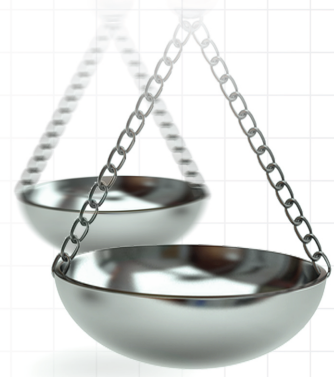
Meta-Analysis

For reprint orders, please contact: info@becaris.com



Journal of **Comparative Effectiveness Research**

Prognostic value of liver histology in steatotic liver disease: a systematic literature review and meta-analysis

A Sidney Barritt IV[1], Anila Qasim[2], Yestle Kim*[,3], Sally D Miller[2], John O'Donnell[3] & Karissa Johnston[2]

[1]Division of Gastroenterology & Hepatology, University of North Carolina, Chapel Hill, NC 27599, USA

[2]Broadstreet HEOR, Vancouver V5Y 1L8, Canada

[3]Madrigal Pharmaceuticals, West Conshohocken, PA 19428, USA

*Author for correspondence: Tel.: +1 267 299 8737; ykim@madrigalpharma.com

Aim: Metabolic dysfunction-associated steatohepatitis (MASH) and steatotic liver disease (MASLD) are progressive liver conditions that can lead to serious long-term outcomes and adverse clinical events. Histologic liver fibrosis is an accepted short-term surrogate end point used in pivotal clinical trials supporting conditionally approved MASH therapies. This study aimed to synthesize up-to-date published associations between histologic liver fibrosis and clinical outcomes, including several novel syntheses. **Materials & methods:** A systematic literature review identified studies of MASH ± MASLD patients published January 2014 to November 2024 from Ovid MEDLINE®, Embase and grey literature. Studies reporting hazard ratios (HRs) comparing the risk of relevant major adverse liver outcomes (MALO; e.g., cirrhosis, hepatic decompensation, hepatocellular carcinoma) and mortality by histologic liver fibrosis stage were included in meta-analysis. Pooled estimates were reported as HRs and 95% CIs. **Results:** Of 2810 returned records, there were 32 eligible studies from 39 articles. Higher fibrosis stage was associated with increased risk of clinical outcomes. Risk of progression to cirrhosis was twofold higher in F3 versus F2 (2.05 [1.45, 2.90]). Risk of hepatic decompensation was >ten-times higher in F3–4 versus F0–2 (10.93 [6.31, 18.92]). Risk of MALO and all-cause mortality were also significantly higher in F4 versus F3 and versus F2, and in F3–4 versus F0–2. Results were directionally consistent among studies of MASH-majority (≥80% MASH) populations and individually reported HRs. **Conclusion:** Increased clinical risk with more advanced fibrosis and/or cirrhosis among patients with significant liver disease supports the value of histologic liver fibrosis as a short-term surrogate for long-term clinical outcomes.

Plain language summary: Using long-term studies to connect changes in liver tissue to risk of serious consequences for patients with metabolic dysfunction-associated steatohepatitis

What is this article about? In chronic liver diseases such as metabolic dysfunction-associated steatohepatitis (MASH), symptoms may not be visible until they have progressed to serious consequences such as liver failure or liver cancer, which may be life-threatening. A challenge of testing new medications for MASH is the length of time it can take for these serious consequences to occur, which can be longer than the 1–2 years of most clinical trials. In such cases, a clinical trial may use a 'proxy' test that is known to be connected to serious long-term consequences, but shows improvements in a shorter timeframe. Previous research has found that a potentially useful proxy in MASH is liver scarring ('fibrosis'), and many studies have individually found a connection between severity of liver fibrosis and risk of serious long-term consequences. This article aimed to combine the data from these studies to determine how consistent and strong this connection is across studies.

How was the research carried out? The authors reviewed previously published studies measuring both liver fibrosis and later occurrence of serious consequences, including liver swelling, bleeding, failure, transplant and death. Then the individual results of each study were pooled together in a combined analysis.

What were the results? Both individual study results and the combined analysis showed that increased severity of liver fibrosis was associated with increased risk of cirrhosis and other liver problems leading to hospitalization, as well as death, indicating that liver fibrosis is likely a useful proxy test for clinical trials.

First draft submitted: 8 July 2026; Accepted for publication: 17 August 2026; Published online: 3 Septemeber 2026

Keywords: liver fibrosis • MASH • MASLD • surrogate endpoints

Metabolic dysfunction-associated steatotic liver disease (MASLD), and its progressive form, metabolic dysfunction-associated steatohepatitis (MASH) are increasingly common liver conditions that may involve inflammation and fibrosis, leading to serious health consequences [1]. The incidence rate of progression from MASL (the mildest form of MASLD) to MASH is 7.4 per 100 person-years (equivalent to 31% of patients over 4.7 years) [2] and patients are at increased risk for long-term consequences, including liver cirrhosis, decompensation (typically involving symptoms such as ascites, variceal bleeding, hepatic encephalopathy), hepatocellular carcinoma (HCC), which can result in liver failure, requiring liver transplant, and death [3]. MASH is the most common reason for liver transplantation among women in the US [4], and is expected to continue increasing in frequency among both men and women, in conjunction with increased prevalence of MASLD risk factors such as obesity, Type 2 diabetes mellitus and dyslipidemia in the US [5] and globally [6]. The most common causes of death in MASLD and MASH include cardiovascular disease, liver disease and cancer [7].

The primary area of unmet need for MASLD and MASH is a treatment that can slow, halt or potentially reverse disease progression, thus preventing serious clinical outcomes [8]. Currently, there are only two conditionally approved therapeutics available for treatment of MASH: resmetirom, a thyroid hormone receptor-beta agonist conditionally approved by the US FDA in 2024 [9] and by the EMA in 2025 [10], and semaglutide, a glucagon-like peptide 1 (GLP-1) receptor agonist conditionally approved by the FDA in August 2025 [11]. Capturing clinical event outcomes in the context of clinical trials is limited by practical considerations. For example, a recent observational study reported that just 2.1% of MASLD patients developed HCC over a mean follow-up time of 3.4 years [12]. The low incidence rate may make it difficult to meet sample size requirements to differentiate outcomes between active and comparator arms, particularly in the time frame of a typical clinical trial.

In order to accommodate this practical limitation, regulatory agencies including the FDA and EMA, have adopted surrogate end points, such as histologic liver fibrosis stage, which can show meaningful change over a shorter timeframe [13,14]. The utility of histologic liver fibrosis stage, particularly in the context of clinical trials, has also been underscored by international clinician consensus [15]. While previously published reviews have quantitatively synthesized the relationship between liver fibrosis and mortality, this has not yet been explored for key outcomes such as cirrhosis, hepatic decompensation or HCC [2,16–18]. The objective of this systematic literature review (SLR) was to synthesize estimates of the relationship between histologic liver fibrosis stage and a variety of relevant clinical outcomes in a population of adults with MASLD and/or MASH, as reported in clinical trials and observational studies.

Materials & methods

SLR & meta-analysis inclusion criteria

The review adheres to methodological standards for conducting SLRs and meta-analyses, including Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) [19], MOOSE Reporting Guidelines [20] and AMSTAR 2 guidelines [21]. The protocol for this review was not registered.

Investigators implemented the literature search on 13 November 2024 in Ovid MEDLINE® and Embase (Supplementary Table 1), with supplementary searches of grey literature sources (ClinicalTrials.gov, World Health Organization International Clinical Trials Registry Platform) and reference lists of relevant published reviews. Clinical trial registry searches were conducted using "NASH or MASH" in the "Title" or "Condition" fields.

Within the study search window, the nomenclature was officially changed from nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH) to MASLD and MASH [22]. To address this, both sets of terms were included in the search strategies. Previous work has demonstrated that the legacy NAFLD/NASH definitions and the new MASLD/MASH definitions overlap in approximately 99% of cases [23]; thus these terms were treated interchangeably in this study's analysis. If both sets of nomenclature were included in a given study, the newer terms were favored.

Studies of interest were identified according to PI/ECOS (Population, Intervention/Exposure, Comparator, Outcomes and Study design) criteria. Eligible study populations included adult patients with a diagnosis of MASH or MASLD (≥1% of patients had to have a MASH diagnosis; "MASH ± MASLD") and reported risk

of outcomes stratified by histologic liver fibrosis stage at least at baseline (follow-up timepoints were included if reported). Studies conducted in patients with MASLD only were excluded from data synthesis. Outcomes of interest included progression to cirrhosis (compensated or decompensated), hepatic decompensation events (ascites, variceal bleeding, hepatic encephalopathy, clinically significant portal hypertension), HCC, liver transplant, end-stage liver disease, composite outcomes of the above liver-related events (i.e., "hepatic decompensation and/or HCC", "major adverse liver outcome [MALO]"), all-cause mortality, liver-specific mortality, CV-related mortality, major adverse cardiovascular events (MACE)/CV-related events and health related quality of life (HRQoL). Outcome definitions are listed in Supplementary Table 2. Eligible study designs were primary studies, either interventional trials or observational longitudinal studies that were published in English. Previously published literature reviews and meta-analyses were excluded from data synthesis, but their reference lists were reviewed to identify additional relevant primary studies.

All studies that were identified by the search strategy were screened by two researchers independently according to the PI/ECOS criteria. Discrepancies between reviewers were discussed, and unresolved conflicts were resolved by a third reviewer. Reasons for exclusion were documented at all stages.

Study characteristics and variables of interest were extracted by a single reviewer and verified by a second reviewer for all studies meeting the inclusion criteria. Extracted data were stored and managed in Microsoft® Excel®. A structured quality control process was applied to every study to ensure completeness, accuracy, and consistency. Data from publications describing the same underlying patient population were mapped in order to avoid double-counting of data contributing to meta-analysis and narrative syntheses. When required, authors were contacted for additional information.

Risk of bias (RoB) assessments were performed with the Cochrane RoB 2 tool for randomized controlled trials (RCTs), and with the ROBINS-I V2 tool for observational studies, by a single reviewer. The certainty of evidence was not formally assessed (i.e., Grading of Recommendations Assessment, Development and Evaluation [GRADE]).

