



Comparative effectiveness of medical treatment vs. metabolic surgery for histologically proven non-alcoholic steatohepatitis and fibrosis: a matched network meta-analysis

Adrian T. Billeter¹, Beatrice Reiners¹, Svenja E. Seide², Pascal Probst^{1,3}, Eva Kalkum³, Christian Rupp⁴, Beat P. Müller-Stich¹

¹Department of General, Visceral and Transplant Surgery, University of Heidelberg, Heidelberg, Germany; ²Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany; ³The Study Center of the German Surgical Society (SDGC), Heidelberg, Germany;

⁴Department of Gastroenterology and Hepatology, University of Heidelberg, Heidelberg, Germany

Contributions: (I) Conception and design: AT Billeter, B Reiners, SE Seide, P Probst, C Rupp, BP Müller-Stich; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: B Reiners, P Probst, E Kalkum; (V) Data analysis and interpretation: AT Billeter, B Reiners, SE Seide, P Probst, C Rupp, BP Müller-Stich; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Adrian T. Billeter, MD, PhD. Department of General, Visceral and Transplant Surgery, University of Heidelberg, Im Neuenheimer Feld 420, 69120 Heidelberg, Germany. Email: adrianbilleter@gmail.com.

Background: Non-alcoholic steatohepatitis (NASH) comprises a major healthcare problem affecting up to 30% of patients with obesity and the associated risk for cardiovascular and liver-related mortality. Several new drugs for NASH-treatment are currently investigated. No study thus far directly compared surgical and non-surgical therapies for NASH. This network meta-analysis compares for the first time the effectiveness of different therapies for NASH using a novel statistical approach.

Methods: The study was conducted according to the PRISMA guidelines for network meta-analysis. PubMed, CENTRAL and Web of Science were searched without restriction of time or language using a validated search strategy. Studies investigating therapies for NASH in adults with liver biopsies at baseline and after at least 12 months were selected. Patients with liver cirrhosis were excluded. Risk of bias was assessed with ROB-2 and ROBINS-I-tools. A novel method for population-adjusted indirect comparison to include and compare single-arm trials was applied. Main outcomes were NASH-resolution and improvement of fibrosis.

Results: Out of 7,913 studies, twelve randomized non-surgical studies and twelve non-randomized surgical trials were included. NASH-resolution after non-surgical intervention was 29% [95% confidence interval (CI): 23–40%] and 79% (95% CI: 72–88%) after surgery. The network meta-analysis showed that surgery had a higher chance of NASH-resolution than medication [odds ratio (OR) =2.68; 95% CI: 1.44–4.97] while drug treatment was superior to placebo (OR =2.24; 95% CI: 1.55–3.24). Surgery (OR =2.18; 95% CI: 1.34–3.56) and medication (OR =1.79; 95% CI: 1.39–2.31) were equally effective to treat fibrosis compared to placebo without difference between them. The results did not change when only new drugs specifically developed for the treatment of NASH were included.

Conclusions: Metabolic surgery has a higher effectiveness for NASH-therapy than medical therapy while both were equally effective regarding improvement of fibrosis. Trials directly comparing surgery with medication must be urgently conducted. Patients with NASH should be informed about surgical treatment options.

Keywords: Non-alcoholic fatty liver disease (NAFLD); non-alcoholic steatohepatitis (NASH); metabolic surgery; fatty liver

Submitted Jan 08, 2021. Accepted for publication May 08, 2021.

doi: 10.21037/hbsn-21-5

View this article at: <https://dx.doi.org/10.21037/hbsn-21-5>

Introduction

Non-alcoholic fatty liver disease (NAFLD), defined as a liver containing an abnormal amount of fat (>5%) (1), is a rapidly growing global health problem with a worldwide prevalence of approximately 25–30% (2). This chronic liver disease is associated with an increased risk for liver failure and hepatocellular carcinoma (HCC) and is expected to become the main indication for liver transplantation in the United States and worldwide in the next years (3). With an estimated prevalence of 70% in obese patients (4), the pathogenesis of NAFLD is linked to multiple factors, including genetics, obesity, metabolic syndrome and type 2 diabetes (T2D) (5). Non-alcoholic steatohepatitis (NASH) is the progressive form of NAFLD and is highlighted by an active inflammation leading to fibrosis (1). About 30–40% of patients with NAFLD develop NASH and NASH is associated with 5–10 folds increased risk for liver-related death (1). In contrast, NAFLD is more harmless and is only dangerous if it progresses to NASH.

To date, there is no approved therapy for NASH apart from lifestyle modification, which requires a weight loss of 7–10% for substantial improvement in liver histology beyond steatosis (6). As most patients do not achieve such long-term weight loss with lifestyle interventions alone, studies have focused on various drug therapies (7). Older drugs for NASH such as pioglitazone were associated with weight gain, bone fractures, and increased incidence of serious heart failure (8–10) while the use of vitamin E was associated with increased mortality (11). Reflecting the urgent need of an effective NASH-therapy, several trials investigating new agents have been conducted including resmetirom, obeticholic acid, emricasan, silymarin, and second-generation insulin sensitizer with promising results (12–16). Furthermore, the GLP-1 receptor (GLP-1R) analogue liraglutide achieves durable weight loss and direct effects on NASH pathogenesis are being discussed (17). The most notable large-scale trials investigating drugs for the treatment of NASH (PIVENS, LEAN, FLINT and REGENERATE) are included in this analysis (15,17–19).

Bariatric-metabolic surgery has been shown to achieve and maintain significant weight loss while also improving the metabolic syndrome and T2D, even in non-severely obese patients (20–22). Furthermore, several studies showed

substantial improvement in liver histology and fibrosis in many patients up to five years after surgery (23). The aim of this systematic review with matched network meta-analysis was to compare the effectiveness of surgical and non-surgical therapies regarding NASH and fibrosis. Since only single-arm surgical studies were available, an innovative statistical approach was applied that allows the inclusion of single-arm trials.

Methods

This systematic review and matched network meta-analysis was reported according to the PRISMA reporting checklist (available at <https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-5/rc>) (24).

Population and outcome of interest

The population of interest was adult patients with NASH. Interventions were bariatric-metabolic operations, pharmacological therapies and/or lifestyle interventions. For the drug trials, a placebo control group was mandatory for inclusion. An overview of the included therapies is provided in <https://cdn.amegroups.cn/static/public/hbsn-21-5-1.pdf>. The outcomes were histological resolution of NASH after 1 year as well as improvement of inflammation and ballooning. Furthermore, change of fibrosis and overall NAFLD-activity score (NAS) (25) were assessed. Resolution of NASH was defined as normalization of liver histology (NAS ≤2) and severe NASH was defined as NAS ≥5 (25). Lastly, surgery-related complications as well as serious adverse events due to the medication were evaluated.

