Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Title: Comparative Analysis of Epidemiological Research Output of Metabolic Dysfunction-Associated Steatotic Liver Disease and Alcohol-Associated Liver Disease: A Global Bibliometric Analysis, 2000–2025

Bhagya Kumarini Silva
Faculty of Medicine, University of Kelaniya
Dulmini Pavithra Senadheera
Faculty of Medicine, University of Kelaniya
Renisha Chrismi Karunanayake
Faculty of Medicine, University of Kelaniya
Tiloka de Silva
Faculty of Business, University of Moratuwa
Arjuna Priyadarsin De Silva
Faculty of Medicine, University of Kelaniya
Hithanadura Janaka de Silva
Faculty of Medicine, University of Kelaniya
Madunil Anuk Niriella
madunil.niriella@kln.ac.lk

Faculty of Medicine, University of Kelaniya

Research Article

Keywords: Bibliometrics, Fatty Liver, Global Health, Healthcare Disparities, Liver Diseases, Alcoholic, Metabolic Dysfunction-Associated Steatotic Liver Disease

Posted Date: September 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10787734/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Fatty liver disease linked to metabolic problems (MASLD) and liver disease caused by alcohol (ALD) are now the two most common causes of long-term liver illness worldwide, contributing heavily to cirrhosis and liver cancer. This study compared how much scientific research has been published on each condition between 2000 and 2025, and where that research came from. We identified 315 studies, covering 292 MASLD and 32 ALD publications. MASLD research was far more abundant, averaging 13.9 papers per year compared to just 1.5 for ALD. Both fields were dominated by wealthier countries (59% of MASLD studies and 72% of ALD studies came from high-income nations), with no ALD studies at all from low-income countries. ALD studies were more likely to involve multiple world regions (22% vs 2%) and to report external funding (31% vs 16%), though citations per year were similar between the two fields (about 4 per year each). These gaps should guide future research funding and priorities.

Introduction

Chronic liver disease is a major global health problem. Metabolic dysfunction-associated steatotic liver disease (MASLD), previously known as non-alcoholic fatty liver disease (NAFLD), and alcohol-associated liver disease (ALD) are its two leading causes worldwide [1]. The 2023 change in nomenclature from NAFLD to MASLD reflects an improved understanding of the metabolic drivers of the disease and aims to reduce the stigma associated with it [1].

MASLD is now widespread, closely linked to rising rates of obesity, type 2 diabetes, and metabolic syndrome [2]. It affects approximately 30% of adults globally, with substantial variation across regions and ethnic groups [3, 4]. The disease spectrum ranges from simple hepatic steatosis to metabolic dysfunction-associated steatohepatitis (MASH), which may progress to cirrhosis and hepatocellular carcinoma [2]. ALD is likewise a major cause of liver-related morbidity and mortality, particularly in regions with high alcohol consumption [5], and continues to present challenges in prevention, early detection, and treatment [6].

Bibliometric analysis is a reproducible method for mapping the scientific landscape of a disease field and identifying knowledge gaps [7]. It has repeatedly shown that hepatology research output is concentrated in high-income countries, reflecting institutional capacity, commercial interest, and established data infrastructure [8], and it provides a transparent and replicable framework for comparisons of this kind [9]. Within hepatology the approach has been applied to NAFLD/MASLD treatment trends [7], MASLD-related hepatocellular carcinoma [10], the gut–liver axis [11], and, most recently, ALD pathogenesis [12], providing structured evidence of which research questions attract sustained investment and which remain neglected. The same geographic inequity has been documented in bibliometric analyses of other chronic diseases, in which sub-Saharan Africa and the Eastern Mediterranean remain persistently underrepresented despite bearing a disproportionate share of global disease burden [13].

Two features of the current literature make a direct comparison timely. First, the 2023 nomenclature shift means that studies using legacy terminology may be missed by contemporary search strategies, complicating the interpretation of publication trends [1, 14]. Second, the emerging category of MetALD, comprising individuals who meet criteria for both metabolic dysfunction and significant alcohol consumption, lacks a coherent epidemiological evidence base [15]. Despite the shared burden of these two conditions, no study to date has compared their epidemiological research landscapes head to head in order to quantify the disparities side by side. We therefore undertook a bibliometric comparison of epidemiological research output on MASLD and ALD published between 2000 and 2025, examining publication volume and growth, geographical and income-group distribution, data sources, journal quartile, citation impact, and funding, with the aim of informing research priority-setting by funders and international liver societies [8, 16].

Methods

Study Design and Search Strategy

This bibliometric study analysed epidemiological research publications on MASLD and ALD. We conducted literature searches in PubMed/MEDLINE and Google Scholar to identify relevant studies published from 1 January 2000 to 31 December 2025; the searches were executed on March 2026. Scopus, Embase, and Web of Science were not searched because institutional access was unavailable, which is acknowledged as a limitation of this study. Search strategies combined Medical Subject Headings (MeSH) terms and keywords related to MASLD (including “non-alcoholic fatty liver disease”, “NAFLD”, “non-alcoholic steatohepatitis”, “NASH”, “metabolic dysfunction-associated steatotic liver disease”, “MASLD”, “metabolic dysfunction-associated fatty liver disease” and “MAFLD”), ALD (including “alcoholic liver disease”, “alcohol-related liver disease”, “alcohol-associated liver disease” and “alcoholic hepatitis”), and study design (including “epidemiology”, “population-based”, “prevalence”, “incidence”, “risk factors”, “morbidity” and “mortality”). The verbatim search string for each database is provided as Supplementary File 1.

Inclusion and Exclusion Criteria

Studies were included if they were original epidemiological research (cross-sectional and cohort), human studies, focused on prevalence, incidence, disease burden, associations, risk factors, mortality, and published in peer-reviewed journals and in English.

Studies were excluded if they were case reports or case series, mechanistic or basic science studies, clinical trials (therapeutic interventions), case-control studies, systematic reviews and meta-analyses, narrative reviews, editorials, or commentaries, or studies focusing exclusively on paediatric populations. Case-control studies were excluded because sampling on outcome precludes the estimation of prevalence, incidence, or burden, which were the measures of interest. Paediatric-only studies were

excluded because their diagnostic thresholds, natural history, and research infrastructure differ substantially from those in adults and have been examined in a separate bibliometric analysis [14].