Meta-analysis

Various fibrosis staging systems were used across included studies, which were converted to the Kleiner/Brunt system (F0–F4) scheme (Supplementary Table 3). In the remainder of this article, all F-stages are reported according to the Kleiner/Brunt staging scheme.

Adjusted HRs (aHRs) and their associated measures of variability were extracted from eligible studies; if they were not reported, unadjusted HRs (uHRs) were instead utilized. Where possible, HRs reported in papers were used to calculate additional comparisons (see Supplementary Methods for details).

In studies where a numeric HR value was not reported, but a Kaplan–Meier (KM) curve was presented, pseudo individual patient data (IPD) was generated and used to calculate the HR value using the method described by Guyot *et al.* [24]. Cox proportional hazards models were then fitted to the reconstructed IPD to derive log HRs and standard errors (SE). KM curve data were extracted using DigitizeIt, and further calculations were performed using R Statistical Software.

For outcomes with a comparison of interest reported in at least three studies, models were fit to calculate a pooled HR, with a 95% CI and a heterogeneity measure (I^2). The model selection process was as follows: if fewer than five HRs could be pooled, then a fixed-effects (FE) model was selected. If five or more HRs could be pooled, then random-effects (RE) models were also fit, using both the Hartung–Knapp–Sidik–Jonkman (HKSJ) method with a Paule–Mendel (PM) estimator [25–27] and DerSimonian-Laird (DL) method [28]. The model that was valid and meaningful was selected, i.e., the model that yielded an informative pooled estimate (CI lies within the union of individual study CIs) that was statistically congruent with the DL model (same significance conclusion) was deemed as the most appropriate meta-analytical model. Pooled HRs were calculated such that the lower F-stage was the reference. In other words, an HR >1 indicates higher risk of outcome in the higher F-stage group compared with the lower F-stage group. All calculations were performed using R Statistical Software (v4.4.2) with "IPDfromKM", "survival", "survminer", "dplyr" and "ggplot2" packages.

Narrative synthesis

Data reported as incidence rates of outcomes of interest, stratified by fibrosis stage, were described narratively, without formal quantitative analysis. These descriptions focused on trends across studies and were included as supportive evidence to the meta-analysis results. For studies that did not report HRs suitable for meta-analysis or

10.57264/cer-2026-0148

incidence rates stratified by fibrosis stage, other measures such as means or patient counts were used to identify trends.

Sensitivity analyses

Sensitivity analyses were conducted to assess the appropriateness of model choice. For models that included more than five effect estimates, a leave-one-out analysis was performed by repeating the meta-analysis and excluding one study at a time, allowing the identification of studies that had greater influence over the pooled effect estimate. Additionally, for models that included more than five or more effect estimates, the following sensitivity analyses were conducted as feasible given the characteristics of included studies: including only studies with at least 5 years of follow-up, excluding studies with serious/high RoB, including only studies that used the same fibrosis staging system, including only observational studies, and excluding pseudo-IPD generated from KM curves.

An additional sensitivity analysis to investigate the consistency of results in MASH patients specifically was conducted by restricting to risk estimates from populations of 80–100% MASH patients ("MASH-majority"). This was conducted to investigate a population more closely related to the populations currently indicated for FDA approved treatments.

Results

Evidence base

Of the 2811 records identified by the search strategy, 279 records were retained for review of full-text articles (Figure 1). With the addition of nine records from hand searching and clinical trial registry searching, ultimately, 32 unique primary studies (with data from these studies reported across 39 publications) were included, which included a range of 17.6–100% of patients with MASH. Of these, there were 10 studies (reported in 14 publications) conducted in MASH-majority populations. The full list of included publications is in Supplementary Table 4.

Among the included studies, there were three RCTs (two pooled analyses and one stand-alone trial) and 29 observational studies (Table 1). Most studies were assessed as having moderate RoB (n = 22). Results of RoB assessments are summarized in Supplementary Figure 1. Nineteen studies were eligible for meta-analysis (Figure 2).

Cirrhosis

Three studies (one clinical trial and two retrospective cohorts), enrolling 66.9–100% patients with MASH, reported sufficient data for fixed effects meta-analysis of aHRs comparing patients with F3 and F2 fibrosis [29–31]. Cirrhosis data from the STELLAR-3/-4 trials were reported in a conference abstract [30], limiting the description of the outcome definition. Patients with F3 fibrosis had an increased risk of progression to cirrhosis, compared with those with F2 (pooled aHR: 2.05 [95% CI: 1.45, 2.90]; Table 2). Of note, the aHR reported by Younossi *et al.* [31] was over tenfold larger than the next highest aHR estimate. This is likely because Younossi *et al.* [31] defined cirrhosis based on histologic progression only, whereas Fujii *et al.* [29] and Goh *et al.* [30] estimated aHRs for decompensated cirrhosis (Supplementary Table 2). The high heterogeneity observed between studies (I^2 = 77.6%) may be explained in part by this variation in outcome definition.

The incidence rate per 1000 person-years of cirrhosis, reported in one study, was 1.0 in F2 fibrosis and 44 in F3 fibrosis (Figure 3A) [31].

Hepatic decompensation

Three studies (two retrospective cohorts and one prospective cohort), enrolling 17.6–100% patients with MASH, reported sufficient data for fixed effects meta-analysis of aHRs comparing patients with F3–4 and F0–2 fibrosis [31–33]. Patients with F3–4 fibrosis had a significantly increased risk of hepatic decompensation compared with those with F0–2 fibrosis, with a pooled aHR of 10.93 (95% CI: 6.31, 18.92) and individual aHR >7 among the included studies (Table 2). Variability in the specific hepatic decompensation events observed in each study varied (Supplementary Table 2), which, in combination with differences in the adjustment factors among the three contributing aHRs, may partly explain the moderate statistical heterogeneity (I^2 = 47.8%) detected.

The incidence rate per 1000 person-years of hepatic decompensation, reported in two studies, was 4 in F2 fibrosis, 10 to 24 in F3 fibrosis and 27 to 71 in F4 fibrosis (Figure 3B) [31,33].

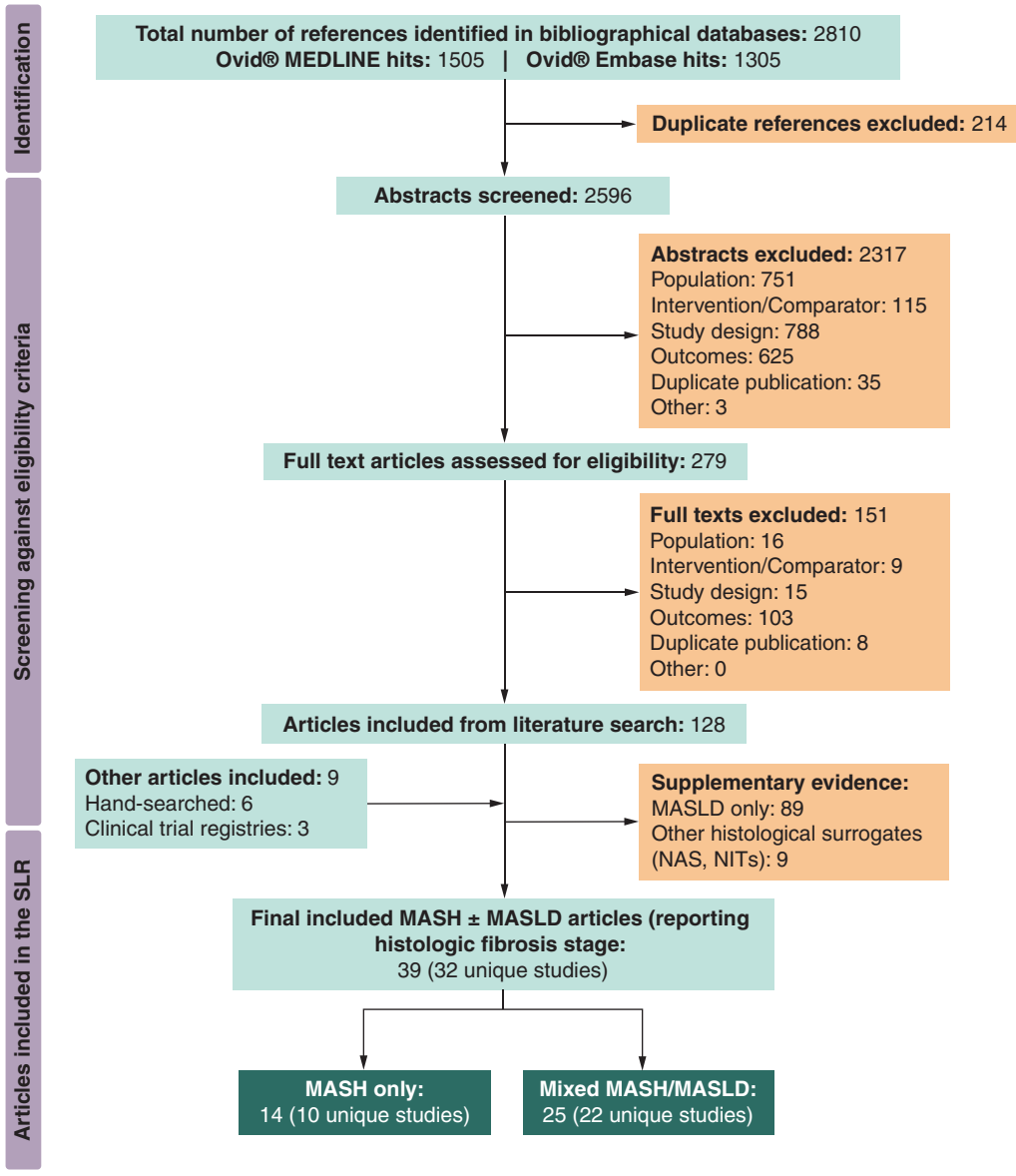


Figure 1. PRISMA diagram of study selection.
MASH: Metabolic dysfunction-associated steatohepatitis; MASLD: Metabolic dysfunction-associated steatotic liver disease; NAFLD: Nonalcoholic fatty liver disease; NAS: NAFLD Activity Score; NIT: Noninvasive test; PRISMA: Preferred Items for Systematic reviews and Meta-Analyses; SLR: Systematic literature review.