Literature search and study selection

Randomized controlled, prospective, and retrospective cohort studies were included. A literature search in MEDLINE, The Cochrane Central Register of Controlled Trials (CENTRAL) and Web of Science was conducted using a validated search strategy (26). The last literature search was performed on December 21st, 2020 without restriction regarding time or language. The comprehensive search strategy can be found in the [Appendix 1](#). Two

reviewers screened article titles and abstracts according to inclusion and exclusion criteria, and the resulting full-text articles were further assessed for eligibility.

Data extraction

Data of population characteristics, liver histology and metabolic parameters were extracted with standard deviation or 95% confidence interval (95% CI). Percentage (i.e., % of patients with diabetes, % NASH-resolution) were converted into absolute values as “n participants”, glucose values into mg/dL, liver enzymes into IU/L and metabolic parameters into mmol/L. Extraction was performed by two reviewers independently. A third reviewer resolved any discrepancies.

Quality assessment

A quality assessment of the trials was performed using the ROBINS-I tool for non-randomized studies (NRS) and ROB-2 tool for randomized trials according to the Cochrane recommendations (27). Reporting, external and internal validity as well as power were critically appraised (Table 1).

Statistical analysis

Continuous baseline characteristics were reported as pooled means adjusted for sample size (pooled weighted mean). For NAS and NAS-components, only corresponding pre- and post-treatment data were included. As not all trials provided standard deviations for NAS and NAS components, range of study means was displayed for these studies. Outcomes were analyzed for each study arm and each individual intervention separately providing a mean effectiveness for each treatment (Figure S1).

All surgical studies were single-arm trials without conservative control group. Thus, a novel but established statistical method for population-adjusted indirect comparison was applied (28-30). To combine evidence on surgery and medication in our investigation while at the same time adjusting for baseline-differences, surgical trials were matched using age, gender, and type 2 diabetes mellitus (T2DM) prior to the network meta-analyses (28-30). Then, random-effects network meta-analyses adjusted for pre-treatment BMI, or, in a second model, adjusted for the change in BMI, were performed on the matched data set using the odds ratio (OR) as effect measure (31). Following the recommendations of Leahy *et al.*, we

included the baseline-odds for the reference group in the estimation (plug-in model) and compared the results to network meta-analyses where the matched trials were treated similar to the randomized controlled trials (RCTs) (naïve pooling) and results when using only RCTs on the comparison of medication and placebo (RCT only) (32). By using the matching on aggregated data level, we were able to obtain combined treatment effect estimates along with their 95% credible intervals on the comparison of placebo with medication, placebo with surgery and medication with surgery. A subgroup analysis for novelty of medical treatment was performed for outcomes where enough data was available. Network meta-analyses were estimated in jags and R (33,34). Network meta-analyses were fit using a Bayesian regime using 3 chains of 1,000,000 observations each, a burn-in of 10,000 iterations and a thinning factor of 10.

Results

Study selection

A total of 7,913 articles were identified by the literature search (last search December 21st, 2020). The study selection is outlined in Figure 1. Ultimately, 24 articles (12 NRS and 12 RCTs) were included in the qualitative and quantitative analysis. An overview is displayed in <https://cdn.amegroups.cn/static/public/hbsn-21-5-1.pdf>.

Qualitative assessment

The ROBINS tool revealed moderate to serious concerns for bias in the surgical NRS, primarily due to lack of randomization and a lack of control groups (Table 1). Also, a selection bias was present because only patients who were successfully operated and that qualified for an operation, both from a medical and insurance point of view, were included. Lastly, patients without follow-up biopsy were not reported or analyzed.

The RoB2 tool revealed no serious concerns for bias in any of the non-surgical trials. All trials were controlled RCTs with blinding of patients and assessors while outcome was reported according to predefined outcomes. Sponsor contribution was disclosed in all trials although conflict of interest may induce a potential bias. Five trials were fully funded by pharmaceutical companies with involvement in study design and data analysis in two trials (12,14) and involvement in study design in 3 trials (13,15,35). Four

Table 1 Quality assessment ROB2 and ROBINS tool

Study ID	Risk of bias (ROB2)	Study ID	Risk of bias (ROBINS tool)
Non-surgical trials		Surgical trials	
LEAN trial 2016	Low	de Almeida <i>et al.</i> 2006	Serious
FLINT trial 2015	Low	Lassailly <i>et al.</i> 2015	Moderate
Promrat <i>et al.</i> 2010	Some concerns	Dixon <i>et al.</i> 2004	Serious
Ratziu <i>et al.</i> 2016	Low	Froylich <i>et al.</i> 2016	Serious
PIVENS trial 2010	Low	Mathurin <i>et al.</i> 2009	Moderate/serious
Torres <i>et al.</i> 2011	Low/some concerns	Salman <i>et al.</i> 2020	Moderate/serious
Younossi <i>et al.</i> 2019	Low	von Schönfels <i>et al.</i> 2018	Serious
Harrison <i>et al.</i> 2019	Low	Tai <i>et al.</i> 2012	Serious
Harrison <i>et al.</i> 2019 EMMINENCE	Low	Vargas <i>et al.</i> 2012	Moderate/serious
Harrison <i>et al.</i> 2019 Emricasan	Low/some concerns	Nikai <i>et al.</i> 2020	Serious
Bril <i>et al.</i> 2019	Low	Salman <i>et al.</i> 2021	Moderate/serious
Wah Kheong <i>et al.</i> 2017	Low	Chaim <i>et al.</i> 2020	Moderate

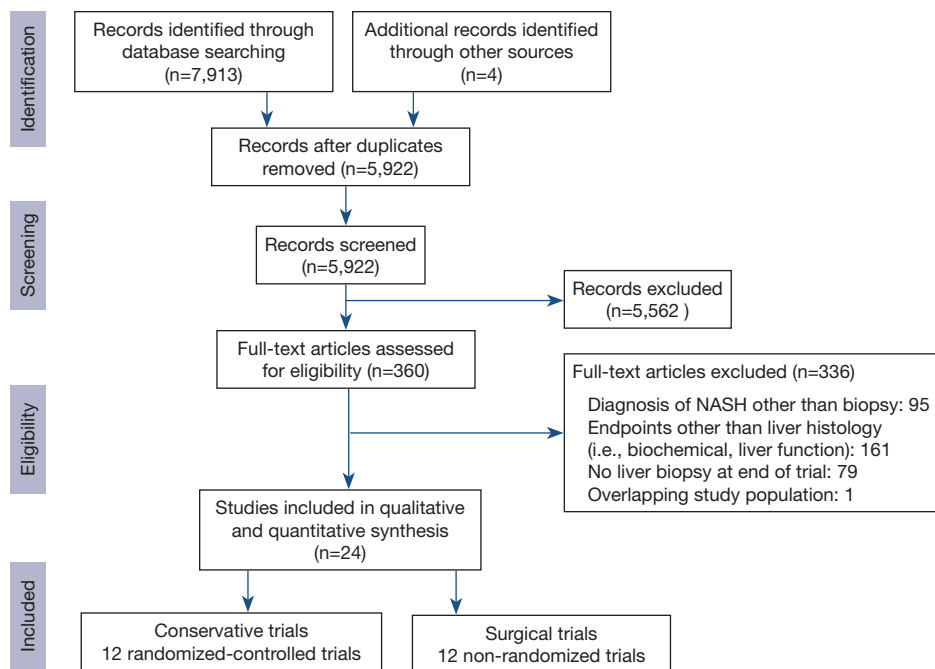


Figure 1 PRISMA flow chart. NASH, non-alcoholic steatohepatitis.