Screening and Data Extraction

Records retrieved were de-duplicated and screened by title and abstract, followed by full-text review against predefined eligibility criteria. A second researcher independently verified final inclusion decisions, and disagreements were resolved by discussion The flow of records through identification, screening, full-text assessment, and inclusion is summarised in a PRISMA-style flow diagram (Figure S1).

Extracted data included bibliometric characteristics (publication year, journal name, SCImago quartile, first author affiliation, citation count), geographical information (country, World Bank income classification, WHO region), and study characteristics (design, sample size, population type, epidemiological measures reported, and funding source). Geographical and income-group classification was based on the country of the study population rather than the affiliation of the first author. For multi-country studies, all participating countries were recorded and classified as single- or multi-region by WHO criteria, and as single- or mixed-income group by World Bank classification. Nine studies reported data on both MASLD and ALD; these were counted within both disease groups, so that 315 unique publications yielded 324 disease-specific records. Citation counts were extracted from Google Scholar in January 2026.

Statistical Analysis

We performed descriptive statistics to summarise publication characteristics, calculating frequencies and percentages for categorical variables and means with standard deviations for continuous variables. Percentages are reported to one decimal place and may not sum to 100 because of rounding.

Temporal trends in publication volume were analysed using the compound annual growth rate (CAGR), calculated as [(Ending Value/Beginning Value) ^ (1/Number of Years)] − 1. For MASLD the CAGR was calculated from 2005 and for ALD from 2015, in each case the year of the first eligible publication in that series. Because the two windows differ in length, and because CAGR is determined solely by the first and last observation in a series, the two values are not directly comparable and are reported descriptively. Year-on-year growth rates were not compared statistically, because eligible ALD publications appeared in only five of the 21 years, leaving the corresponding ratios either undefined or highly unstable.

Mean annual publication output was compared using a paired t-test, pairing on calendar year. We used chi-square tests to compare categorical variables (region and income group of the country in which the study is conducted, data source, journal quartile, and funding declaration). Independent t-tests were used to compare continuous variables (citation counts, sample sizes).

Several comparisons involved sparse cells: no ALD study originated from the African or Eastern Mediterranean regions or from a low-income country, and expected counts fell below five in more than 20% of cells for the regional, income-group, and journal-quartile comparisons. The corresponding chi-square statistics are therefore reported as descriptive rather than confirmatory. Sample-size distributions were markedly right-skewed, and mean-based comparisons of sample size are presented for completeness only. No adjustment was made for multiple comparisons, and p-values should be interpreted as exploratory.

Citation analysis included both raw citation counts and time-normalised citations calculated by dividing total citations by the years since publication. This normalisation accounts for the advantage older publications have in accumulating citations.

We set the statistical significance threshold at $p < 0.05$ for all analyses, which were performed in R version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Overview of Literature Search and Study Selection

Searches identified 315 unique publications meeting the inclusion criteria (Figure S1). Because nine studies reported on both conditions, these yielded 324 disease-specific records: 292 for MASLD [21–312] and 32 for ALD [304–335], of which nine addressed both conditions [304–312]. All tables present disease-specific records. The earliest eligible MASLD study was published in 2005 and the earliest eligible ALD study in 2015; no eligible study was published between 2000 and 2004. Eligible ALD publications appeared in only five of the 21 years, and 87.5% (28/32) of ALD records were published in 2024 or 2025. The temporal distribution of publications differed markedly between the two conditions in volume, consistency, and recent trajectory (Fig. 1).

Publication Volume and Temporal Trends

MASLD research output exceeded that of ALD throughout the study period, with a mean of 13.9 publications per year compared with 1.5 for ALD. A paired t-test across all years from 2005 onwards confirmed that this difference in mean annual output was statistically significant ($p < 0.001$). The difference was less significant when the comparison was restricted to the five years in which ALD output was non-zero ($p = 0.065$), indicating that the difference is driven mainly by years of complete ALD absence; this restricted comparison rests on few observations and should be interpreted cautiously.

MASLD publications grew steadily, with a compound annual growth rate (CAGR) of 14.5% from 2005 to 2025. Output remained low in the initial years (two publications in each of 2005 and 2006), reached double figures by 2012, and peaked at 34 publications in 2024, with 30 in 2025. ALD research followed a far more sporadic pattern, with eligible publications in only five years (Fig. 1) and a CAGR of 25.9% from

2015 to 2025. ALD output did not exceed two publications in any year before 2024, when eight appeared, rising to 20 in 2025; together these two years account for 87.5% (28/32) of all ALD records.

Because the two CAGRs were calculated over windows of different length and are determined solely by the first and last year of each series, they are not directly comparable and are presented for description only. The intermittency of the ALD series, with no eligible publications in 16 of the 21 years, precludes any stable model-based estimate of its growth trajectory.

Geographical Distribution

Geographical analysis, based on WHO regions and including studies spanning multiple regions, showed striking disparities between MASLD and ALD research (Table 1). MASLD studies were predominantly region-specific, with 97.9% (286/292) conducted within a single WHO region and only 2.1% (6/292) involving multiple regions. In contrast, ALD research more frequently adopted a broader scope, with 21.9% (7/32) spanning multiple regions, accounting for 53.8% (7/13) of all multi-region records despite ALD comprising just 9.9% (32/324) of the dataset.