Hepatic decompensation &/or HCC

Three retrospective cohort studies, enrolling 45.9–74.9% patients with MASH, reported sufficient data for fixed effects meta-analysis of aHRs comparing patients with F3–4 and F0–2 fibrosis [29,34,35]. Meta-analysis of unadjusted estimates suggested a 30-fold increase in risk of hepatic decompensation and/or HCC in patients with F3–4 fibrosis compared with those with mild-moderate fibrosis (F0–2) (pooled uHR 32.37, 95% CI: 13.51, 77.59; I^2 = 60.6%; Table 2). Pooling aHRs from these same studies attenuated the magnitude of effect, but the association remained consistent (pooled aHR: 20.50, 95% CI: 8.25, 50.93, I^2 = 59.9%). The substantial heterogeneity detected may be due to study-level differences in the specific hepatic decompensation events observed across studies and variation in the adjustment factors among the three contributing aHRs (Supplementary Table 2). Despite these differences, HRs reported in individual studies were consistent with the pooled HR in terms of direction and statistical significance.

10.57264/cer-2026-0148

10.57264/cer-2026-0148

Table 1. Characteristics of included studies.

Study name/author, year	Sample size/population composition/Country(ies)	Study design/year(s) of study conduct	Baseline fibrosis stage, n (%)†	Comorbidities, %	Outcomes reported	Risk of bias	Included in the meta-analysis
Clinical trials							
GS-US-321-0105 (NCT01672866) and GS -US-321-0106 (NCT01672879) Sanyal, 2019; Sanyal, 2016	n = 475 NASH 100% North America and Europe (9 countries)	Phase IIb, placebo-controlled, double-blind 2012–2017	Bridging fibrosis: F2: 124 (37) F3: 93 (43) Compensated cirrhosis: F3: 86 (33) F4: 171 (66)	N/R	All-cause mortality; cirrhosis; MALO (composite); end-stage liver disease	Low	Yes (Sanyal 2019)
REGENERATE (NCT02548351) Younossi, 2022	n = 1218 NASH 100% North America, Europe, Australia and Israel (21 countries)	Phase III, placebo-controlled, double-blind 2015–2023	F1: 96 (24) F2: 142 (35) F3:169 (42)	N/R	HRQoL	Low	No
STELLAR-3 (NCT03053050) and STELLAR-4 (NCT03053063) Goh, 2020; Iyer, 2020; Noureddin, 2020; Younossi, 2019	n = 1679 NASH 100% North America, South America, Europe, Asia, Australia and New Zealand (27 countries)	Phase III, placebo-controlled, double-blind 2017–2019	F3: 725 (47.7) F4: 794 (52.3)	N/R	Cirrhosis; Hepatic decompensation (composite); MALO (composite); MACE or CV-related events (composite); HRQoL	Low	Yes (Goh 2020)
Observational studies							
Akbari, 2024; Akbari 2023	n = 1260 NASH 62.7% Sweden	Retrospective cohort 1974–2020	F0: 222 (24.6) F1: 372 (41.2) F2: 210 (23.3) F3: 100 (11)	N/R	MALO (composite)	Moderate	Yes (Akbari 2024)
Akuta, 2020	n = 441 NASH 90.2% Japan	Retrospective cohort 1976–2019	F0: 50 (11.3) F1: 182 (41.3) F2: 68 (15.4) F3: 108 (24.5) F4: 33 (7.5)	Hyperlipidemia: 34.0	Overall survival	Critical	No
Ampuero, 2021	n = 1893 NASH 46.7% Spain	Prospective cohort N/R-2020	F0: 649 (34.3) F1: 511 (27) F2: 312 (16.5) F3: 284 (15) F4: 137 (7.2)	Obesity: 61.5	All-cause mortality; Hepatic decompensation (composite); Cirrhosis	Serious	No
Ampuero, 2019	n = 568 NASH 46.7% Spain	Prospective cohort 2015–2017	F0: 208 (36.6) F1: 152 (26.8) F2: 104 (18.3) F3: 63 (11.1) F4: 41 (7.2)	Obesity: 56.7 Dyslipidemia: 32.4	All-cause mortality; CV events	Moderate	No
Angulo, 2015	n = 619 NASH 45.9% North America, Europe, Asia and Australia (6 countries)	Retrospective cohort 1975–2012	F0: 322 (52) F1: 141 (22.8) F2: 85 (13.7) F3: 53 (8.6) F4: 18 (2.9)	N/R	Hepatic decompensation and/or HCC (composite); MALO (composite)	Moderate	Yes

†Fibrosis stages converted according to Supplementary Table 3.

CV: Cardiovascular; HCC: Hepatocellular carcinoma; HRQoL: Health related quality of life; MACE: Major adverse cardiovascular event; MALO: Major adverse liver outcomes; MASH: Metabolic dysfunction-associated steatotic hepatitis; N/R: Not reported.

Becaris Publishing

Table 1. Characteristics of included studies (cont.).

Observational studies							
Buzzetti, 2019	n = 437 NASH 39% Europe (3 countries)	Retrospective cohort N/R-2016	F0: 233 (53) F1: 95 (22) F2: 37 (8) F3: 34 (8) F4: 38 (9)	Dyslipidemia: 60.2	MALO (composite)	Serious	Yes
NASH CRN Database 2 study Chalasani, 2020	n = 1010 NASH 58% USA	Prospective cohort 2014–2019	F0: 254 (25.1) F1: 253 (25.0) F2: 185 (18.3) F3: 231 (22.9) F4: 85 (8.4)	N/R	Cirrhosis	Moderate	No
Chayama, 2023	n = 393 NASH 100% Japan	Prospective cohort 2016–2018	N/R	Dyslipidemia: 47.1	Cirrhosis, compensated; Cirrhosis, decompensated; HCC; Liver transplant; MALO (composite); MACE or CV-related events (composite)	Moderate	No
De Roza, 2021	n = 267 NASH 100% New Zealand	Prospective cohort 2005–2017	F0: 53 (20.3) F1: 95 (36.3) F2: 31 (11.9) F3: 48 (18.4) F4: 34 (13.1)	N/R	All-cause mortality; HCC; Hepatic decompensation (composite)	Moderate	No
CLIONE study Fujii, 2023	n = 1398 NASH 66.9% Japan	Retrospective cohort 1994–2021	F0: 241 (17.2) F1: 539 (38.6) F2: 394 (28.2) F3: 198 (14.2) F4: 26 (1.9)	Dyslipidemia: 57.5	All-cause mortality; Liver-specific mortality; Cirrhosis, decompensated; HCC; Hepatic decompensation and/or HCC (composite)	Moderate	Yes
Grimaudo, 2020	n = 471 NASH 76.2% Italy	Prospective cohort 2004–2018	F3–4: 162 (34.4)	Obesity: 48.5	All-cause mortality; Liver-specific mortality; Hepatic decompensation (composite); HCC; MACE of CV-related events (composite)	Serious	Yes
Hagström, 2019; Hagström, 2018a; Hagström, 2017	n = 646 NASH 66.4% Sweden	Retrospective cohort 1971–2009	F0: 164 (25.4) F1: 256 (23.6) F2: 149 (23.1) F3: 58 (9.0) F4: 20 (3.1)	Obesity: 29.1 Hyperlipidemia: 8.8	All-cause mortality; Liver-specific mortality; Hepatic decompensation (composite); MALO (composite)	Moderate	Yes (Hagström 2017)
Hagström, 2018b	n = 60 NASH 100% Sweden	Retrospective cohort 2009–2016	F0: 15 (25) F1: 22 (37) F2: 14 (23) F3: 6 (10) F4: 3 (5)	N/R	All-cause mortality	Moderate	No

†Fibrosis stages converted according to Supplementary Table 3.

CV: Cardiovascular; HCC: Hepatocellular carcinoma; HRQoL: Health related quality of life; MACE: Major adverse cardiovascular event; MALO: Major adverse liver outcomes; MASH: Metabolic dysfunction-associated steatotic hepatitis; N/R: Not reported.

10.57264/cer-2026-0148

Table 1. Characteristics of included studies (cont.).

Observational studies							
Henson, 2020	n = 285 NASH 82.5% USA	Prospective cohort 2011–2018	F0: 86 (30.2) F1: 91 (31.9) F2: 52 (18.2) F3: 30 (10.5) F4: 26 (9.1)	Dyslipidemia: 52.6	MACE or CV related events (composite)	Serious	No
Hirose, 2020	n = 223 NASH 74.9% Japan	Retrospective cohort 1975–2012	F3–4: 36 (16.1)	Overweight: 54.3 Dyslipidemia: 7.2	All-cause mortality; Hepatic decompensation and/or HCC (composite)	Moderate	Yes
Hirose, 2018	n = 233 NASH 76% Japan	Retrospective cohort 1975–2012	N/R	N/R	All-cause mortality; Hepatic decompensation and/or HCC (composite)	Moderate	No
Ito, 2019	n = 246 NASH 63.4% Japan	Retrospective cohort 1999–2014	F0: 61 (24.8) F1: 97 (39.4) F2: 38 (15.4) F3: 42 (17.1) F4: 8 (3.3)	Dyslipidemia: 71.4	All-cause mortality; Hepatic decompensation (composite); HCC	Moderate	Yes
Kamarajah, 2018	n = 113 NASH 78% Malaysia	Prospective cohort 2012–2017	F0: 35 (51) F1: 45 (40) F2: 8 (7) F3: 23 (20) F4: 2 (2)	Obesity: 86 Dyslipidemia: 90	Hepatic decompensation and/or HCC (composite)	Serious	No
Kumazaki, 2024	n = 518 MASH 54.8% Japan	Retrospective cohort 2002–2021	F0: 106 (20.5) F1: 160 (30.9) F2: 116 (22.4) F3: 84 (16.2) F4: 52 (10.0)	N/R	HCC	Moderate	Yes
Nasr, 2018	n = 129 NASH NR% Sweden	Prospective cohort 1988–2015	F0: 60 (46.5) F1: 31 (24) F2: 22 (17.1) F3: 12 (9.3) F4: 4 (3.1)	Obesity: 29	Hepatic decompensation and/or HCC (composite)	Critical	No
Peleg, 2018	n = 153 NASH 17.6% Israel	Prospective cohort 2005–2012	F0–2: 121 (79.1) F3–4: 32 (20.9)	N/R	All-cause mortality; Hepatic decompensation (composite)	Moderate	Yes
Pennisi, 2023	n = 1938 MASH 39.6% North America, Europe, Asia and Australia (7 countries)	Retrospective cohort N/R	F0–F1: 872 (45.0) F2–F4: 1066 (55.0)	Obesity: 46.1	Hepatic decompensation (composite); HCC; Hepatic decompensation and/or HCC (composite); CV-related events (composite)	Moderate	No

†Fibrosis stages converted according to Supplementary Table 3.

