.g. using GRADE approach).</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>13d</td>
<td>Describe the methods (e.g. piloted forms, independently, in duplicate) used for the quality assessment.</td>
<td>11 &amp; 30</td>
</tr>
<tr>
<td>14. Meta-bias(es)</td>
<td>14</td>
<td>Specify any planned assessment of meta-bias(es) (such as publication bias or selective reporting across studies, ROBIS tool).</td>
<td>NA</td>
</tr>
<tr>
<td>15. Data synthesis</td>
<td>15a</td>
<td>Specify clearly the method (narrative, meta-analysis or network meta-analysis) of handling or synthesizing data and their details (e.g. state the principal summary measures that were extracted or calculated, how</td>
<td>11-12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>heterogeneity was assessed, what statistical approaches were used if a quantitative synthesis has been conducted.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describe the software that was used to analyze the data if a quantitative synthesis has been conducted.</td>
<td>NA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Report if zero events are included in the studies and how they were handled in statistical analyses, if relevant.</td>
<td>NA.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describe methods of any pre-specified additional analyses (such as sensitivity or subgroup analyses, meta-regression).</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### RESULTS

#### 16. Review and primary study selection

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Provide the details of review selection (e.g. numbers of reviews screened, retrieved, and included and excluded in the overview) and the number of the additional eligible primary studies that were included, ideally with a flow diagram of the overview process.</td>
<td>12-13</td>
</tr>
<tr>
<td>Present a flow diagram that gives separately the number of studies focused on harms outcomes.</td>
<td>24 (Figure 3)</td>
</tr>
<tr>
<td>List the studies (full citation) that were excluded after reading the full text and provide reasons.</td>
<td>21</td>
</tr>
</tbody>
</table>

#### 17. Review and primary study characteristics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe characteristics of each included SR in tables (such as title or author, search date, PICOS, design and number of studies included, number and age range of participants, dose/frequency, follow up period [treatment duration], review limitations, results or conclusion) and of each additional primary study.</td>
<td>14 -15 (Tables 1 &amp; 2)</td>
</tr>
<tr>
<td>For each included SR report language and publication status restrictions that have been used.</td>
<td>NA</td>
</tr>
</tbody>
</table>

#### 18. Overlapping

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Present or/discuss about overlapping of studies within SRs (at least one of the following):</td>
<td>13</td>
</tr>
<tr>
<td>- Present measures of overlap (such as CCA).</td>
<td>13</td>
</tr>
<tr>
<td>- Provide citation matrix</td>
<td>14-15</td>
</tr>
<tr>
<td>- Give the number of index publications or/discuss about overlapping</td>
<td>14-15 &amp; 29</td>
</tr>
</tbody>
</table>

#### 19. Present assessment of methodological quality and quality of evidence

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Present results in text or/and tables of any quality assessment (see also subitems 13a-c):</td>
<td>22, Figure 2, Appendix D</td>
</tr>
<tr>
<td>- Reporting or/and methodological quality of the included SRs.</td>
<td>14-15 &amp; 16-21</td>
</tr>
<tr>
<td>- Inform for the quality of the individual studies that were included in the SRs (report results for sequence generation, allocation concealment, blinding, withdrawals, bias etc.) and for the additional included primary studies.</td>
<td></td>
</tr>
<tr>
<td>- Quality of evidence.</td>
<td>NA</td>
</tr>
</tbody>
</table>

#### 20. Present meta-bias(es)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Present results of any assessment of meta-bias(es) (such as publication bias or selective reporting across studies, ROBIS assessment).</td>
<td>NA</td>
</tr>
</tbody>
</table>

#### 21. Synthesis of results

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summarize and present the main findings of the overview for benefits and harms. If a quantitative synthesis has been conducted, present each summary measure with a confidence interval, prediction interval or a credible interval and measures of heterogeneity or inconsistency.</td>
<td>23-27 (Tables 3,4,5)</td>
</tr>
<tr>
<td>Give results of any additional analyses (such as sensitivity, subgroup analyses, or meta-regression).</td>
<td>NA.</td>
</tr>
<tr>
<td>Report results for adverse events separately for each intervention.</td>
<td>27, Tables 4,5</td>
</tr>
</tbody>
</table>