Table 1
Distribution of papers by WHO region

Research area	AFRO	AMRO	EMRO	EURO	SEARO	WPRO	Multiple regions	Total
MASLD, n	6	70	18	71	24	97	6	292
MASLD, % of row	2.1	24.0	6.2	24.3	8.2	33.2	2.1	100
ALD, n	0	11	0	8	1	5	7	32
ALD, % of row	0	34.4	0	25.0	3.1	15.6	21.9	100
Total, n	**6**	**81**	**18**	**79**	**25**	**102**	**13**	**324**
Total, % of row	**1.9**	**25.0**	**5.6**	**24.4**	**7.7**	**31.5**	**4.0**	**100**

AFRO = African Region; AMRO = Region of the Americas; EMRO = Eastern Mediterranean Region; EURO = European Region; SEARO = South-East Asia Region; WPRO = Western Pacific Region; Multiple regions = studies involving countries from more than one WHO region; n = number. *$\chi^2 = 35.93$, $df = 6$, $p < 0.001$ for the difference in regional distribution between research areas. Percentages may not sum to 100 because of rounding.*

Among MASLD records, the Western Pacific region (WPRO) accounted for 33.2% (97/292), driven primarily by contributions from China and the Republic of Korea. The European (EURO) and the Americas (AMRO) regions followed at 24.3% (71/292) and 24.0% (70/292), respectively, with the United States and European countries such as Spain and Germany playing key roles. South-East Asia (SEARO) contributed 8.2% (24/292), while the Eastern Mediterranean (EMRO) and African (AFRO) regions were notably

underrepresented at 6.2% (18/292) and 2.1% (6/292), respectively. For ALD, the Americas region led with 34.4% (11/32), heavily influenced by studies from the United States, followed by Europe at 25.0% (8/32). The Western Pacific accounted for 15.6% (5/32) and South-East Asia for 3.1% (1/32); neither the African nor the Eastern Mediterranean region contributed any single-region ALD study.

MASLD records outnumbered ALD records in every single-region category (for example, 95.1% of Western Pacific records were MASLD), whereas ALD accounted for the majority of multi-region records (53.8% vs 46.2%), which may indicate a tendency for ALD research to address broader, cross-regional questions. A chi-square test showed a significant association between research area and regional distribution (χ^2=35.93, df = 6, $p < 0.001$); as noted in the Methods, this test involves several cells with expected counts below five and is reported descriptively.

Income Classification Analysis

Analysis by World Bank income groups further highlighted disparities in research representation (Table 2). HICs dominated both fields, accounting for 58.6% (171/292) of MASLD and 71.9% (23/32) of ALD records. UMICs were more prominent in MASLD (26.0%; 76/292) than in ALD (6.3%; 2/32), reflecting a stronger MASLD focus in countries such as China. LMICs accounted for 11.3% (33/292) of MASLD records and only 3.1% (1/32) of ALD records. LICs were virtually absent, with 1.0% (3/292) of MASLD records and no ALD records. Mixed-income studies were more common in ALD (18.8%; 6/32) than in MASLD (3.1%; 9/292), aligning with ALD's multi-region emphasis. A chi-square test demonstrated a significant association between research area and income group distribution (χ^2=23.00, df = 4, $p < 0.001$); this p-value applies to the distribution as a whole rather than to any individual cell, and is again reported descriptively given the sparse cells involved.

Table 2
Distribution of papers by income group of study country

Research area	HIC	LIC	LMIC	UMIC	Mixed	Total
MASLD, n	171	3	33	76	9	292
MASLD, % of row	58.6	1.0	11.3	26.0	3.1	100
ALD, n	23	0	1	2	6	32
ALD, % of row	71.9	0	3.1	6.3	18.8	100
Total, n	**194**	**3**	**34**	**78**	**15**	**324**
Total, % of row	**59.9**	**0.9**	**10.5**	**24.1**	**4.6**	**100**

HIC = High-Income Country; LIC = Low-Income Country; LMIC = Lower-Middle-Income Country; UMIC = Upper-Middle-Income Country; Mixed = studies involving countries from more than one income group; n = number. χ^2 = 23.00, df = 4, $p < 0.001$ for the difference in income-group distribution between research areas. Percentages may not sum to 100 because of rounding.

At the country level, the United States accounted for 17% (50/292) of MASLD and 31% (10/32) of ALD records, underscoring its outsized role in ALD research. China accounted for 16% (48/292) of MASLD but only 3% (1/32) of ALD, while the Republic of Korea contributed 8% (23/292) to MASLD and 9% (3/32) to ALD. The United Kingdom showed minimal MASLD involvement (2%; 5/292) but higher involvement in ALD (6%; 2/32).

Data Sources and Sample Sizes

The studies utilised diverse data sources: primary data collection in 209 (64.5%), secondary data (e.g., national surveys, electronic medical records, health insurance data, or biobanks) in 99 (30.6%), and the Global Burden of Disease (GBD) database in 16 (4.9%) records (Table 3). Data sources were identifiable from abstracts, with secondary sources noted only if explicitly mentioned.

Table 3
Distribution of papers by type of data source

Research area	GBD	Primary	Secondary	Total
MASLD, n	10	201	81	292
MASLD, % of row	3.4	68.8	27.7	100
ALD, n	6	8	18	32
ALD, % of row	18.8	25.0	56.3	100
Total, n	**16**	**209**	**99**	**324**
Total, % of row	**4.9**	**64.5**	**30.6**	**100**

GBD = Global Burden of Disease; n = number. Secondary data sources include national surveys, electronic medical records, health insurance data, and biobanks. $\chi^2 = 29.98$, df = 2, $p < 0.001$ for the difference in data-source distribution between research areas.

Distributions differed significantly across research areas ($\chi^2=29.98$, df = 2, $p < 0.001$). MASLD studies heavily favoured primary data (68.8%; 201/292), suitable for targeted clinical or community-based investigations. ALD studies relied more on secondary sources (56.3%; 18/32) and GBD data (18.8%; 6/32), together accounting for 75.0% of ALD records. This may reflect the difficulty of collecting primary data on alcohol-related conditions, including under-reporting and stigma, or a preference for large-scale burden analyses; the present data cannot distinguish between these explanations.

Sample sizes were extracted from abstracts where available, excluding GBD-based studies and a small number of indeterminate cases (primarily secondary data papers). Mean sample size did not differ significantly overall (MASLD: 260,754 [n = 274]; ALD: 298,404 [n = 23]; p = 0.87), among primary data studies (MASLD: 9,276 [n = 199]; ALD: 9,907 [n = 8]; p = 0.97), or among secondary data studies (MASLD: 928,009 [n = 75]; ALD: 452,269 [n = 15]; p = 0.47). These distributions were highly skewed and several strata contained very few ALD studies, so the comparisons are descriptive and should not be read as

evidence of equivalence. Distinguishing primary from secondary sources on the basis of abstracts alone may also have introduced minor classification errors.