CV: Cardiovascular; HCC: Hepatocellular carcinoma; HRQoL: Health related quality of life; MACE: Major adverse cardiovascular event; MALO: Major adverse liver outcomes; MASH: Metabolic dysfunction-associated steatotic hepatitis; N/R: Not reported.

Table 1. Characteristics of included studies (cont.).

Observational studies							
NAFLD Adult Database 2 (NCT01030484) Sanyal, 2021	n = 1773 NASH 75.5% USA	Prospective cohort 2009–2019	F0–F2: 1237 (69.8) F3: 369 (20.8) F4: 167 (9.4)	Obesity: 33	All-cause mortality; Liver-specific mortality; Hepatic decompensation (composite); Ascites; Variceal bleeding; Hepatic encephalopathy; HCC; End-stage liver disease; MACE or CV-related events (composite)	Moderate	Yes
Sebastiani, 2015	n = 148 NASH 100% Canada	Retrospective cohort 2004–2014	F0: 23 (15.5) F1: 53 (35.8) F2: 22 (14.9) F3: 28 (18.9) F4: 22 (14.9)	Obesity: 56.1	MALO (composite)	Moderate	Yes
Seko, 2017	n = 238 NASH 70.2% Japan	Prospective cohort 1999–2014	*Among MASH:* F0: 2 (1.2) F1: 85 (50.9) F2: 38 (22.8) F3: 27 (16.2) F4: 15 (9.0)	N/R	HCC	Moderate	Yes
Seko, 2014	n = 312 NASH 56.4% Japan	Retrospective cohort 1999–2013	*Among MASH:* F1: 89 (51) F2: 45 (26) F3: 27 (15) F4: 15 (8)	N/R	HCC	Moderate	Yes
Tsutsumi, 2024	n = 1349 MASH 65.6% Japan	Retrospective cohort 1974–2020	F0: 239 (17.7) F1: 493 (36.5) F2: 383 (28.4) F3: 211 (15.6) F4: 23 (1.7)	Obesity: 87.6 Dyslipidemia: 78.9	All-cause mortality	Moderate	Yes
Younes, 2022	n = 1339 NASH 68.4% Europe and Australia (4 countries)	Retrospective cohort 1990–2016	F0: 359 (26.8) F1–2: 674 (50.3) F3–4: 306 (22.9)	N/R	All-cause mortality	Moderate	Yes
Younossi, 2024	n = 702 MASH 100% USA	Retrospective cohort 1984–2021	F0–1: 400 (57.0) F2: 123 (17.5) F3: 128 (18.2) F4: 51 (7.3)	Obesity: 80 Dyslipidamia: 54	Cirrhosis, End-stage liver disease, HCC, Hepatic decompensation (composite), Hepatic decompensation and/or HCC (composite), Liver transplant, MACE or CV-related events (composite)	Moderate	Yes

†Fibrosis stages converted according to Supplementary Table 3.

CV: Cardiovascular; HCC: Hepatocellular carcinoma; HRQoL: Health related quality of life; MACE: Major adverse cardiovascular event; MALO: Major adverse liver outcomes; MASH: Metabolic dysfunction-associated steatotic hepatitis; N/R: Not reported.

10.57264/cer-2026-0148

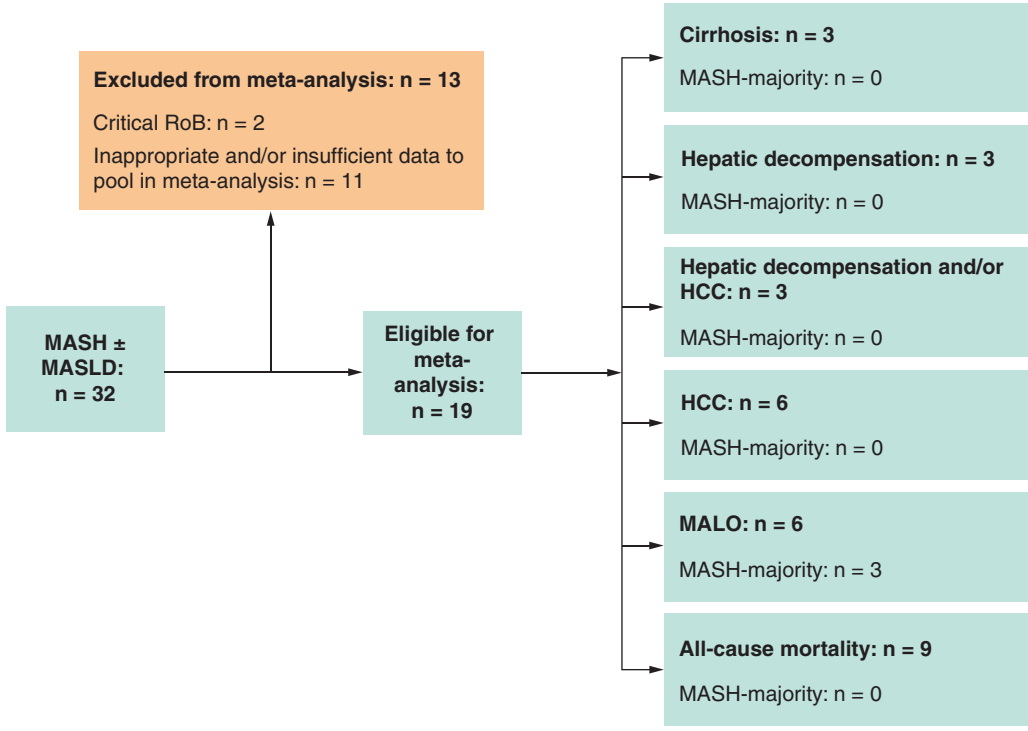


Figure 2. Eligibility of studies for meta-analysis.
HCC: Hepatocellular carcinoma; MALO: Major adverse liver outcome; MASH: Metabolic dysfunction-associated steatohepatitis; MASLD: Metabolic dysfunction-associated steatotic liver disease; RoB: Risk of bias.

Hepatocellular carcinoma

Three studies (two retrospective cohorts and one prospective cohort), enrolling 66.9–100% patients with MASH, reported sufficient data for fixed effects meta-analysis of aHRs comparing patients with F4 and F3 fibrosis [29,31,33]. All three aHRs were estimated by indirect calculation from directly reported aHRs. Patients with F4 fibrosis showed an uncertain risk of HCC compared with those with F3 fibrosis; while the pooled aHR was 1.52 (95% CI: 0.33, 7.10), individual estimates varied (0.52 [95% CI: 0.03, 9.05], 0.53 [95% CI: 0.02, 12.53] and 5.00 [95% CI: 0.53, 46.91]). Statistical heterogeneity (I^2 = 2.05%) was low.

Five studies (three retrospective [29,36,37] and two prospective [33,38] cohorts), enrolling 49–70% patients with MASH, were included in an RE meta-analysis (PM-HKSJ model) of aHRs comparing patients with F3–4 and F0–2 fibrosis (see Supplementary Table 5 for model selection). The resulting pooled aHR was 5.57 (95% CI: 1.87, 16.60), indicating an increased risk of HCC in patients with advanced fibrosis (F3–4) compared with mild-moderate fibrosis (F0–2; Table 2). The heterogeneity was low (I^2 = 17.0%), suggesting limited between-study variability. Variation in the pooled aHR was observed in the leave-one-out sensitivity analysis, with estimates ranging from 3.64 to 10.72 (Supplementary Table 6). The study by Kumazaki *et al.* [36] yielded the most precise individual estimate and had the largest influence on the pooled result. The variability in pooled estimates indicates that results may be sensitive to individual study inclusion, perhaps due to differences in covariate adjustments. Only two of the five studies contributing to the HCC meta-analysis reported more than 5 years of follow-up. Moreover, the median follow-up times among these studies lacked the variation for this sensitivity analysis to be meaningful (range: 4.0 to 6.1 years). As such, stratified or sensitivity analyses by follow-up duration were not feasible for this outcome. Sensitivity analysis based on RoB was not conducted, since all included studies were assessed as having a moderate RoB.

The incidence rate per 1,000 person-years of HCC, reported in three studies, was 2.2 in F1 fibrosis, 4.5 in F2 fibrosis, 3 to 14.2 in F3 fibrosis and 1 to 19 in F4 fibrosis (Figure 3C) [29,31,33].

Table 2. Meta-analysis results for eligible metabolic dysfunction-associated steatotic hepatitis ± metabolic dysfunction-associated steatotic liver disease studies.