trials were partially funded with contribution to study design [FLINT (18)] and no relevant role of sponsor in PIVENS (19), LEAN (17) and Kheong (16). Promrat (36), Torres (37) and Bril (38) received no pharmaceutical funding.

Descriptive characteristics and baseline liver histology

The surgical trials comprised of ten prospective cohort trials (39-48) and two with comparative (49,50) designs

Table 2 Mean NAS score at first and second biopsy

Study and group	NAS pre ¹	NAS post ¹
<i>Surgical trials</i>		
de Almeida <i>et al.</i> 2006	–	–
Lassailly <i>et al.</i> 2015	5.00	1.00
Dixon <i>et al.</i> 2004	–	–
Froylich <i>et al.</i> 2016 RYGB	4.40	1.40
Froylich <i>et al.</i> 2016 SG	2.60	0.90
Mathurin <i>et al.</i> 2009	3.71	2.13
Salman <i>et al.</i> 2020	5.20	2.63
von Schönfels <i>et al.</i> 2018	2.40	0.90
Tai <i>et al.</i> 2012	2.60	0.80
Vargas <i>et al.</i> 2012	–	–
Nikai <i>et al.</i> 2020	3.9	1
Salman <i>et al.</i> 2021	6	3
Chaim <i>et al.</i> 2020	–	–
<i>Mean</i> ²	4.4 ²	2.0 ²
<i>Non-surgical trials</i>		
Armstrong <i>et al.</i> 2016		
Liraglutide 1.8 mg/d	4.90	3.60
Neuschwander-Tetri <i>et al.</i> 2015		
Obi acid 25 mg/d	5.30	3.60
Promrat <i>et al.</i> 2010		
Lifestyle modification	4.40	1.90
Ratziu <i>et al.</i> 2016		
Elaf 80 mg/d	4.90	–
Elaf/120 mg/d	5.00	–
Sanyal <i>et al.</i> 2010		
Pioglitazone	5.00	4.50
Vit E	5.10	3.20
Torres <i>et al.</i> 2011		
Rosiglitazone	5.10	3.33
Ros. + MTF	4.10	2.78
Ros. + Los.	4.90	3.53
Harrison <i>et al.</i> 2019		
Resmetiron	4.90	–

Table 2 (continued)**Table 2** (continued)

Study and group	NAS pre ¹	NAS post ¹
Younossi <i>et al.</i> 2019		
Obitecholic acid 10 mg	–	–
Obitecholic acid 25 mg	–	–
Harrison <i>et al.</i> 2019		
MSDC-0602 K 62.5 mg	5.5	4.7
MSDC-0602 K 125 mg	5.20	4.1
MSDC-0602 K 250 mg	5.40	4.3
Harrison <i>et al.</i> 2019		
Emricasan 5 mg	5.60	5.34
Emricasan 50 mg	5.50	5.46
Bril <i>et al.</i> 2019		
Vit E	3.90	–
Vit E + Pioglitazone	3.70	–
Wah Kheong <i>et al.</i> 2017		
Silymarin	5.00	4.4
<i>Mean</i> ²	5.2 ²	4.2 ²

¹, means as reported by the trials; ², pooled mean adjusted for sample size. NAS, NAFLD-activity score; NAFLD, non-alcoholic fatty liver disease.

comparing Roux-Y gastric bypass (RYGB) with sleeve gastrectomy. Mean time between surgery and re-biopsy was 13.6 (12–25.6) months. Mean age of the patients was 41.1±9.2 years with a BMI of 48.3±7.8 kg/m². NAS was reported in eight trials with a mean of 4.4 (2.4–6.0) at baseline (*Table 2*).

The twelve non-surgical studies comprised of five two-armed (13,16–18,36), one four-armed (12), five three-armed placebo-controlled trials (14,15,19,35,38), and one three-armed randomized trial (37). Mean time between the two biopsies was 16 (9–24) months. The average age was 53.2±11.2 years and BMI of 34.0±4.8 kg/m². Mean NAS was 5.2±0.9 at baseline (*Table 2*).

Mean NAS decreased by –2.4 (–1.5 to –4) after surgery. The steatosis-score fell by –1.1 (–1.0 to –1.6), ballooning by –0.6 (–0.3 to –1.1) and inflammation by –0.4 (–0.1 to –1). In the non-surgically studies, NAS improved by –1 (0.04 to –2.4). The steatosis-score fell by –0.5 (0.1 to –1.3), ballooning by –0.2 (0.1 to –0.6) and inflammation by –0.3 (0.09 to –0.6).

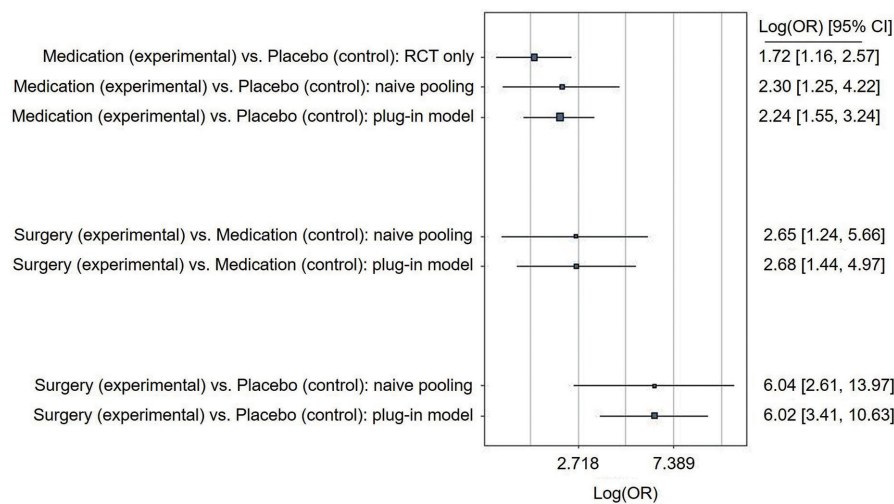


Figure 2 Matched network meta-analysis for resolution of NASH. NASH, non-alcoholic steatohepatitis; RCT, randomized controlled trial; OR, odds ratio; CI, confidence interval.