### DISCUSSION
<table>
<thead>
<tr>
<th>22. Summary of evidence</th>
<th>22</th>
<th>Provide a concise summary of the main findings with the strength and shortcomings of evidence for each main outcome.</th>
<th>27-28</th>
</tr>
</thead>
<tbody>
<tr>
<td>23. Limitations</td>
<td>23a</td>
<td>Discuss limitations of either the overview or included studies (or both) (e.g. different eligibility criteria, limitations of searching reviews, language restrictions, publication and selection bias).</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>23b</td>
<td>Report possible limitations of the included reviews related to harms (issues of missing data and information, definitions of harms, rare adverse effects).</td>
<td>28-29</td>
</tr>
<tr>
<td>24. Conclusions</td>
<td>24a</td>
<td>Provide a general interpretation of the results in coherence with the review findings and present implications for practice; consider the harms equally as carefully as the benefits and in the context of other evidence.</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>24b</td>
<td>Present implications for future research.</td>
<td>30</td>
</tr>
<tr>
<td><strong>AUTHORSHIP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Contributions of authors</td>
<td>25</td>
<td>Provide contributions of authors.</td>
<td>8</td>
</tr>
<tr>
<td>26. Dual (co-)authorship</td>
<td>26</td>
<td>Report about dual (co-)authorship in the limitation or declarations of interest section.</td>
<td>NA</td>
</tr>
<tr>
<td><strong>FUNDING</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Funding or other support</td>
<td>27a</td>
<td>Indicate sources of financial and other support for the OoSRS (direct funding) or for the authors (indirect funding) or report no funding.</td>
<td>NA.</td>
</tr>
<tr>
<td></td>
<td>27b</td>
<td>Provide name for the overview funder and/or sponsor, or for the authors’ supporters.</td>
<td>NA.</td>
</tr>
<tr>
<td></td>
<td>27c</td>
<td>Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in conducted the OoSRS</td>
<td>NA</td>
</tr>
</tbody>
</table>
Appendix B. Search strategy for MEDLINE via PubMed

Non-alcoholic fatty liver disease

1. Non-alcoholic fatty liver disease [MeSH, all subheadings and categories included]
2. NAFLD OR Non-alcoholic fatty liver disease OR non alcoholic fatty liver disease OR nonalcoholic fatty liver disease OR Hepatic steatosis OR liver steatosis OR fatty liver OR NASH OR non-alcoholic steatohepatitis OR non alcoholic steatohepatitis OR nonalcoholic steatohepatitis
3. #1 OR #2

Pioglitazone

4. Pioglitazone [MeSH, all subheadings and categories included]
5. Pioglitazone OR TZD* OR thiazolidinedione* OR glitazone* OR insulin sensitizer* OR PPAR gamma agonist* OR PPAR-gamma agonist* OR PPARgamma OR PPARγ agonist* OR PPAR-γ agonist OR Peroxisome proliferator – activated receptor gamma OR peroxisome proliferator activated receptor gamma
6. #4 OR #5

Vitamin E

7. Vitamin E [MeSH, all subheadings and categories included]
8. Vitamin E OR Alpha tocopherol* OR alpha-tocopherol* OR α tocopherol* OR a-tocopherol* OR αtocopherol* OR tocopherol* OR Antioxidant* OR anti-oxidant*
9. #7 OR #8

Pharmacological intervention

10. Drug therapy [MeSH, all subheadings and categories included]
11. Pharmacologic* intervention* OR pharmacologic* treatment* OR pharmacologic* agent*
12. #10 OR #11

Pioglitazone or Vitamin E or lifestyle intervention or pharmacological intervention

13. #6 OR #9 OR #12

NAFLD and pioglitazone or vitamin E or lifestyle intervention or pharmacological intervention

14. #3 AND #13
Appendix C. Data extraction list

- Citation
- Author/year
- The objective of the review
- Participants (characteristics / total number)
- Setting / context
- Description of intervention
- Phenomena of interest
- Number of databases searched
- The date range of database searching
- The date range of the included studies
- Number of studies included in the review
- Types of studies included in the review
- Country of origin of primary studies
- Tool for quality assessment
- Quality rating
- Method of analysis/synthesis
- Outcome report
- Significance / direction
- Heterogeneity
- Adverse outcomes
- Any author’s comments
- Funding source
- Conflicts of interest
## Appendix D. Quality appraisal per AMSTAR 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 2</td>
<td>No</td>
<td>Partial Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Partial Yes</td>
<td>No</td>
<td>Partial Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Item 3</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Item 4</td>
<td>No</td>
<td>Yes</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 5</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 6</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 7</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Item 8</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Partial Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 9</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Partial Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Partial Yes</td>
</tr>
<tr>
<td>Item 10</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Item 11</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Item 12</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>MA</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 13</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Item 14</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Item 15</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>MA</td>
</tr>
<tr>
<td>Item 16</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Failed Critical domains
- Six
- Two
- Five
- None
- Five
- Two
- Three
- Three
- None
- Two
- One
- None

### Failed Non critical domains
- Four
- Three
- Four
- None
- Six
- Three
- Seven
- Three
- Three
- None
- One
- One
- Two

### Rating of confidence
- Critically low
- Critically low
- Critically low
- High
- Critically low
- Critically low
- Critically low
- Critically low
- Critically low
- High
- Critically low
- Low
- High

Items in bold are considered critical. MA: Meta-analysis