Outcome	F-stage comparisons	Adjusted estimates				Unadjusted estimates			
		Individual study estimates, HR (95% CI)	Final model selected (number of studies)	Pooled aHR (95% CI)	I^2, %	Individual study estimates, HR (95% CI)	Final model selected (number of studies)	Pooled uHR (95% CI)	I^2, %
Cirrhosis, decompensated	F3 vs F2	Fujii, 2023: 1.23 (0.04, 39.63)	Fixed effects (n = 3)	2.05 (1.45, 2.90)‡,§	77.6	–	–	–	–
		Goh, 2020†: 1.84 (1.29, 2.63)							
		Younossi, 2024: 25.00 (4.64, 134.63)							
Hepatic decompensation	F3–4 vs F0–2	Peleg, 2018: 11.10 (4.65, 26.47)	Fixed effects (n = 3)	10.93 (6.31, 18.92)‡	47.8	–	–	–	–
		Sanyal, 2021: 36.10 (8.90, 146.37)							
		Younossi, 2024: 7.14 (3.14, 16.24)							
Hepatic decompensation/HCC	F3–4 vs F0–2	Angulo, 2015: 47.46 (11.94, 188.63)	Fixed effects (n = 3)	20.50 (8.25, 50.93)‡	59.9	Angulo, 2015: 52.89 (13.31, 210.16)	Fixed effects (n = 3)	32.37 (13.51, 77.59)‡	60.6
		Fujii, 2023: 51.30 (5.18, 508.05)				Fujii, 2023: 146.90 (18.31, 1178.65)			
		Hirose, 2020: 5.87 (1.41, 24.42)				Hirose, 2020: 10.80 (2.81, 41.46)			
HCC	F4 vs F3	Fujii, 2023: 0.52 (0.03, 9.05)	Fixed effects (n = 3)	1.52 (0.33, 7.10)§	2.7	–	–	–	–
		Sanyal, 2021: 0.53 (0.02, 12.53)							
		Younossi, 2024: 5.00 (0.53, 46.91)							
	F3–4 vs F0–2	Fujii, 2023: 5.11 (0.34, 76.70)	PM-HKSJ (n = 5)	5.57 (1.87, 16.60)‡	17	–	–	–	–
		Kumazaki, 2024: 2.71 (1.16, 6.35)							
		Sanyal, 2021: 4.90 (0.39, 61.59)							
		Seko, 2014: 12.30 (3.52, 43.04)							
		Seko, 2017: 24.40 (2.08, 285.86)							
MALO	F4 vs F3	Angulo, 2015: 3.34 (1.49, 7.49)	Fixed effects (n = 3)	3.27 (1.98, 5.41)‡,§	45.9	Angulo, 2015: 2.90 (1.38, 6.08)	Fixed effects (n = 5)	2.43 (1.70, 3.46)‡,§	64.6
		Hagstrom, 2017: 15.21 (2.76, 83.78)				Goh, 2020†: 7.76 (1.83, 32.87)			
						Hagstrom, 2017: 7.19 (1.87, 27.66)			
		Younossi, 2024: 2.50 (1.25, 5.00)				Sanyal, 2019: 1.20 (0.67, 2.16)			
						Younossi, 2024: 3.12 (1.59, 6.15)			

†Data acquired from nonpeer reviewed source (conference abstract/poster).
‡$p < 0.05$.
§At least one estimate was derived from an indirect calculation.
¶At least one estimate was derived from pseudo-IPD generated from a KM curve.
aHR: Adjusted hazard ratio; HCC: Hepatocellular carcinoma; IPD: Individual patient data; KM: Kaplan–Meier; MALO: Major adverse liver outcome; PM-HKSJ: Paule–Mendel-Hartung–Knapp–Sidik–Jonkman.

10.57264/cer-2026-0148

Table 2. Meta-analysis results for eligible metabolic dysfunction-associated steatotic hepatitis ± metabolic dysfunction-associated steatotic liver disease studies (cont.).

Outcome	F-stage comparisons	Adjusted estimates				Unadjusted estimates			
		Individual study estimates, HR (95% CI)	Final model selected (number of studies)	Pooled aHR (95% CI)	I^2, %	Individual study estimates, HR (95% CI)	Final model selected (number of studies)	Pooled uHR (95% CI)	I^2, %
	F4 vs F2	Angulo, 2015: 3.32 (1.51, 7.31)	Fixed effects (n = 3)	6.31 (3.38, 11.79)‡,§	72.3	Angulo, 2015: 3.77 (1.85, 7.69)	Fixed effects (n = 3)	6.62 (3.72, 11.78)‡,§	71.2
		Hagstrom, 2017: 27.94 (5.49, 142.23)				Hagstrom, 2017: 18.66 (4.89, 71.20)			
		Younossi, 2024: 14.29 (3.82, 53.45)				Younossi, 2024: 20.00 (4.71, 84.86)			
	F3–4 vs F0–2	Akbari, 2024: 8.90 (4.58, 17.31)	PM-HKSJ (n = 6)	6.66 (4.04, 10.99)‡,§	34.1	–	–	–	–
		Angulo, 2015: 6.35 (3.35, 12.04)							
		Buzzetti, 2019: 15.40 (4.99, 47.54)							
		Hagstrom, 2017 5.37 (2.73, 10.56)							
		Sebastiani, 2015: 3.10 (1.39, 6.93)							
		Younossi, 2024: 8.33 (3.73, 18.63)							
All-cause mortality	F4 vs F3	Fujii, 2023: 2.47 (0.27, 22.91)	Fixed effects (n = 3)	2.13 (1.10, 4.10)‡,§	0	Fuji, 2023: 2.79 (0.46, 16.96)	Fixed effects (n = 4)	2.30 (1.35, 3.94)‡,§,¶	0
		Hagstrom, 2017: 2.13 (0.86, 5.29)				Hagstrom, 2017: 2.15 (1.01, 4.59)			
		Sanyal, 2021: 2.05 (0.72, 5.83)				Sanyal, 2021: 2.00 (0.75, 3.21)			
						Tsutsumi, 2024: 4.06 (0.78, 21.00)			
	F4 vs F0–2	–	–	–	–	Fujii, 2023: 5.95 (1.42, 24.92)	Fixed effects (n = 3)	5.63 (3.15–10.06)‡,§,¶	0
						Sanyal, 2021: 5.60 (2.76, 11.35)			
						Tsutsumi, 2024: 5.44 (1.28, 23.16)			
	F3–4 vs F0–2	Fujii, 2023: 2.96 (0.52, 16.88)	PM-HKSJ (n = 9)	3.36 (2.37–4.75)‡	0	–	–	–	–
		Grimaudo, 2020: 4.10 (0.77, 21.84)							
		Hagstrom, 2017: 3.80 (1.84, 7.86)							
		Hirose, 2020: 2.39 (0.71, 8.02)							
		Ito, 2019: 6.51 (1.43, 29.59)							
		Peleg, 2018: 5.54 (1.82, 16.89)							
		Sanyal, 2021: 3.90 (1.81, 8.43)							
		Tsutsumi, 2024: 1.80 (0.89, 3.65)							
		Younes, 2022: 7.40 (1.31, 41.71)							

†Data acquired from nonpeer reviewed source (conference abstract/poster).
‡$p < 0.05$.
§At least one estimate was derived from an indirect calculation.
¶At least one estimate was derived from pseudo-IPD generated from a KM curve.
aHR: Adjusted hazard ratio; HCC: Hepatocellular carcinoma; IPD: Individual patient data; KM: Kaplan–Meier; MALO: Major adverse liver outcome; PM-HKSJ: Paule–Mendel-Hartung–Knapp–Sidik–Jonkman.

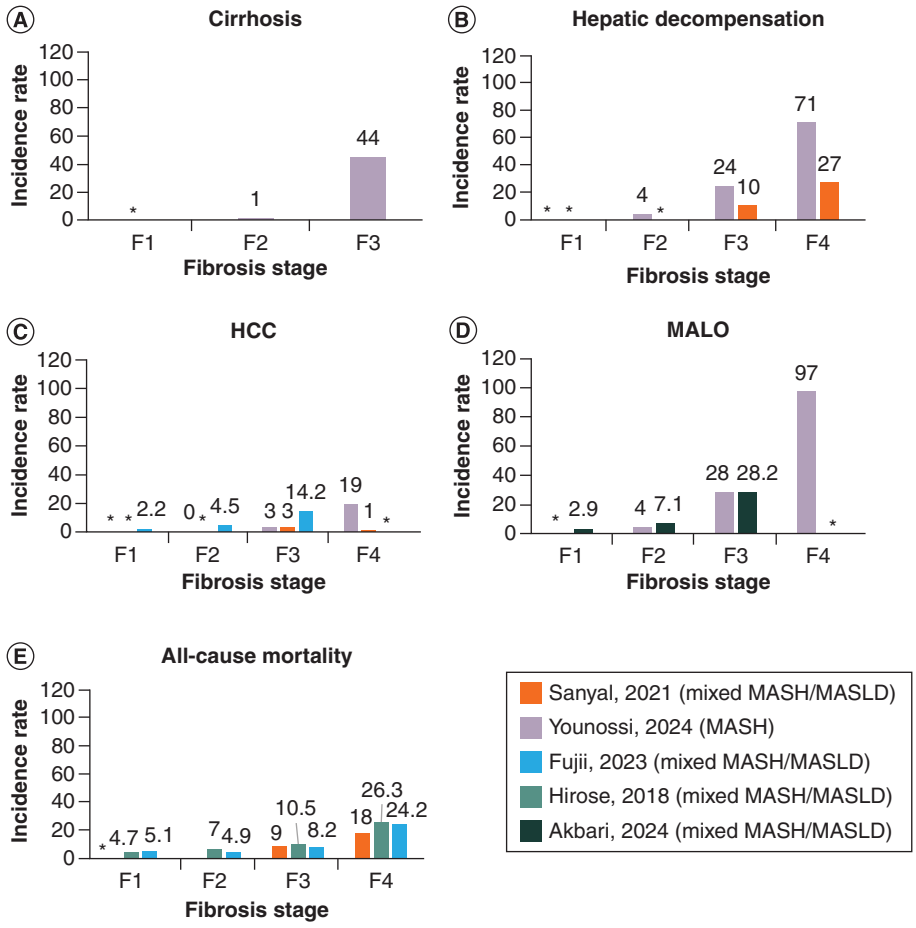


Figure 3. Incidence rates of clinical outcomes by fibrosis stage in MASH with or without MASLD patients. Published incidence rates (per 1000 patient-years) by fibrosis stage in MASH ± MASLD patients of **(A)** cirrhosis, **(B)** hepatic decompensation, **(C)** HCC, **(D)** MALO and **(E)** all-cause mortality.
*Incidence not reported at this stage.
HCC: Hepatocellular carcinoma; MALO: Major adverse liver outcome; MASH: Metabolic dysfunction-associated steatohepatitis; MASLD: Metabolic dysfunction-associated steatotic liver disease.