Resolution of NASH

In the surgical trials, 501 patients had liver biopsies with NASH at baseline. One year after surgery, 383 of 501 (76%; 95% CI: 72–88%) patients had NASH-resolution ($NAS \leq 2$; [Figure S1A](#)). In patients with $NAS \geq 5$ at baseline ($n=141$), NASH-resolution was 76% (95% CI: 60–92%; [Figure S1B](#)).

In the non-surgical intervention groups, 1,637 patients had biopsy-proven NASH with a NASH-resolution of 29% (95% CI: 23–40%; [Figure S1A](#)) with the lifestyle intervention of the Promrat-trial having the strongest effect while emricasan had the least effect (14,36). In patients with $NAS \geq 5$ at baseline ($n=758$), NASH-resolution was 28% (95% CI: 16–39%; [Figure S1B](#)).

Matched network meta-analysis

Resolution of NASH

NASH-resolution adjusted for pre-treatment BMI was estimated in the first model. Due to the large baseline difference in BMI, the model provided huge intervals that rendered any statistical analysis futile. In this model, medication was superior regarding NASH-resolution compared to placebo in the plug-in model (OR =2.63; 95% CI: 1.55–4.45), which was not observed in either the naïve model (OR =1.80; 95% CI: 0.36–8.97) or the RCT-only model (OR =1.46; 95% CI: 0.56–3.77). Neither when using the plug-in model, nor when using the naïve pooling, a difference between surgery and medication or surgery and control was observed ([Figure S2](#)). Between-trial

heterogeneity was high in all models (plug-in: 0.80, naïve: 1.06, RCT-only: 0.48) and could not be explained by pre-treatment BMI.

Since pre-treatment BMI did not reduce the observed heterogeneity, change in BMI before and after treatment (BMI-difference) was included as covariate in a second model. Here, medication was superior for NASH-resolution compared to placebo using the plug-in model (OR =2.24; 95% CI: 1.55–3.24). Surgery was superior regarding NASH-resolution in the plug-in model compared to placebo (OR =6.02; 95% CI: 3.41–10.63) and medication (OR =2.68; 95% CI: 1.44–4.97; [Figure 2](#)). When adjusting for the BMI-difference, results for the naïve model and the RCT-only model pointed in a similar direction ([Figure 2](#)). Heterogeneity was slightly reduced by adjusting for BMI-difference instead of pre-treatment BMI in all three models (plug-in: 0.78, naïve: 1.01, RCT-only: 0.42).

When only the newly developed drugs for NASH-treatment (MSDC-0602K; obeticholic acid; emricasan; liraglutide; elafibranor) were analyzed, the estimated treatment effects pointed in a similar direction. However, all estimates were associated with higher imprecision due to the limited availability of data and credible intervals are very wide for this subgroup analysis (e.g., surgery *vs.* medication: OR =6.16; 95% CI: 2.26–16.73).

Improvement of ballooning and inflammation

Detailed histological information was given for 272 surgical and 1,212 non-surgical patients. Ballooning was improved

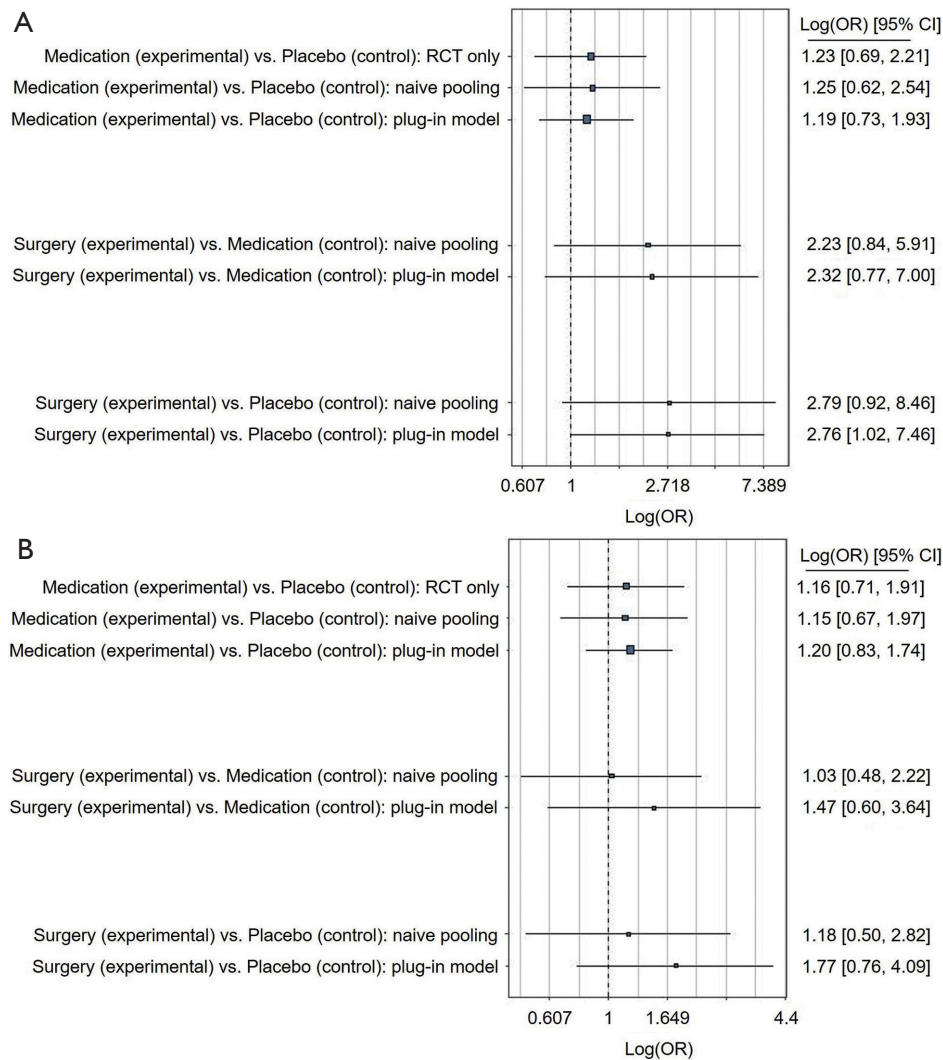


Figure 3 Matched network meta-analysis for (A) ballooning and (B) inflammation. RCT, randomized controlled trial; OR, odds ratio; CI, confidence interval.

in 78% (95% CI: 69–88%) after surgery compared to 40% (95% CI: 31–49%) after non-surgical therapy (Figure S3A). The network meta-analysis revealed no difference in resolution of ballooning in the plug-in model between medication and placebo (OR =1.19; 95% CI: 0.73–1.93) and surgery *vs.* medication (OR =2.32; 95% CI: 0.77–7.00; Figure 3A). Surgery *vs.* placebo was borderline significant in the plug-in model (OR =2.76; 95% CI: 1.02–7.46); however, the credible interval included the OR of no-effect more clearly when using the naïve model (OR =2.79; 95% CI: 0.92–8.46; Figure 3A). Between-trial heterogeneity was 0.75 in the plug-in model, 0.81 in the naïve pooling model, and 0.62 in the RCT-only model.