Citation Analysis and Research Impact

MASLD records accrued significantly higher mean total citations (30.4 per paper) than ALD records (9.2 per paper) ($p < 0.001$), consistent with MASLD's larger volume and longer publication history. Normalising for years since publication eliminated this difference (MASLD: 4.0 citations per year; ALD: 4.1 citations per year; $p = 0.91$). Given the small number of ALD records, this analysis had limited statistical power, and the absence of a detectable difference should not be interpreted as demonstrating equivalence of impact.

Journal quality, evaluated via SCImago quartiles, showed ALD records were more likely to appear in top-tier (Q1) journals (71.9%; 23/32) than MASLD records (52.7%; 154/292), with fewer ALD records in Q2 (9.4%) and a similar spread across Q3 (15.6%) and Q4 (3.1%) compared with MASLD (Q2: 30.8%, Q3: 14.0%, Q4: 2.4%) (Table 4). The difference in journal quartile distribution did not reach statistical significance at the pre-specified threshold of $p < 0.05$ ($\chi^2=6.67$, $df = 3$, $p = 0.08$) and should be interpreted cautiously as a non-significant trend.

Table 4
Distribution of papers by SCImago quartile of the publishing journal

Research area	Q1	Q2	Q3	Q4	Total
MASLD, n	154	90	41	7	292
MASLD, % of row	52.7	30.8	14.0	2.4	100
ALD, n	23	3	5	1	32
ALD, % of row	71.9	9.4	15.6	3.1	100
Total, n	**177**	**93**	**46**	**8**	**324**
Total, % of row	**54.6**	**28.7**	**14.2**	**2.5**	**100**

Q1 = highest quartile; Q4 = lowest quartile. SCImago quartile based on SCImago Journal Rank (SJR) indicator. n = number. *$\chi^2 = 6.67$, $df = 3$, $p = 0.08$ for the difference in quartile distribution between research areas.*

Journal quartile was associated with data source type ($\chi^2=16.78$, $df = 6$, $p = 0.01$) (Table 5). Records based on secondary data were relatively more concentrated in higher-quartile journals: secondary data accounted for 33.9% of Q1 and 32.3% of Q2 records, but only 19.6% of Q3 records and none of the Q4 records, all of which used primary data. This association involves empty cells and is reported descriptively.

Table 5
Distribution of papers by journal quartile and type of data source

Journal quartile	GBD	Primary	Secondary	Total
Q1, n	14	103	60	177
Q1, % of row	7.9	58.2	33.9	100
Q2, n	2	61	30	93
Q2, % of row	2.2	65.6	32.3	100
Q3, n	0	37	9	46
Q3, % of row	0	80.4	19.6	100
Q4, n	0	8	0	8
Q4, % of row	0	100	0	100
Total, n	**16**	**209**	**99**	**324**
Total, % of row	**4.9**	**64.5**	**30.6**	**100**

GBD = Global Burden of Disease; n = number. Q1 = highest quartile; Q4 = lowest quartile. $\chi^2 = 16.78$, df = 6, $p = 0.01$ for the association between data source type and journal quartile (chi-square test).

Funding Sources and Support

Funding declarations recorded in PubMed metadata revealed disparities in resource support. ALD records were significantly more likely to report funding (31%; 10/32) than MASLD records (16%; 48/292) ($p = 0.04$). It should be noted that the comparison is based on funding declaration and not funding itself as an absent declaration is not necessarily equivalent to an absence of funding.

Discussion

In this head-to-head bibliometric comparison of 315 publications yielding 324 disease-specific records, MASLD and ALD epidemiological research differed sharply in volume, geographical scope, and methodology, while sharing the same fundamental blind spots.

The dominance of MASLD research (292 records) is unsurprising given its connection to the global obesity and metabolic syndrome epidemics [3]. However, the field tends to repeat similar questions across similar populations [17]. This trend is driven by commercial interest in therapies and existing data infrastructure [18]. Recent bibliometric analyses of MASLD-related hepatocellular carcinoma [10] and of the gut–liver axis in hepatology [11] identify the same dominant contributors, principally the United States and China, with a near-total absence of sub-Saharan African and South Asian institutions, irrespective of the hepatological subtopic studied. By contrast, ALD, recognised for decades as a major cause of liver-related death, contributed only 32 records, disproportionately few relative to its burden [5].

ALD research was, however, more globally framed, with 21.9% of records spanning multiple WHO regions compared with 2.1% for MASLD. The concentration of ALD output in 2024–2025 is striking, but its interpretation warrants caution: it may reflect a genuine research response to post-pandemic increases in alcohol consumption and alcohol-related liver disease, but it is equally compatible with the greater visibility of recent literature in the databases searched and with the general expansion of steatotic liver disease publishing that followed the 2023 nomenclature change [1].

Both fields are predominantly concentrated in high-income countries [8]. The African region contributed 2.1% of MASLD records and the Eastern Mediterranean 6.2%, and neither contributed a single-region ALD study, even though both are projected to experience among the largest future increases in chronic liver disease [19]. This gap is not just a matter of capacity [8]. It highlights how international funding and journal visibility tend to favour populations with well-established healthcare systems. This structural inequity is not unique to liver disease: comparable bibliometric analyses of cardiovascular disease and other chronic conditions in the Arab region and low-income settings have documented identical cycles of underfunding and data scarcity [13]. The hepatology community has called for a reconfiguration of global research priorities, including targeted funding streams, capacity-building consortia, and open-access data mandates, but implementation has been slow and uneven [16, 17]. This creates a cycle: without data from low-income and middle-income countries, investment cannot be justified, and without investment, data cannot be generated [8]. Findings from high-income countries may not apply to low-income ones where genetics, diet, healthcare access, and alcohol consumption vary greatly. Future research should move away from convenience samples in tertiary hospitals and focus instead on population-based studies [17].

Methodologically, MASLD relies heavily on primary data collection (68.8%), which allows for detailed clinical profiling but limits replication in resource-limited areas. In comparison, ALD studies more often depend on secondary data and global burden estimates (75.0%), sacrificing detail for wider coverage. Despite these different strategies, citation impact per paper is comparable once adjusted for publication age, indicating that ALD papers are not less influential individually, merely far fewer. ALD papers also appeared more frequently in top-tier journals (71.9% vs 52.7%), although this difference was not statistically significant and should not be over-interpreted; if real, it may indicate that policy-relevant analyses of large secondary datasets attract greater editorial attention [18]. A recent analysis of the ALD pathogenesis literature using data-mining strategies similarly highlighted the dominance of mechanistic themes and the relative neglect of epidemiological and population-based work [12], a finding that aligns with the pattern observed here.