Major adverse liver outcomes

Overall, five studies enrolling 39–100% patients with MASH reported an association between histologic fibrosis stage and MALO. Three retrospective cohort studies enrolling 46–100% patients with MASH [31,34,39], and two RCTs enrolling 100% patients with MASH [30,40] were included in a FE meta-analysis of uHRs comparing patients with F4 and F3 fibrosis. There was a greater than twofold increased risk of MALO in patients with F4 versus F3 fibrosis (pooled uHR of 2.43, 95% CI: 1.70, 3.46; I^2 = 64.4%; Table 2). Overall, leave-one-out influencer sensitivity analysis showed consistency in pooled uHRs across all iterations, ranging from 2.6 to 3.63; however, in most instances, the pooled estimate did not reach statistical significance. Of these five studies reporting HRs comparing risk of MALO in F4 and F3 fibrosis, three retrospective cohorts reported aHRs [29,33,39]. Based on FE meta-analysis, patients with F4 fibrosis had a significantly increased risk of MALO compared with those with F3 fibrosis (pooled aHR 3.27, 95% CI: 1.98, 5.41; I^2 = 45.9%).

Meta-analysis of other F-stage comparisons showed consistency in terms of direction and significance of the effect while pooling uHRs and aHRs, showing increased risk of MALO among patients with higher fibrosis stages. A FE meta-analysis of three retrospective cohort studies (46–100% patients with MASH) showed a sixfold increase in risk of MALO in patients with F4 versus F2 fibrosis (pooled uHR: 6.62, 95% CI: 3.72, 11.78; I^2 = 71.22%; pooled aHR 6.31, 95% CI: 3.38, 11.79; I^2 = 72.33%; Table 2) [31,34,39]. Based on a RE meta-analysis (PM-HKSJ model) of six retrospective cohort studies reporting aHRs patients with advanced fibrosis (F3–4) had a significantly increased risk of MALO compared with those with mild-moderate fibrosis (F0–2) with a pooled aHR

10.57264/cer-2026-0148

of 6.66 (95% CI: 4.04, 10.99) and low-to-moderate statistical heterogeneity (I^2 = 34.13%; Table 2; Supplementary Table 5) [31,34,39,41–43]. When limiting analysis to studies judged to be at moderate or low RoB, the effect estimate was directionally consistent and statistically significant, albeit slightly attenuated (pooled aHR: 6.14, 95% CI: 3.76, 10.04). Leave-one-out influencer sensitivity analysis showed high consistency in pooled aHRs across all iterations, ranging from 6.14 to 7.48 (Supplementary Table 6). Limiting analysis to studies with at least 5 years of follow-up (n = 5) yielded a pooled aHR of 6.43 (95% CI: 3.33–12.41), compared with 6.66 (95% CI: 2.79, 15.90) in the full sample, indicating that the relationship between fibrosis stage and MALO was consistent across longer periods of observation. However, given the observed statistical heterogeneity in pooled HRs comparing F4 and F3 or F4 and F2, the magnitude of increased risk of MALO at higher fibrosis stages should be interpreted with caution.

The incidence rate per 1000 person-years of MALO, reported in two studies, was 2.9 in F1 fibrosis, 4 to 7.1 in F2 fibrosis, 28 to 28.2 in F3 fibrosis and 97 in F4 fibrosis (Figure 3D) [31,41].

All-cause mortality

In total, 15 studies reported on all-cause mortality. Four studies (three retrospective [29,39,44] and one prospective [33] cohorts), enrolling 56–67% patients with MASH, were included in a FE meta-analysis of uHRs comparing patients with F4 and F3 fibrosis. A uHR was estimated indirectly for one study, while uHRs were estimated from digitized KM curves for all other studies (Table 2). There was a twofold increased risk of all-cause mortality in patients with F4 versus F3 fibrosis (pooled uHR of 2.30; 95% CI: 1.35, 3.94; I^2 = 0%). Pooled uHRs remained stable across all iterations of leave-one-out influencer analyses, ranging from 2.15 to 2.47 with a majority of pooled estimates retaining statistical significance (Supplementary Table 6). The effect remained consistent when pooling aHRs from three of these four studies (pooled aHR: 2.13, 95% CI: 1.10, 4.10; I^2 = 0%) [29,33,39].

Meta-analysis of other F-stage comparisons consistently showed an increased risk of all-cause mortality in patients at higher versus lower fibrosis stages. Three studies (two retrospective cohorts [29,44] and one prospective cohort [33], enrolling 56–67% patients with MASH) were included in a FE meta-analysis of uHRs comparing patients with F4 and F0–2 fibrosis. A fivefold increase in risk of all-cause mortality was found in patients with F4 versus F0–2 (pooled uHR: 5.63, 95% CI: 3.15, 10.06; I^2 = 0%; Table 2). When comparing patients with F3–4 and F0–2 fibrosis, the pooled estimate of effect was attenuated. Based on a RE meta-analysis (PM-HKSJ model) pooling nine studies (six retrospective [29,35,39,44–46] and three prospective [32,33,47] cohorts), enrolling 18–75% patients with MASH, patients with advanced fibrosis (F3–4) had a threefold increase risk of all-cause mortality compared with patients with mild-moderate fibrosis (F0–2) (pooled aHR: 3.36, 95% CI: 2.37, 4.75; I^2 = 0%; Table 2; Supplementary Table 5). When the analysis was limited to studies judged as having moderate or low RoB, the pooled aHR was similar to the overall estimate (3.33, 95% CI: 2.26, 4.90) [29,32,33,35,39,44,46,47]. Leave-one-out sensitivity analysis showed minimal variation, with pooled aHRs ranging from 3.18 to 4.06 (Supplementary Table 6). When limited to studies with at least 5 years of follow-up (n = 6), the pooled aHR was 4.22 (95% CI: 2.28, 7.83), compared with 3.36 (95% CI: 2.37, 4.75) in the full analysis [32,35,39,45–47].

The incidence rate per 1000 person-years of all-cause mortality, reported in three studies, was 4.7 to 5.1 in F1 fibrosis, 4.9 to 7 in F2 fibrosis, 8.2 to 10.5 in F3 fibrosis and 18 to 26.3 in F4 fibrosis (Figure 3E) [29,33,48].

Other outcomes

Three studies reported HRs of liver-specific mortality [33,47,49]; however, these studies did not report the same fibrosis stage comparisons, which prevented meta-analysis (Supplementary Table 7). HRs for end-stage liver disease (one study) [33] and MACE/CV-related events (two studies) [4,31] were limited; therefore, meta-analysis was not feasible. None of the identified studies reported HRs of liver transplant. Only two studies reported HRQoL; however, data were presented as mean HRQoL scores by liver fibrosis stage at baseline [50] and 18 months after randomization [51]. These data trended toward worse outcomes among patients with higher fibrosis stage.

MASH-majority sensitivity analysis

Of the outcomes reported in MASH-majority populations, only data for MALO were sufficient for meta-analysis. Three studies (two RCTs [30,40] and one retrospective cohort study [31]) reported uHRs comparing patients with F4 and F3 fibrosis. Based on a fixed effects meta-analysis, patients with F4 fibrosis had a significantly increased risk of MALO compared with those with F3 fibrosis (pooled uHR: 2.06, 95% CI: 1.34, 3.14). Substantial statistical heterogeneity (I^2 = 74.8%) was detected. This heterogeneity appeared to be driven by the data reported by Goh *et al.* [30] from the STELLAR-3/STELLAR-4 trials, which did not clearly report whether this HR was adjusted or

unadjusted (7.76, 95% CI: 1.83, 32.87). Variation in the definition of MALO across studies also contributes to the observed heterogeneity, although the other two contributing estimates showed a more modest increase in risk of MALO (uHRs 1.20 and 3.12; Supplementary Table 2). This warrants caution while interpreting the pooled effect estimate.

The variation in fibrosis stage comparisons and general paucity of data reported prevented the calculation of any further pooled HRs in MASH-majority populations. However, trends in individually reported HRs were directionally consistent with the rest of the MASH ± MASLD meta-analysis results reported in this manuscript (Supplementary Table 8).

Discussion

This meta-analysis assessed the association between liver fibrosis stage and key patient-relevant outcomes to provide evidence for the utility of histologic liver fibrosis as a surrogate marker in patients with MASH. Across both MASH ± MASLD and MASH-majority populations, available evidence consistently supported a positive association between advancing fibrosis stage and increased risk of several clinical outcomes. This association was demonstrated in the meta-analyses of MASH ± MASLD study populations. For example, the comparison of patients with F3–4 versus F0–2 fibrosis showed over a threefold increased risk of all-cause mortality and more than a fivefold increased risk of HCC. Pooled comparisons of patients with stage F4 versus F3 fibrosis also yielded a statistically significant increase in all-cause mortality risk, while the F3 versus F2 comparison similarly showed a significant association with cirrhosis. These associations remained directionally consistent across leave-one-out sensitivity analyses, and no single study altered statistical significance. In the MASH majority population, meta-analysis was generally not feasible due to limited data. One exploratory meta-analysis was conducted for MALO, which found that patients with stage F4 fibrosis showed twofold higher risk of MALO compared with those with stage F3 fibrosis.