Lobular inflammation improved in 66% (95% CI: 55–78%) after surgery and 43% (95% CI: 32–55%) after conservative therapy (Figure S3B). The network meta-analysis revealed no difference in inflammation-resolution for any treatments using any of the models (Figure 3B).

Improvement of fibrosis

A total of 266 patients in the surgical group had fibrosis at baseline with improvement in 63% (95% CI: 47–80%, Figure S3C). However, Mathurin demonstrated worsening in fibrosis in 19.8% at 1 year after surgery although the rate of fibrosis was low and thus likely not representative. Froylich reported an overall improvement of fibrosis of

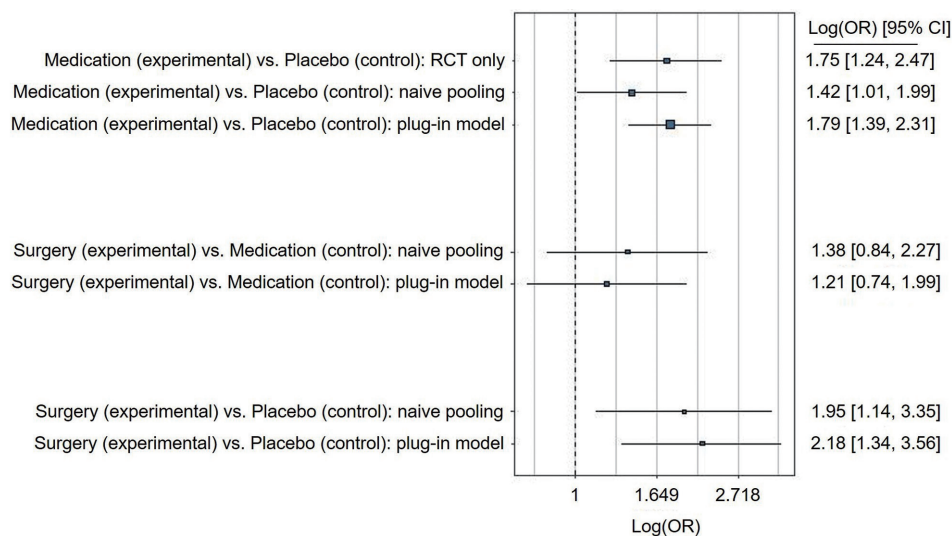


Figure 4 Matched network meta-analysis for fibrosis. RCT, randomized controlled trial; OR, odds ratio; CI, confidence interval.

-0.86 ± 1.3 from 1.9 ± 1.5 . However, a worsening in 3 patients (12%) after RYGB was also observed (49). Nikai reported overall improvement of fibrosis after laparoscopic sleeve gastrectomy (LSG). However, ten patients with fibrosis without steatosis and low initial NAS had persistence of fibrosis after LSG (45).

In the non-surgical group, 1,505 patients had fibrosis at baseline. Improvement was observed in 23% (95% CI: 18–28%) except in the Promrat-trial, in which fibrosis remained unchanged in all patients. The LEAN-trial reported worsening of fibrosis in 2 patients (9%) after liraglutide and 8 patients (36%) in the placebo group (17). Younossi observed progression of fibrosis in 17% after obeticholic acid 10 mg/d and 13% after 25 mg/d (15). Harrison reported lower rates of worsening of fibrosis after placebo (20%) compared to emricasan 5 mg (41%) and 50 mg (38%) (14).

The network meta-analysis revealed a better improvement of fibrosis in the plug-in model comparing medication with placebo (OR =1.79; 95% CI: 1.39–2.31) and for surgery *vs.* placebo (OR =2.18; 95% CI: 1.34–3.56; *Figure 4*). However, metabolic surgery had a comparable effect compared to medication (OR =1.21; 95% CI: 0.74–1.99; *Figure 4*). Similar results were observed in the naïve model and the RCT-only model. Between-trial heterogeneity were considerably lower than for the previously described outcomes (plug-in: 0.28, naïve: 0.37, and RCT-only: 0.28).

When only the newly developed drugs for NASH-treatment (MSDC-0602K; obeticholic acid; emricasan; liraglutide; elafibranor) were compared to surgery, the

effects remained comparable (surgery *vs.* medication: OR =1.32; 95% CI: 0.64–2.73).

Comparison of surgical procedures

von Schönfels reported a decrease in NAS of -1 in the RYGB group (baseline NAS =2) and -2.5 after LSG (baseline NAS =2.5) with similar histological improvement after both procedures (50). Similar results were observed by Froylich *et al.* with -3 after RYGB (baseline NAS =4.4) and -1.7 after LSG (baseline NAS =2.6) with a stronger improvement of NAS and fibrosis after RYGB although without statistical significance (49). No relevant difference in NAS improvement between LSG and RYGB was observed in either trial.

Surgery-related complications and serious adverse events

Two surgical studies provided information about complications. Tai reported complications in 3/21 patients (2 gastrojejunostomic stenosis, 1 bleeding of gastrojejunostomic ulcer), none of which were life-threatening or required re-operation (43). Vargas reported no complications (44). There were no deaths reported in the surgical studies.

Treatment in the non-surgical trials was overall well tolerated and no trial reported significantly higher rates of serious adverse events in the treatment groups compared to placebo. Most frequent adverse event with obeticholic acid was pruritus [28–51% (15,18)] although discontinuation

because of pruritus was maximum 9% (15,18). Gastrointestinal symptoms (diarrhea, constipation, loss of appetite) occurred more frequently with liraglutide than placebo (81% *vs.* 65%). However, study withdrawal and serious events were similar (8% liraglutide *vs.* 4% placebo) (17). PIVENs reported similar frequency of cardiovascular events (12–14%) and did not observe an increase of bone fractures (19). Yet, Brill reported higher occurrence of edema (13% after combination *vs.* 0% after vitamin E and 3% after placebo) and weight gain (+5.7 kg in pioglitazone + vitamin E *vs.* +0.5 kg in vitamin E group and -0.8 kg in placebo) (38). Diarrhea was more frequent after resmetiron (33% *vs.* 9% placebo) (13) and emricasan (9–16% *vs.* 10% placebo) (14) and nausea with elafibranor (10–13% *vs.* 9% placebo) (35).