There are significant gaps in both fields. The dominance of cross-sectional studies limits insights into disease progression and the modification of risk factors, and prospective cohort studies remain scarce. The growing category of MetALD, which includes individuals who experience both metabolic dysfunction and significant alcohol consumption, is largely missing from the literature, even though evidence suggests that these combined exposures worsen liver injury and complicate treatment [15, 20]. ALD records reported external funding more often than MASLD records (31% versus 16%), which may reflect

a greater reliance on government and public health funding for ALD compared with the commercial and institutional mix seen in metabolic liver disease research [8]. Both areas nonetheless remain significantly underfunded relative to their global impact [8, 16].

This analysis has several limitations. First, only PubMed/MEDLINE and Google Scholar were searched; Scopus, Embase, and Web of Science were unavailable, and their inclusion would probably have identified additional records and altered the citation metrics. Google Scholar in particular is not a curated bibliographic index, and both its coverage and its citation counts change over time, which limits the exact reproducibility of the figures reported here. Second, the analysis was restricted to English-language publications, which is likely to have undercounted research from China, Japan, Brazil, and parts of Europe, and therefore probably understates output from upper-middle-income countries. Third, the transition in nomenclature from NAFLD to MASLD may have reduced search sensitivity despite the inclusion of legacy terms [1, 14]. Fourth, several key comparisons involved sparse or empty cells, sample-size distributions were highly skewed, and no adjustment was made for multiple testing; the reported p-values should therefore be regarded as exploratory rather than confirmatory. Fifth, nine studies contributed to both disease groups, so the two groups are not fully independent, which mildly violates the independence assumption of the chi-square test; a sensitivity analysis excluding these studies would strengthen the regional and income-group comparisons. Sixth, funding was ascertained from PubMed metadata, which is incomplete and records declared rather than actual support. Finally, classification of data source and extraction of sample size relied on abstracts rather than full texts, and the CAGR is determined only by the first and last year of each series and is therefore sensitive to the choice of window.

Future research should determine whether the increase in ALD output in 2024–2025 is sustained, model publication trends using count-based regression rather than endpoint-anchored growth rates, assess the methodological quality of included studies and their influence on policy [18], and explore the practical barriers researchers face in low-resource settings. Extending this comparative framework to viral hepatitis, autoimmune, and cholestatic liver diseases would establish whether the patterns observed here are specific to MASLD and ALD or reflect a more general misalignment between hepatology research activity and global disease burden.

Conclusions

This first head-to-head bibliometric comparison of MASLD and ALD epidemiological research reveals profound asymmetries and shared blind spots, with important implications for global liver disease research. MASLD dominates in volume, yet this lead masks geographic and methodological limitations: research is overwhelmingly region-specific, concentrated in high-income settings, and dependent on primary data collection. ALD, despite fewer publications, demonstrates greater international breadth and higher rates of declared funding; a higher proportion of ALD papers appeared in first-quartile journals, but this difference did not reach statistical significance. Time-normalised citation impact did not differ significantly between the two fields, although the small number of ALD records limits the strength of that

comparison. Both literatures share a critical failure: near-total absence of research from LICs and negligible representation from Africa and the Eastern Mediterranean, regions facing the steepest projected rises in liver disease burden. The dominance of cross-sectional designs and limited investigation of metabolic–alcohol interactions (MetALD) further constrain the evidence base.

Closing these gaps demands targeted funding for underrepresented regions, equitable international consortia, investment in prospective cohort studies, and open-access data sharing. Without coordinated action, the global hepatology research enterprise will remain misaligned with the actual disease burden.

Declarations

Author Contribution

All authors contributed to the study design. MAN, APDS and HJdS supervised the project. BKS, DPS and RCK conducted material preparation and data collection. TDS performed the analysis. BKS, DPS and RCK wrote the first draft of the manuscript. All authors revised the manuscript. All authors read and approved the final manuscript.

Conflict of Interest Statement:

On behalf of all authors, the corresponding author states that there is no conflict of interest

References

1. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F et al (2023) A multisociety Delphi consensus statement on new fatty liver disease nomenclature. Hepatology 78(6):1966–1986
2. Tilg H, Petta S, Stefan N, Targher G (2026) Metabolic dysfunction-associated steatotic liver disease in adults: a review. JAMA 335(2):163–174
3. Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L (2023) The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. Hepatology 77(4):1335–1347
4. Riazi K, Azhari H, Charette JH, Underwood FE, King JA, Afshar EE et al (2022) The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 7(9):851–861
5. Asrani SK, Devarbhavi H, Eaton J, Kamath PS (2019) Burden of liver diseases in the world. J Hepatol 70(1):151–171
6. Crabb DW, Im GY, Szabo G, Mellinger JL, Lucey MR (2020) Diagnosis and treatment of alcohol-associated liver diseases: 2019 practice guidance from the American Association for the Study of Liver Diseases. Hepatology 71(1):306–333