The present meta-analysis strengthens the existing literature by expanding on the estimates calculated previously, by including more recently published primary data, more clinical and additional novel outcome types (some of which had previously not been included in a published meta-analysis of HRs, i.e., cirrhosis, hepatic decompensation, HCC and MALO), calculating pooled HRs for multiple fibrosis stage comparisons, and aiming to provide greater nuance to the understanding of progression from moderate fibrosis (stages F2–F3) to cirrhosis (stage F4). The present meta-analysis was also the first, to the authors' knowledge, to publish a pooled risk estimate in a MASH-majority population. As the evidence base for the MASH population increases, future meta-analyses can provide estimates for more outcome types to provide more accurate estimates in this higher-risk group. Findings from this analysis are consistent with previous meta-analyses that found significant associations between increased fibrosis stage and clinical outcomes, including mortality and liver-related events [2,16–18]. These included an exponential increase in pooled risk of both all-cause and liver-specific mortality incrementally with increasing fibrosis stage [16,17], increases in pooled risk of all-cause mortality and liver-related events in advanced fibrosis versus mild-moderate or no fibrosis (as well as between moderate fibrosis [F2] and mild or no fibrosis [F0–1]) [18], and an estimated incidence of progression from stage F3 fibrosis to F4 fibrosis (cirrhosis) of 5.1 patients per 100 person-years [2]. The consistent association with increased risk of progression to cirrhosis is particularly relevant as it is, itself, associated with risk of serious clinical outcomes such as mortality [52].

Notably, the estimates of association between fibrosis stage and clinical outcomes in many of the meta-analyses showed substantial heterogeneity. Potential sources of heterogeneity include differences in the proportion of patients with a MASH diagnosis, staging system for liver fibrosis, number and nature of variables used in statistical adjustment of aHRs, and outcome definitions across studies. The proportion of patients with MASH in each study population varied widely, which allowed the inclusion of more studies for meta-analysis; however, patients with progressive disease (i.e., MASH) have greater risk of clinical outcomes than those with MASLD alone [53]. While there were various liver fibrosis staging systems used across studies, the harmonization scheme is likely to have mitigated major impacts on study conclusions, though residual misclassification of patients may have increased heterogeneity. Outcome definitions also varied across studies for outcomes such as hepatic decompensation, defined in one study as incidence of esophageal varices, hepatic encephalopathy, ascites or trans-jugular intrahepatic portosystemic shunt, and in another as incidence of ascites, encephalopathy or variceal hemorrhage.

Strengths of this study include a rigorous SLR and the broad list of outcomes and fibrosis stage comparisons that were considered for meta-analysis, maximizing the use of the evidence base. The sensitivity analyses also provided confirmation that the direction of association for each fibrosis stage comparison was consistent and was

not driven by outlier studies alone, or by studies with different lengths of follow-up or serious RoB. The limitations of this meta-analysis were largely related to the lack of available primary studies reporting the risk of outcomes in populations with MASH. While a larger number of studies are available for MASLD-only populations, the risks of adverse outcomes in those with progressive disease (i.e., MASH) are expected to be substantially greater than in MASLD, due to the advanced inflammation, liver injury and fibrosis [53]. Additionally, as with all literature reviews there was a risk of publication bias, limiting the inclusion of negative results. While this study attempted to minimize this by including data from grey literature sources, including nonpeer reviewed sources, formal assessment was not feasible for several pooled outcomes because of the limited number of contributing studies. It is difficult to quantify the impact of nonpublication of relevant data [54], however if negative results (statistically nonsignificant values or significant values showing an inverse relationship between fibrosis stage and clinical outcomes) were less likely to be published, this may have led to an overestimation of the positive association between fibrosis stage and clinical outcomes [55]. In addition, the exclusion of nonEnglish language publications may have excluded relevant international studies, limiting the generalizability of results. This SLR was conducted according to a protocol and search strategy drafted a priori, though it was not prospectively registered. This study included data from both clinical trial and observational study designs, which maximized the use of available longitudinal data. While observational study designs more frequently report longer follow-up times with sufficient outcome incidence to conduct analysis, there was a risk of residual confounding by variables not accounted for in each study's statistical adjustment model. If present, residual confounding would lead to overestimation of the positive association between fibrosis stage and clinical outcomes.

In conclusion, both meta-analysis and narrative synthesis derived from studies involving mixed populations of MASH and MASLD patients indicate that a progressively worsening fibrosis (particularly cirrhosis) is linked to increased risk of adverse clinical outcomes, including hepatic decompensation, HCC, MALO, all-cause and liver-specific mortality and HRQoL. These findings support the notion that liver fibrosis is an important and useful short-term surrogate for long-term clinically relevant events.

Summary points

- Metabolic dysfunction-associated steatohepatitis (MASH), a progressive form of metabolic dysfunction-associated liver disease (MASLD), is associated with serious long-term adverse clinical outcomes which may occur over a longer timeframe than is practical to measure in a clinical trial.
- Key adverse clinical outcomes of MASH include liver cirrhosis, decompensation (typically involving symptoms such as ascites, variceal bleeding, hepatic encephalopathy) and hepatocellular carcinoma (HCC), which can result in liver failure, requiring liver transplant, and death.
- Histologic liver fibrosis staging has been identified in pivotal clinical trials as a surrogate marker that is observable over a shorter timeframe, therefore, allowing earlier identification of clinical efficacy.
- The objective of this systematic literature review (SLR) was to synthesize estimates of the relationship between histologic liver fibrosis stage and a variety of relevant clinical outcomes in a population of adults with MASLD and/or MASH.
- The search strategy identified articles from 1 January 2014 – 13 November 2024, from Ovid MEDLINE® and Embase as well as clinical trial registries.
- Meta-analyses were conducted pooling together hazard ratios (HRs) using fixed and/or random effects models (if $n \geq 5$ estimates were available) for each outcome, and separately for available fibrosis stage comparisons (F4 vs F3, F4 vs F0–2, etc.).
- This SLR identified 32 unique studies (n = 22 studies mixed MASH and MASLD patient populations, n = 10 studies MASH-majority [≥80% MASH] patient populations), with a range of 17.6–100% of patients having MASH.
- Meta-analyses were performed for risk of progression to cirrhosis, hepatic decompensation, HCC, MALO and all-cause mortality, with a consistent pattern of increased risk of adverse outcomes among patients with higher fibrosis stage, relative to lower stage.
- The only outcome for which meta-analysis was feasible among MASH-majority study populations was MALO, patients with F4 fibrosis had a significantly increased risk of MALO compared with those with F3 fibrosis (2.06, 95% CI: 1.34, 3.14).
- Increased risk of clinical outcomes (several synthesized for the first time) in advanced stage fibrosis, versus earlier stage, further supports the utility of liver fibrosis as a surrogate of disease progression in clinical trials of MASH.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at https://becarispublishing.com/doi/epdf/10.57264/cer-2026-0148

Author contributions

All authors contributed to study conception and design; A Qasim, Y Kim and SD Miller were responsible for investigation; all authors contributed to writing and editing of the manuscript.

Financial disclosure

This study was funded by Madrigal Pharmaceuticals (PA, USA).

Competing interests disclosure

AS Barritt has served as a speaker, a consultant and an advisory board member for Madrigal, Target RWE, Boehringer Ingelheim, Ionis and Mirium. Y Kim and J O'Donnell are employees of and own stocks and shares in Madrigal Pharmaceuticals. A Qasim, SD Miller and K Johnston are employees of Broadstreet HEOR. The authors have no other competing interests or relevant affiliations with any organization or entity with the subject matter or materials discussed in the manuscript apart from those disclosed.

Writing disclosure

No funded writing assistance was utilized in the production of this manuscript.

Data transparency statement

Study protocol, template data collection forms, data extracted from included studies, data used for all analyses, and analytic code may be requested from the authors.

References

Papers of special note have been highlighted as: • of interest

1. Hagström H, Shang Y, Hegmar H, Nasr P. Natural history and progression of metabolic dysfunction-associated steatotic liver disease. *Lancet Gastroenterol. Hepatol.* 9(10), 944–956 (2024).
2. Le P, Payne JY, Zhang L *et al.* Disease State transition probabilities across the spectrum of NAFLD: a systematic review and meta-analysis of paired biopsy or imaging studies. *Clin. Gastroenterol. Hepatol.* 21(5), 1154–1168 (2023).
3. Tamaki N, Kimura T, Wakabayashi SI *et al.* Long-term clinical outcomes in steatotic liver disease and incidence of liver-related events, cardiovascular events and all-cause mortality. *Aliment. Pharmacol. Ther.* 60(1), 61–69 (2024).
4. Noureddin M, Vipani A, Bresee C *et al.* NASH leading cause of liver transplant in women: updated analysis of indications for liver transplant and ethnic and gender variances. *Am. J. Gastroenterol.* 113(11), 1649–1659 (2018).
5. Younossi ZM, Stepanova M, Al Shabeeb R *et al.* The changing epidemiology of adult liver transplantation in the United States in 2013–2022: the dominance of metabolic dysfunction-associated steatotic liver disease and alcohol-associated liver disease. *Hepatol. Commun.* 8(1), e0352 (2024).
6. Ferretti S, Barreyro FJ. Worldwide increasing prevalence of non-alcoholic steatohepatitis as an indication of liver transplantation: epidemiological view and implications. *Curr. Hepatol. Rep.* 23(1), 193–203 (2024).
7. Ochoa-Allemant P, Hubbard RA, Kaplan DE, Serper M. Cause-specific mortality in patients with steatotic liver disease in the United States. *J. Hepatol.* 83(4), 860–869 (2025).
8. US Food and Drug Administration. Noncirrhotic Nonalcoholic Steatohepatitis With Liver Fibrosis: Developing Drugs for Treatment Guidance for Industry. (2018). https://www.fda.gov/media/119044/download
9. Madrigal Pharmaceuticals Inc. *REZDIFFRA: Package Insert*. US Food and Drug Administration website, PA, USA (2025). [updated December 2025]. https://www.accessdata.fda.gov/drugsatfda_docs/label/2025/217785s002s004s005lbl.pdf
10. European Medicines Agency. Summary of Product Characteristics: rezdiffra. (2025). https://www.ema.europa.eu/en/documents/product-information/rezdiffra-epar-product-information_en.pdf
11. Novo Nordisk Inc. *WEGOVY: Package Insert*. US Food and Drug Administration website, NJ, USA (2025). [updated December 2025]. https://www.accessdata.fda.gov/drugsatfda_docs/label/2025/218316Orig1s000lbl.pdf
12. Behari J, Gougol A, Wang R *et al.* Incidence of hepatocellular carcinoma in nonalcoholic fatty liver disease without cirrhosis or advanced liver fibrosis. *Hepatol. Commun.* 7(7), e00183 (2023).
13. US Food and Drug Administration. Table of surrogate endpoints that were the basis of drug approval or licensure. (2024). https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure

- **This guidance from the US FDA includes the list of acceptable surrogate endpoints for various disease areas, including metabolic dysfunction-associated steatohepatitis.**

14. European Medicines Agency. Reflection papers on regulatory requirements for the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH) - Scientific guideline. (2023). https://www.ema.europa.eu/en/reflection-papers-regulatory-requirements-development-medicinal-products-chronic-non-infectious-liver-diseases-pbc-psc-nash-scientific-guideline
15. Sanyal AJ, Nakajima A, Bugianesi E *et al.* Expert Delphi consensus on surrogate endpoints for treatment assessment in metabolic dysfunction-associated steatohepatitis. *Clin. Gastroenterol. Hepatol.* 24(8), 2140–2153 (2026).