Discussion

This systematic review with matched network meta-analysis demonstrated the beneficial and reproducible effects of metabolic surgery on NASH with a resolution in 79% compared to 29% in non-surgical interventions. After adjusting for baseline differences, the network meta-analysis demonstrated a higher rate of NASH-resolution after metabolic surgery (2.5-fold *vs.* medication; 6-fold *vs.* placebo) while medication had a 2-fold higher chance of NASH-resolution than placebo. Regarding fibrosis, both surgery and medication had a higher chance for improvement of fibrosis compared to placebo while there was no difference between surgery and medication.

In contrast to prior meta-analyses, this work applied a novel but established method for population-adjusted indirect comparisons to include and compare single-arm trials. This approach allowed to compare the effectiveness of various treatment options for NASH and fibrosis (28–30). Furthermore, the study focused on histologically-proven NASH and fibrosis while simple NAFLD was excluded. This is of relevance since NASH and fibrosis are the strongest risk factors for progression to HCC and cirrhosis (51,52). Comparing the efficacy of the treatments, this analysis indicated that surgery is more effective to treat NASH than the currently tested drugs which appear also less effective than the older drugs and thus seem not to constitute a promising option to treat the NASH epidemic (Figure S1) (12–15,17). However, regarding fibrosis, there is an equal effectiveness of surgery and medication (OR 1.21; 95% CI: 0.74–1.99) while both treatments are superior to placebo. Therefore, an RCT comparing metabolic surgery with medication is urgently needed to define the role of

surgery in treatment of NASH and reversal of fibrosis.

While these results are generally encouraging, a worsening of fibrosis was also observed in all treatments. Froylich observed a worsening only after RYGB while Nikai identified presence of steatosis in the first biopsy and HbA1c as the only predictive factor for improvement of fibrosis. Mathurin hypothesized that rapid weight loss may induce progression of fibrosis but also found a refractory insulin resistance to be a predictive factor. Another explanation may be that bypass procedures interfere too strongly with the enterohepatic circulation which may impede liver recovery. This hypothesis would be in line with the observation that acute liver failure after bariatric surgery was thus far only observed after bypass procedures (53,54). Nevertheless, recent data showed that patients after bariatric surgery had a lower risk of developing cirrhosis compared to a well-matched cohort (55). Since most patients with NAFLD-diagnosis have a BMI <35 kg/m², the authors even discuss extending the indication for bariatric surgery to lower BMI-patients with NAFLD. Furthermore, long-term follow-up from the Lille-bariatric cohort with repeat liver biopsy five years after bariatric surgery in patients with NASH also revealed a sustained NASH-resolution and even better improvement of fibrosis over time (23). Furthermore, progression of fibrosis was also reported in several non-surgical trials and in an overall higher rate [9% after liraglutide, 13–17% after obeticholic acid, 40% after emricasan (14,15,17)]. The placebo groups had even higher rates of worsening of fibrosis with 20–36% although only three trials reported this outcome (14,15,17) while Kheon *et al.* observed new onset fibrosis in 6% of the placebo group *vs.* 2% in the treatment group (16). In addition, recent analysis indicated that an effective treatment for NASH should also target cardiovascular disease to reduce mortality because the NASH treatment only reduced mortality in patients with cirrhosis (56). This observation is another argument for a role of metabolic surgery in the treatment of NASH since it is well established that metabolic surgery reduces cardiovascular mortality (57,58). These findings underline the urgent need for a trial comparing metabolic surgery with medication for NASH.

This analysis also provided some insight into the mechanisms of NASH-resolution and the impact of weight loss. Since the network meta-analysis was adjusted for baseline differences and BMI-loss, the persistently shown superior effect of surgery suggests that metabolic surgery has additional effects on the liver beyond weight loss. This hypothesis is in line with a recent experimental study (59).

Finally, the safety profile of surgical and non-surgical

therapies was analyzed. The two surgical studies reported minor complications without any deaths and bariatric-metabolic surgery is among the safest type of surgeries with a mortality of <1% and severe complications <5% (60). Even in advanced liver failure, metabolic surgery can be safely performed in experienced centers in selected patients (61,62). Nonetheless, the potential benefits and risks of metabolic surgery in patients with NASH must be carefully assessed in any patient. All medication trials reported good tolerability, with the most frequent adverse events being liver-unrelated and comprising of pruritus, diarrhea, and nausea. However, the findings in these studies are contradicted by the long-term data of large cohorts that found higher mortality in vitamin E treated patients (11) while glitazones had a higher rate of heart failure, weight gain, and fractures (8-10) which essentially led to the cessation of these drugs as NASH-therapy. Therefore, directly comparing safety, tolerability and quality of life is another important outcome of future trials.

There were several limitations of this analysis. All surgical trials were single-arm and therefore lacked randomization and a control group resulting in a risk of bias. While aggregating level-matching allows to estimate the treatment effects of surgery *vs.* medication and surgery *vs.* placebo that could otherwise not be estimated, these comparisons should clearly be investigated in RCTs. Another concern is the differences in baselines characteristics that are addressed by the matching procedure. The main difference is the higher BMI in the surgical trials since all surgically treated patients require a BMI ≥ 35 kg/m². However, since NASH is strongly associated with obesity, the patients in the non-surgical trials were also overweight and obese. Furthermore, surgical patients were roughly ten years younger than the patients in the non-surgical trials. This point raises the question whether duration of NASH may have an influence on the results since age is a risk factors for advanced disease and HCC (51,52,63). Furthermore, there is no long-term follow-up from these studies beyond the time point of the follow-up biopsy. Therefore, a reduction in disease progression, HCC or liver failure cannot be assessed in these cohorts. Lastly, for most medical therapies only one study was available which did not allow to compare the effectiveness of the different drugs. These issues were addressed as far as possible by the network meta-analysis, which was matched for baseline age, gender and T2D while the baseline BMI-difference was addressed by a covariate analysis using BMI-difference. In line with these concerns is the slightly lower mean NAS in the surgical trials (NAS 4.4)

compared to the non-surgical trials (NAS 5.1). Due to the low number of available studies, matching for baseline NAS was not possible. However, the severity of NASH (NAS ≥ 5) did not affect the effectiveness in surgical or non-surgical treatments (Figure S1). Despite these limitations, the results shown here represent the best evidence regarding the efficacy of treatment options for NASH.