7. Dai JJ, Zhang YF, Zhang ZH (2023) Global trends and hotspots of treatment for nonalcoholic fatty liver disease: a bibliometric and visualization analysis (2010–2023). World J Gastroenterol 29(37):5339–5360
8. Devarbhavi H, Asrani SK, Arab JP, Nartey YA, Pose E, Kamath PS (2023) Global burden of liver disease: 2023 update. J Hepatol 79(2):516–537
9. Ninkov A, Frank JR, Maggio LA (2022) Bibliometrics: methods for studying academic publishing. Perspect Med Educ 11(3):173–176
10. Kalligeros M, Henry L, Younossi ZM (2025) Trends and visualization analysis of research on MASLD-related hepatocellular carcinoma: a bibliometric analysis. Discov Oncol 16:853
11. Wen Y, Zhang T, Zhang B, Wang F, Wei X, Wei Y et al (2024) Comprehensive bibliometric and visualized analysis of research on gut-liver axis published from 1998 to 2022. Heliyon 10(6):e27819
12. Zhang J, Yang Z, Liu X, Yang X, Jin X, Duan H et al (2025) New insights into the pathogenesis of alcoholic liver disease based on global research. Dig Dis Sci 70(3):903–918
13. Bani Hani A, Tarazi A, Hamadneh Y, Al-Samahan A, Awad R, Kan'an M et al (2024) Analyzing cardiovascular disease research in the Arab region: a bibliometric review from 2012 to 2022. Int J Vasc Med 2024:5915132
14. Gong X, Bai S, Lei E, Lu T, Chen Y, Cai J et al (2025) A bibliometric analysis of metabolic dysfunction-associated steatotic liver disease in children from 2004 to 2024. Front Pediatr 13:1468788
15. Israelsen M, Torp N, Johansen S, Hansen CD, Hansen ED, Thorhauge K et al (2024) GALAXY consortium. Validation of the new nomenclature of steatotic liver disease in patients with a history of excessive alcohol intake: an analysis of data from a prospective cohort study. Lancet Gastroenterol Hepatol 9(3):218–228
16. Lazarus JV, Mark HE, Allen AM, Arab JP, Carrieri P, Noureddin M et al (2024) A global action agenda for turning the tide on fatty liver disease. Hepatology 79(2):502–523
17. Lazarus JV, Mark HE, Anstee QM, Arab JP, Batterham RL, Castera L et al (2022) Advancing the global public health agenda for NAFLD: a consensus statement. Nat Rev Gastroenterol Hepatol 19(1):60–78
18. Balakrishnan M, Rehm J (2024) A public health perspective on mitigating the global burden of chronic liver disease. Hepatology 79(2):451–459
19. Huang DQ, Terrault NA, Tacke F, Gluud LL, Arrese M, Bugianesi E et al (2023) Global epidemiology of cirrhosis - aetiology, trends and predictions. Nat Rev Gastroenterol Hepatol 20(6):388–398
20. Feng Y, Xu HQ, Han PS, Liu T, Gao YH (2025) Metabolic and alcohol-associated liver disease: a novel liver disease entity. J Clin Hepatol 41(11):2235–2239

Figures

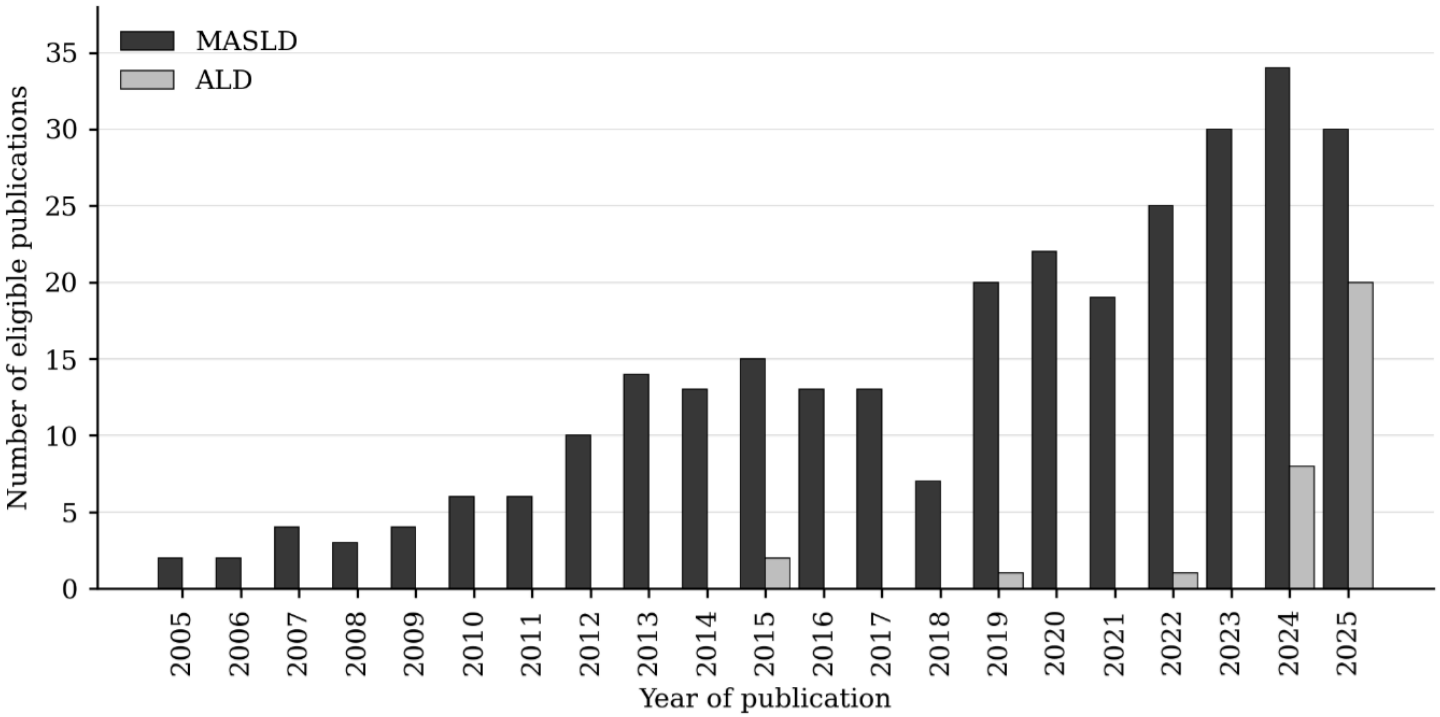


Figure 1

Number of eligible epidemiological publications per year for MASLD and ALD, 2005–2025.

MASLD = metabolic dysfunction-associated steatotic liver disease; ALD = alcohol-associated liver disease. Bars show disease-specific records (292 MASLD, 32 ALD); the nine studies reporting on both conditions are counted in each series. No eligible study met the inclusion criteria between 2000 and 2004.