- **This expert Delphi panel emphasizes the value of histologic fibrosis staging as a surrogate endpoint for clinical trials of treatments for metabolic dysfunction-associated steatohepatitis.**

16. Dulai PS, Singh S, Patel J *et al.* Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. *Hepatology* 65(5), 1557–1565 (2017).
17. Ng CH, Lim WH, Lim GEH *et al.* Mortality outcomes by fibrosis stage in nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Clin. Gastroenterol. Hepatol.* 21(4), 931–939.e935 (2023).

- **This systematic review and meta-analysis, focused on a broader clinical population of nonalcoholic fatty liver disease, found similar results showing increased risk of mortality (all-cause and liver-specific) with increased histologic fibrosis stage, supporting its value as a surrogate endpoint.**

18. Taylor RS, Taylor RJ, Bayliss S *et al.* Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Gastroenterology* 158(6), 1611–1625.e1612 (2020).
19. Page MJ, McKenzie JE, Bossuyt PM *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, n71 (2021).
20. Brooke BS, Schwartz TA, Pawlik TM. MOOSE reporting guidelines for meta-analyses of observational studies. *JAMA Surg.* 156(8), 787–788 (2021).
21. Shea BJ, Reeves BC, Wells G *et al.* AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 358, j4008 (2017).
22. Rinella ME, Lazarus JV, Ratziu V *et al.* A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *J. Hepatol.* 79(6), 1542–1556 (2023).
23. Barritt AS, Yu F, Mospan AR *et al.* High concordance between nonalcoholic fatty liver disease and metabolic dysfunction-associated steatotic liver disease in the TARGET-NASH real-world cohort. *Am. J. Gastroenterol.* 119(8), 1624–1627 (2024).
24. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med. Res. Methodol.* 12(1), 9 (2012).
25. Hartung J, Knapp G. On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat. Med.* 20(12), 1771–1782 (2001).
26. Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Stat. Med.* 20(24), 3875–3889 (2001).
27. Sidik K, Jonkman JN. A simple confidence interval for meta-analysis. *Stat. Med.* 21(21), 3153–3159 (2002).
28. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control. Clin. Trials* 7(3), 177–188 (1986).
29. Fujii H, Iwaki M, Hayashi H *et al.* Clinical outcomes in biopsy-proven nonalcoholic fatty liver disease patients: a multicenter registry-based cohort study. *Clin. Gastroenterol. Hepatol.* 21(2), 370–379 (2023).
30. Goh GBB, Anstee QM, Trauner M *et al.* Validation of histologic and noninvasive measures of fibrosis as surrogate endpoints of disease progression in patients with nonalcoholic steatohepatitis (NASH). *Hepatol. Int.* 14(Suppl. 1), S317–S318 (2020).
31. Younossi ZM, Mangla KK, Berentzen TL *et al.* Liver histology is associated with long-term clinical outcomes in patients with metabolic dysfunction-associated steatohepatitis. *Hepatol. Commun.* 8(6), e0423 (2024).

- **This observational study was conducted in a MASH population, and reported increased risk of progression to cirrhosis for patients with F3 vs. F2 baseline histologic fibrosis stage.**

32. Peleg N, Sneh Arbib O, Issachar A *et al.* Noninvasive scoring systems predict hepatic and extra-hepatic cancers in patients with nonalcoholic fatty liver disease. *PLoS One* 13(8), e0202393 (2018).
33. Sanyal AJ, Van Natta ML, Clark J *et al.* Prospective study of outcomes in adults with nonalcoholic fatty liver disease. *N. Engl. J. Med.* 385(17), 1559–1569 (2021).

- **This large, prospective multicenter observational study followed a population of nonalcoholic fatty liver disease (metabolic dysfunction-associated liver disease) patients for a median of 4 years, reporting increased risk of all-cause mortality and various liver-related complications associated with increased histologic fibrosis stage.**

34. Angulo P, Kleiner DE, Dam-Larsen S *et al.* Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology* 149(2), 389–397 (2015).

- **This retrospective international cohort study followed patients for over 35 years to gather longitudinal data on risk of liver-related events associated with increased histologic fibrosis stage.**

35. Hirose S, Matsumoto K, Tatemichi M *et al.* Nineteen-year prognosis in Japanese patients with biopsy-proven nonalcoholic fatty liver disease: lean versus overweight patients. *PLoS ONE* 15(11), e0241770 (2020).
36. Kumazaki S, Hikita H, Tahata Y *et al.* Serum growth differentiation factor 15 is a novel biomarker with high predictive capability for liver cancer occurrence in patients with MASLD regardless of liver fibrosis. *Aliment. Pharmacol. Ther.* 60(3), 327–339 (2024).
37. Seko Y, Sumida Y, Tanaka S *et al.* Predictors of malignancies and overall mortality in Japanese patients with biopsy-proven non-alcoholic fatty liver disease. *Hepatol. Res.* 45(7), 728–738 (2015).
38. Seko Y, Sumida Y, Tanaka S *et al.* Development of hepatocellular carcinoma in Japanese patients with biopsy-proven non-alcoholic fatty liver disease: association between PNPLA3 genotype and hepatocarcinogenesis/fibrosis progression. *Hepatol. Res.* 47(11), 1083–1092 (2017).
39. 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.* 67(6), 1265–1273 (2017).
40. Sanyal AJ, Harrison SA, Ratziu V *et al.* The natural history of advanced fibrosis due to nonalcoholic steatohepatitis: data from the simtuzumab trials. *Hepatology* 70(6), 1913–1927 (2019).
41. Akbari C, Dodd M, Stal P *et al.* Long-term major adverse liver outcomes in 1260 patients with non-cirrhotic NAFLD. *JHEP Reports* 6(2), 100915 (2024).
42. Buzzetti E, Hall A, Ekstedt M *et al.* Collagen proportionate area is an independent predictor of long-term outcome in patients with non-alcoholic fatty liver disease. *Aliment. Pharmacol. Ther.* 49(9), 1214–1222 (2019).
43. Sebastiani G, Alshaalan R, Wong P *et al.* Prognostic value of non-invasive fibrosis and steatosis tools, hepatic venous pressure gradient (HVPG) and histology in nonalcoholic steatohepatitis. *PLoS One* 10(6), e0128774 (2015).
44. Tsutsumi T, Kawaguchi T, Fujii H *et al.* Hepatic inflammation and fibrosis are profiles related to mid-term mortality in biopsy-proven MASLD: a multicenter study in Japan. *Aliment. Pharmacol. Ther.* 59(12), 1559–1570 (2024).
45. Ito T, Ishigami M, Ishizu Y *et al.* Utility and limitations of noninvasive fibrosis markers for predicting prognosis in biopsy-proven Japanese non-alcoholic fatty liver disease patients. *J. Gastroenterol. Hepatol.* 34(1), 207–214 (2019).
46. Younes R, Govaere O, Petta S *et al.* Caucasian lean subjects with non-alcoholic fatty liver disease share long-term prognosis of non-lean: time for reappraisal of BMI-driven approach? *Gut* 71(2), 382–390 (2022).
47. Grimaudo S, Pipitone RM, Pennisi G *et al.* Association between PNPLA3 rs738409 C>G variant and liver-related outcomes in patients with nonalcoholic fatty liver disease. *Clin. Gastroenterol. Hepatol.* 18(4), 935–944.e933 (2020).
48. Hirose S, Kagawa T, Tsuruya K. Up to 40 years prognosis in Japanese patients with biopsy-proven nonalcoholic fatty liver disease. *Hepatology* 68(Suppl. 1), 1326A–1327A (2018).
49. Hagström H, Nasr P, Ekstedt M *et al.* Accuracy of noninvasive scoring systems in assessing risk of death and liver-related endpoints in patients with nonalcoholic fatty liver disease. *Clin. Gastroenterol. Hepatol.* 17(6), 1148–1156.e1144 (2019).
50. Younossi ZM, Stepanova M, Anstee QM *et al.* Reduced patient-reported outcome scores associate with level of fibrosis in patients with nonalcoholic steatohepatitis. *Clin. Gastroenterol. Hepatol.* 17(12), 2552–2560.e2510 (2019).
51. Younossi ZM, Stepanova M, Nader F *et al.* Obeticholic acid impact on quality of life in patients with nonalcoholic steatohepatitis: REGENERATE 18-month interim analysis. *Clin. Gastroenterol. Hepatol.* 20(9), 2050–2058.e2012 (2022).
52. Luthra R, Sheth A. Understanding MASH: an examination of progression and clinical outcomes by disease severity in the TARGET-NASH Database. *Adv. Ther.* 42(2), 1165–1195 (2025).
53. Ekstedt M, Hagström H, Nasr P *et al.* Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* 61(5), 1547–1554 (2015).
54. Page M, Higgins J, Sterne J. Chapter 13: assessing risk of bias due to missing evidence in a meta-analysis [last updated August 2024]. In: *Cochrane Handbook for Systematic Reviews of Interventions, Version 6.5* Higgins J, Thomas J, Chandler J (Eds). Cochrane, (2024). https://cochrane.org/handbook
55. Dickersin K, Min YI. Publication bias: the problem that won't go away. *Ann. NY Acad. Sci.* 703, 135–148 (1993).