Conclusions

Metabolic surgery appears to be the most effective treatment for NASH with a higher chance of NASH-resolution than medication or placebo while medication was superior compared to placebo. Furthermore, both metabolic surgery and medication improved fibrosis better than placebo without difference between surgery and medication. Despite these promising results in favor of metabolic surgery, the results must be interpreted cautiously due to the sole availability of single-arm trials for surgery. Furthermore, comparing the effectiveness of different medical therapies was also not possible since only one trial was available for most drugs. While underlying baseline differences were addressed through aggregated data matching and the inclusion as covariate, confirmatory studies directly comparing surgery with medication are urgently needed. Nonetheless, patients with NASH should be informed about surgery as treatment option before starting a pharmaceutical therapy.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at <https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-5/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://hbsn.amegroups.com/article/view/10.21037/hbsn-21-5/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Byrne CD, Targher G. NAFLD: a multisystem disease. *J Hepatol* 2015;62:S47-64.
2. Younossi ZM, Koenig AB, Abdelatif D, et al. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 2016;64:73-84.
3. Bellentani S, Scaglioni F, Marino M, et al. Epidemiology of non-alcoholic fatty liver disease. *Dig Dis* 2010;28:155-61.
4. Machado M, Marques-Vidal P, Cortez-Pinto H. Hepatic histology in obese patients undergoing bariatric surgery. *J Hepatol* 2006;45:600-6.
5. Marchesini G, Brizi M, Bianchi G, et al. Nonalcoholic fatty liver disease: a feature of the metabolic syndrome. *Diabetes* 2001;50:1844-50.
6. Vilar-Gomez E, Martinez-Perez Y, Calzadilla-Bertot L, et al. Weight Loss Through Lifestyle Modification Significantly Reduces Features of Nonalcoholic Steatohepatitis. *Gastroenterology* 2015;149:367-78.e5; quiz e14-5.
7. Francque S, Vonghia L. Pharmacological Treatment for Non-alcoholic Fatty Liver Disease. *Adv Ther* 2019;36:1052-74.
8. Balas B, Belfort R, Harrison SA, et al. Pioglitazone treatment increases whole body fat but not total body water in patients with non-alcoholic steatohepatitis. *J Hepatol* 2007;47:565-70.
9. Erdmann E, Charbonnel B, Wilcox RG, et al. Pioglitazone use and heart failure in patients with type 2 diabetes and preexisting cardiovascular disease: data from the PROactive study (PROactive 08). *Diabetes Care* 2007;30:2773-8.
10. Meymeh RH, Woollorton E. Diabetes drug pioglitazone (Actos): risk of fracture. *CMAJ* 2007;177:723-4.
11. Bjelakovic G, Nikolova D, Gluud LL, et al. Mortality in randomized trials of antioxidant supplements for primary and secondary prevention: systematic review and meta-analysis. *JAMA* 2007;297:842-57.
12. Harrison SA, Alkhoury N, Davison BA, et al. Insulin sensitizer MSDC-0602K in non-alcoholic steatohepatitis: A randomized, double-blind, placebo-controlled phase IIb study. *J Hepatol* 2020;72:613-26.
13. Harrison SA, Bashir MR, Guy CD, et al. Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* 2019;394:2012-24.
14. Harrison SA, Goodman Z, Jabbar A, et al. A randomized, placebo-controlled trial of emricasan in patients with NASH and F1-F3 fibrosis. *J Hepatol* 2020;72:816-27.
15. Younossi ZM, Ratziu V, Loomba R, et al. Obeticholic acid for the treatment of non-alcoholic steatohepatitis: interim analysis from a multicentre, randomised, placebo-controlled phase 3 trial. *Lancet* 2019;394:2184-96.
16. Wah Kheong C, Nik Mustapha NR, Mahadeva S. A Randomized Trial of Silymarin for the Treatment of Nonalcoholic Steatohepatitis. *Clin Gastroenterol Hepatol* 2017;15:1940-9.e8.
17. Armstrong MJ, Gaunt P, Aithal GP, et al. Liraglutide safety and efficacy in patients with non-alcoholic steatohepatitis (LEAN): a multicentre, double-blind, randomised, placebo-controlled phase 2 study. *Lancet* 2016;387:679-90.
18. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet* 2015;385:956-65.
19. Sanyal AJ, Chalasani N, Kowdley KV, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. *N Engl J Med* 2010;362:1675-85.
20. Billeter AT, Scheurlen KM, Probst P, et al. Meta-analysis of metabolic surgery versus medical treatment for microvascular complications in patients with type 2 diabetes mellitus. *Br J Surg* 2018;105:168-81.
21. Müller-Stich BP, Senft JD, Warschkow R, et al. Surgical versus medical treatment of type 2 diabetes mellitus in nonseverely obese patients: a systematic review and meta-analysis. *Ann Surg* 2015;261:421-9.
22. Vidal J, Corcelles R, Jiménez A, et al. Metabolic and Bariatric Surgery for Obesity. *Gastroenterology* 2017;152:1780-90.
23. Lassailly G, Caiazzo R, Ntandja-Wandji LC, et al. Bariatric Surgery Provides Long-term Resolution of Nonalcoholic Steatohepatitis and Regression of Fibrosis. *Gastroenterology* 2020;159:1290-301.e5.
24. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA

- extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162:777-84.
25. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313-21.
 26. Goossen K, Tenckhoff S, Probst P, et al. Optimal literature search for systematic reviews in surgery. *Langenbecks Arch Surg* 2018;403:119-29.
 27. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions*. version 6.0 (updated July 2019). Cochrane, 2019.
 28. Phillippo DM, Ades AE, Dias S, et al. Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal. *Med Decis Making* 2018;38:200-11.
 29. Phillippo D, Ades T, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Decision Support Unit, ScHARR, University of Sheffield: NICE Decision Support Unit, 2016. 81 p. (Technical Support Documents).
 30. Schmitz S, Maguire Á, Morris J, et al. The use of single armed observational data to closing the gap in otherwise disconnected evidence networks: a network meta-analysis in multiple myeloma. *BMC Med Res Methodol* 2018;18:66.
 31. Dias SDAE, Welton NJ, Jansen JP, et al. *Network Meta-Analysis for Decision Making*. John Wiley & Sons Ltd, 2018.
 32. Leahy J, Thom H, Jansen JP, et al. Incorporating single-arm evidence into a network meta-analysis using aggregate level matching: Assessing the impact. *Stat Med* 2019;38:2505-23.
 33. Plummer M. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. 2003. Accessed 2020.07.20. Available online: <https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf>
 34. R. The R Project for Statistical Computing. Accessed 2020.07.20. Available online: <https://www.r-project.org/>
 35. Ratziu V, Harrison SA, Francque S, et al. Elafibranor, an agonist of the peroxisome proliferator-activated receptor- α and- δ , induces resolution of nonalcoholic steatohepatitis without fibrosis worsening. *Gastroenterology* 2016;150:1147-59.e5. Erratum in: *Gastroenterology* 2017;152:2084.
 36. Promrat K, Kleiner DE, Niemeier HM, et al. Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. *Hepatology* 2010;51:121-9.
 37. Torres DM, Jones FJ, Shaw JC, et al. Rosiglitazone versus rosiglitazone and metformin versus rosiglitazone and losartan in the treatment of nonalcoholic steatohepatitis in humans: a 12-month randomized, prospective, open-label trial. *Hepatology* 2011;54:1631-9.
 38. Bril F, Biernacki DM, Kalavalapalli S, et al. Role of Vitamin E for Nonalcoholic Steatohepatitis in Patients With Type 2 Diabetes: A Randomized Controlled Trial. *Diabetes Care* 2019;42:1481-8.
 39. de Almeida SR, Rocha PR, Sanches MD, et al. Roux-en-Y gastric bypass improves the nonalcoholic steatohepatitis (NASH) of morbid obesity. *Obes Surg* 2006;16:270-8.
 40. Dixon JB, Bhathal PS, Hughes NR, et al. Nonalcoholic fatty liver disease: Improvement in liver histological analysis with weight loss. *Hepatology* 2004;39:1647-54.
 41. Mathurin P, Hollebecque A, Arnalsteen L, et al. Prospective study of the long-term effects of bariatric surgery on liver injury in patients without advanced disease. *Gastroenterology* 2009;137:532-40.
 42. Salman MA, Salman AA, Abdelsalam A, et al. Laparoscopic Sleeve Gastrectomy on the Horizon as a Promising Treatment Modality for NAFLD. *Obes Surg* 2020;30:87-95.
 43. Tai CM, Huang CK, Hwang JC, et al. Improvement of nonalcoholic fatty liver disease after bariatric surgery in morbidly obese Chinese patients. *Obes Surg* 2012;22:1016-21.
 44. Vargas V, Allende H, Lecube A, et al. Surgically induced weight loss by gastric bypass improves non alcoholic fatty liver disease in morbid obese patients. *World J Hepatol* 2012;4:382-8.
 45. Nikai H, Ishida K, Umemura A, et al. Effects of Laparoscopic Sleeve Gastrectomy on Non-Alcoholic Steatohepatitis and Liver Fibrosis in Japanese Patients with Severe Obesity. *Obes Surg* 2020;30:2579-87.
 46. Lassailly G, Caiazzo R, Buob D, et al. Bariatric Surgery Reduces Features of Nonalcoholic Steatohepatitis in Morbidly Obese Patients. *Gastroenterology* 2015;149:379-88; quiz e15-6.
 47. Chaim FDM, Pascoal LB, Chaim FHM, et al. Histological grading evaluation of non-alcoholic fatty liver disease after bariatric surgery: a retrospective and longitudinal observational cohort study. *Sci Rep* 2020;10:8496.
 48. Salman MA, Salman AA, Omar HSE, et al. Long-term effects of one-anastomosis gastric bypass on liver histopathology in NAFLD cases: a prospective study. *Surg Endosc* 2021;35:1889-94.
 49. Froylich D, Corcelles R, Daigle C, et al. Effect of

- Roux-en-Y gastric bypass and sleeve gastrectomy on nonalcoholic fatty liver disease: a comparative study. *Surg Obes Relat Dis* 2016;12:127-31.
50. von Schönfels W, Beckmann JH, Ahrens M, et al. Histologic improvement of NAFLD in patients with obesity after bariatric surgery based on standardized NAS (NAFLD activity score). *Surg Obes Relat Dis* 2018;14:1607-16.
 51. Alexander M, Loomis AK, van der Lei J, et al. Risks and clinical predictors of cirrhosis and hepatocellular carcinoma diagnoses in adults with diagnosed NAFLD: real-world study of 18 million patients in four European cohorts. *BMC Med* 2019;17:95.
 52. Hagström H, Nasr P, Ekstedt M, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J Hepatol* 2017;67:1265-73.
 53. Mahawar KK. Liver Dysfunction with Both Roux-en-Y and One-Anastomosis Gastric Bypass Is Almost Exclusively Seen with Longer Than Standard Limb Lengths. *Obes Surg* 2018;28:548-9.
 54. Eilenberg M, Langer FB, Beer A, et al. Significant Liver-Related Morbidity After Bariatric Surgery and Its Reversal—a Case Series. *Obes Surg* 2018;28:812-9.
 55. Wirth KM, Sheka AC, Kizy S, et al. Bariatric Surgery is Associated With Decreased Progression of Nonalcoholic Fatty Liver Disease to Cirrhosis: A Retrospective Cohort Analysis. *Ann Surg* 2020;272:32-9.
 56. Rowe I, Hagstrom H, Ekstedt M, et al. The predicted benefits of treatments for nonalcoholic steatohepatitis are limited in patients without cirrhosis. *Hepatology* 2019;70:755A-6A.
 57. Billeter AT, Eichel S, Scheurlen KM, et al. Meta-analysis of metabolic surgery versus medical treatment for macrovascular complications and mortality in patients with type 2 diabetes. *Surg Obes Relat Dis* 2019;15:1197-210.
 58. Sjöström L, Peltonen M, Jacobson P, et al. Bariatric surgery and long-term cardiovascular events. *JAMA* 2012;307:56-65.
 59. Watanabe K, Ohta M, Takayama H, et al. Effects of Sleeve Gastrectomy on Nonalcoholic Fatty Liver Disease in an Obese Rat Model. *Obes Surg* 2018;28:1532-9.
 60. Gero D, Raptis DA, Vleeschouwers W, et al. Defining Global Benchmarks in Bariatric Surgery: A Retrospective Multicenter Analysis of Minimally Invasive Roux-en-Y Gastric Bypass and Sleeve Gastrectomy. *Ann Surg* 2019;270:859-67.
 61. Vuppalanchi R, McCabe ME 4th, Tandra SR, et al. Safety and Efficacy of Bariatric Surgery in Cirrhosis Patients With Extreme Obesity. *Ann Surg* 2022;275:e174-80.
 62. Agarwal L, Sahu AK, Baksi A, et al. Safety of metabolic and bariatric surgery in obese patients with liver cirrhosis: a systematic review and meta-analysis. *Surg Obes Relat Dis* 2021;17:525-37.
 63. Ioannou GN, Green P, Kerr KF, et al. Models estimating risk of hepatocellular carcinoma in patients with alcohol or NAFLD-related cirrhosis for risk stratification. *J Hepatol* 2019;71:523-33.

Cite this article as: Billeter AT, Reiners B, Seide SE, Probst P, Kalkum E, Rupp C, Müller-Stich BP. Comparative effectiveness of medical treatment *vs.* metabolic surgery for histologically proven non-alcoholic steatohepatitis and fibrosis: a matched network meta-analysis. *HepatoBiliary Surg Nutr* 2022;11(5):696-708. doi: 10.21037/hbsn-21-5