Compound annual growth rate was 14.5% (2005–25) for MASLD and 25.9% (2015–25) for ALD. *The annual average is calculated over the full 21-year period (2005–2025). Each CAGR window begins at the year of the first eligible publication in that series.*

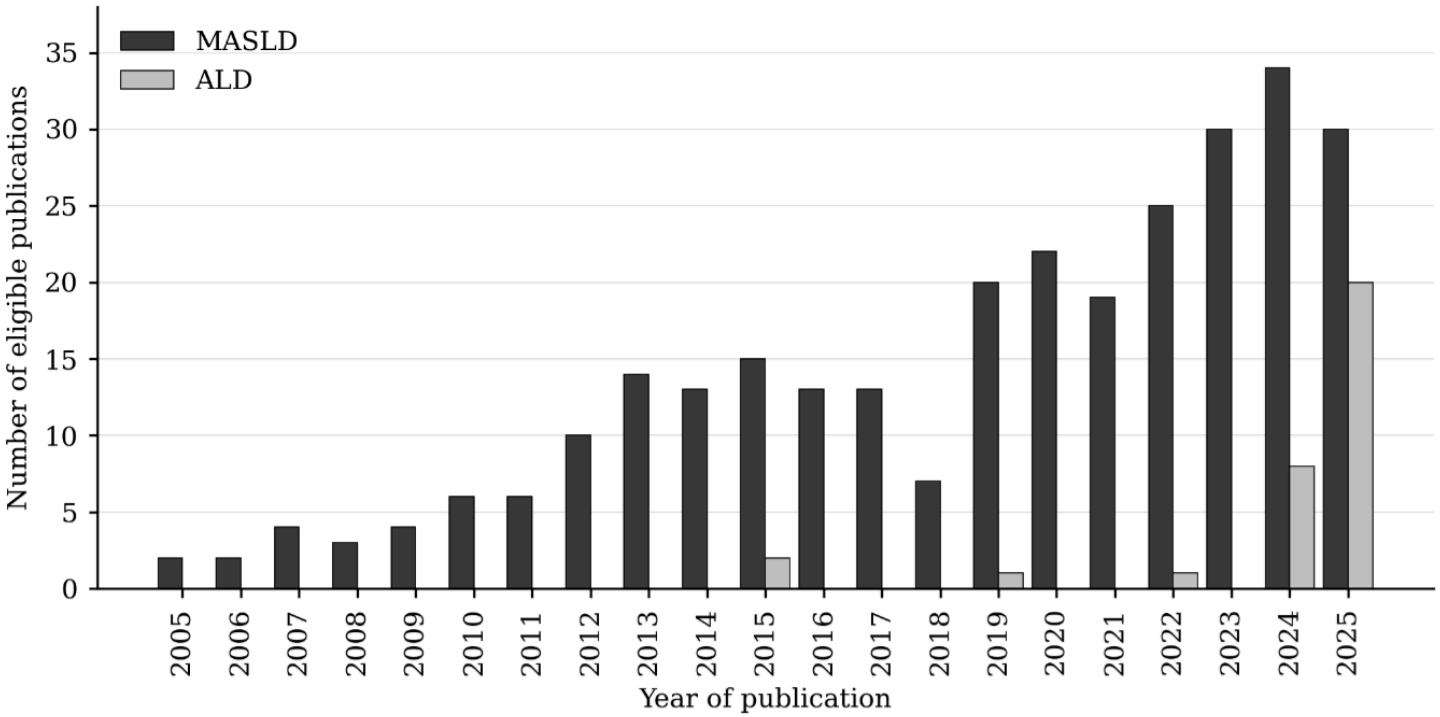


Figure 1

Number of eligible epidemiological publications per year for MASLD and ALD, 2005–2025.

MASLD = metabolic dysfunction-associated steatotic liver disease; ALD = alcohol-associated liver disease. Bars show disease-specific records (292 MASLD, 32 ALD); the nine studies reporting on both conditions are counted in each series. No eligible study met the inclusion criteria between 2000 and 2004.

Compound annual growth rate was 14.5% (2005–25) for MASLD and 25.9% (2015–25) for ALD. *The annual average is calculated over the full 21-year period (2005–2025). Each CAGR window begins at the year of the first eligible publication in that series.*

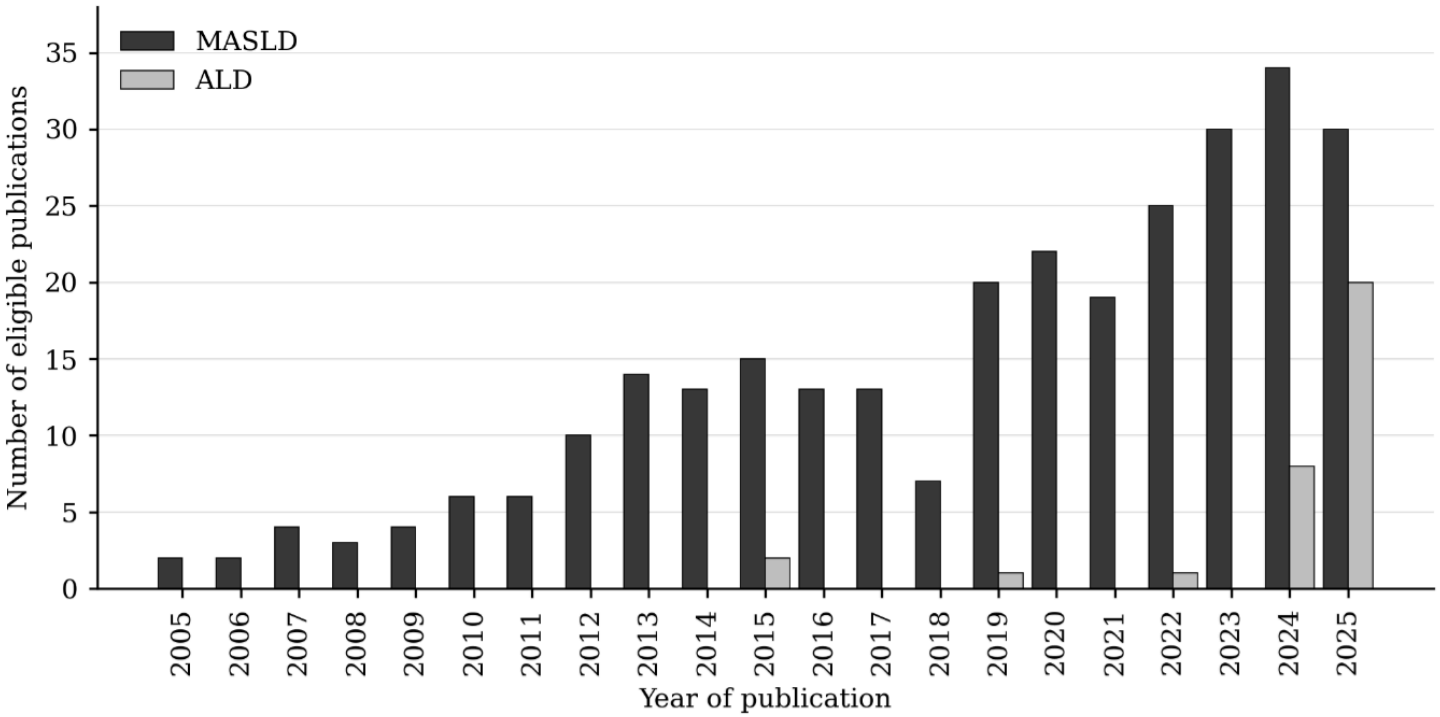


Figure 1

Number of eligible epidemiological publications per year for MASLD and ALD, 2005–2025.

MASLD = metabolic dysfunction-associated steatotic liver disease; ALD = alcohol-associated liver disease. Bars show disease-specific records (292 MASLD, 32 ALD); the nine studies reporting on both conditions are counted in each series. No eligible study met the inclusion criteria between 2000 and 2004.

Compound annual growth rate was 14.5% (2005–25) for MASLD and 25.9% (2015–25) for ALD. *The annual average is calculated over the full 21-year period (2005–2025). Each CAGR window begins at the year of the first eligible publication in that series.*

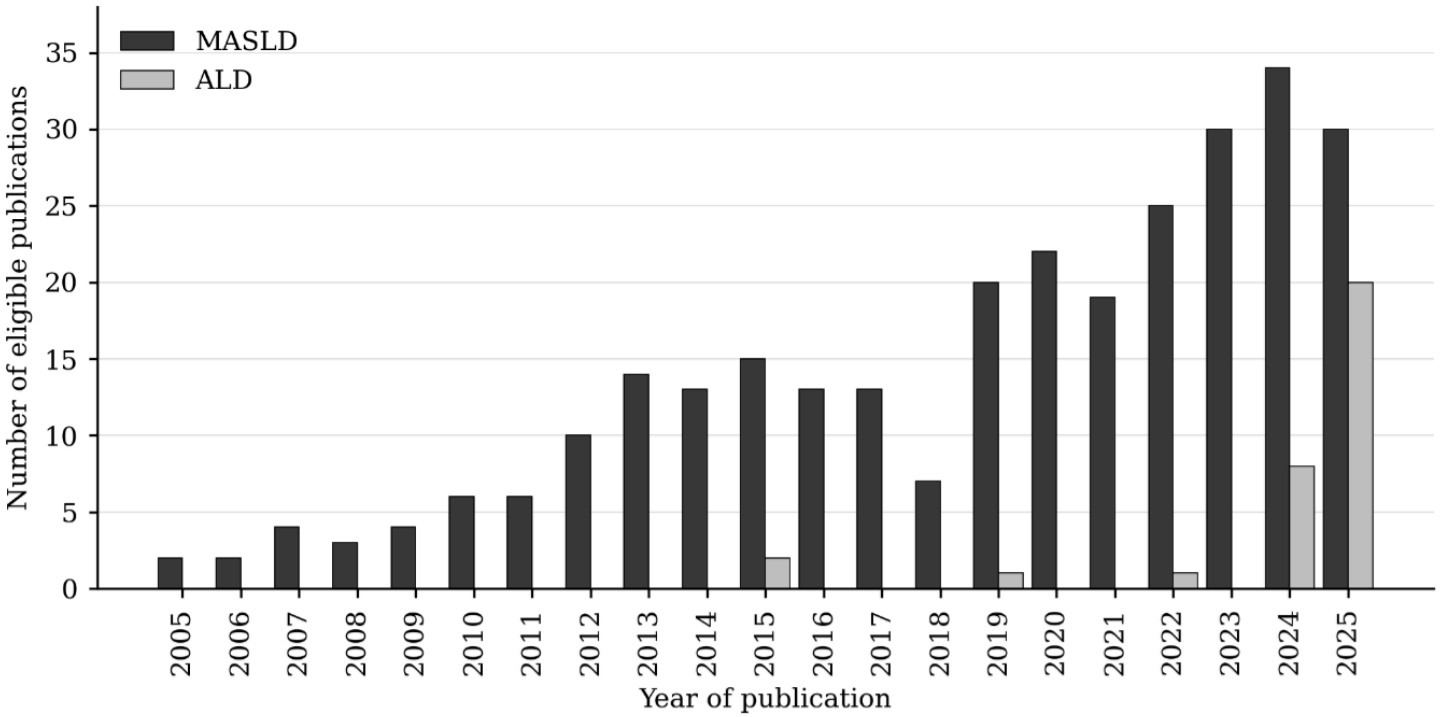


Figure 1

Number of eligible epidemiological publications per year for MASLD and ALD, 2005–2025.

MASLD = metabolic dysfunction-associated steatotic liver disease; ALD = alcohol-associated liver disease. Bars show disease-specific records (292 MASLD, 32 ALD); the nine studies reporting on both conditions are counted in each series. No eligible study met the inclusion criteria between 2000 and 2004.

Compound annual growth rate was 14.5% (2005–25) for MASLD and 25.9% (2015–25) for ALD. *The annual average is calculated over the full 21-year period (2005–2025). Each CAGR window begins at the year of the first eligible publication in that series.*

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFig1PRISMAflowdiagram06.08.26.docx
- SupplementaryFig1PRISMAflowdiagram06.08.26.docx
- SupplementaryFig1PRISMAflowdiagram06.08.26.docx
- SupplementaryFig1PRISMAflowdiagram06.08.26.docx
- SupplementaryFile2IncludedStudies06.08.26.docx
- SupplementaryFile2IncludedStudies06.08.26.docx

- SupplementaryFile2IncludedStudies06.08.26.docx
- SupplementaryFile2IncludedStudies06.08.26.docx