OPEN ACCESS

AI portal tract detection and characterisation for a regional analysis of steatosis and inflammation in MASLD, MASH and AIH

Dylan Windell ,[1] Alastair Magness,[1] Cayden Beyer,[1] Helena Thomaides Brears,[1] Sarah Larkin,[1] Kezia Hobson,[1,2] Paul Aljabar,[1] Kenneth Fleming,[1,3] Eve Fryer,[4] Timothy James Kendall ,[5] Reema Kainth,[1] Phil Wakefield,[1] Caitlin Rose Langford ,[1] Pierre Bedossa,[6] Robert Goldin[2]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jcp-2025-210311).

[1]Perspectum Ltd, Oxford, UK
[2]Imperial College London, London, UK
[3]University of Oxford, Oxford, UK
[4]Cellular Pathology, Oxford University Hospitals NHS Trust, Oxford, UK
[5]Centre for Inflammation Research, University of Edinburgh, Edinburgh, UK
[6]LiverPat, Paris, France

Correspondence to
Dr Dylan Windell; dylan.windell@perspectum.com

Received 20 June 2025
Accepted 12 October 2025

To cite: Windell D, Magness A, Beyer C, *et al*. *J Clin Pathol* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jcp-2025-210311

ABSTRACT

Aims Annotation of liver biopsies for disease staging is increasingly aided by digital pathology; however, existing systems do not quantify inflammation and steatosis within an anatomical framework. We hypothesise that an artificial intelligence (AI) system that quantifies portal tracts (PT) and the anatomical distribution of steatotic vesicles and inflammatory cells will align with manual pathologist scoring and stratify liver diseases.

Methods In this observational, cross-sectional study, digitised images of haematoxylin and eosin-stained specimens were pooled from four independent cohorts of metabolic dysfunction-associated steatotic liver disease (MASLD) or steatohepatitis (MASH) or autoimmune hepatitis (AIH) (n=390: 89 MASLD, 238 MASH, 63 AIH). PT, steatosis, and inflammation were quantified using a proprietary AI system and scored by expert pathologists.

Results The percentage of steatosis was higher in MASH (7.5%) than in MASLD (3.2%). Lobular regions had larger steatotic vesicles (260 vs 190 μm^2). AI-derived steatosis quantification correlated with manual grading (r_s=0.72). The inflammatory cell number (ICN) was twofold higher in AIH than MASLD/MASH in interface (390 vs 140), portal (4600 vs 1500) and lobular (1500 vs 650) regions. Portal inflammation from manual grading correlated with ICN count at PT (r_s=0.71) but not lobular regions ($r_s \leq 0.29$). For equivalent grades of portal inflammation, the ICN was up to threefold higher in AIH than in MASLD/MASH (r_s=0.71).

Conclusion A new AI system for anatomical quantification of liver biopsy features measured variation in fat and inflammation across the lobule. It showed that inflammation burden was higher in AIH than MASLD/MASH, despite equivalent portal grades, providing objective support for histological scoring.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Grading of inflammation in metabolic dysfunction-associated steatotic liver disease/steatohepatitis and autoimmune hepatitis is limited by low inter-rater reliability and challenges in identifying key anatomical features like portal tracts, contributing to delays and errors in clinical trial endpoint assignation.

WHAT THIS STUDY ADDS

⇒ This study demonstrates a system that automatically detects and quantifies portal tracts and surrounding patterns of inflammation and steatosis. Artificial intelligence (AI) showed that inflammation occurred in similar regions but was higher in autoimmune hepatitis than in metabolic dysfunction-associated steatohepatitis, despite similar grading from manual scoring.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This AI system provides granular and regional information that can aid biopsy scoring and support pathologists in liver disease staging and diagnosis for clinical trials.

INTRODUCTION

Metabolic dysfunction-associated steatotic liver disease (MASLD) affects one-third of the global population[1] and is characterised by increased fat in and around hepatocytes.[2] Progression to steatohepatitis (MASH), with concurrent liver inflammation, hepatocyte ballooning and increased risk of fibrosis, is estimated to occur in 5%–14% of adults and projected to increase.[1 3 4] Liver biopsy remains the reference standard for diagnosing and staging MASLD/MASH, particularly in complex cases.[2 5 6]

The metabolic comorbidities and multiplicity of disease features in MASH complicate drug efficacy, leading to a proliferation of antisteatotic[7] and anti-inflammatory[8] candidates for clinical trials. In late-phase clinical trials for MASH, resolution of histological features is often a primary endpoint. However, while clearly improved by blinded and centralised liver histology reading, both inter-rater and intra-rater reliability still adversely affect biopsy-based evaluation.[9–11] This highlights a continued need for standardisation of biopsy sampling and consensus in grading.[12]

Identification of portal tracts (PT) in liver biopsy samples allows both the assessment of biopsy adequacy and the staging of disease based on the relative anatomical distribution of liver disease features.[13 14] Identification of PT is challenging because they vary greatly in their size, structure and shape,[13] and their appearance is affected by biopsy needle size, cutting angle[15] and disease stage.[16] In recent audits,[13 17] at least 50% of biopsies failed

adequacy criteria on the basis of a minimal number of PT, so improving PT detection is also desirable.

Artificial intelligence (AI)-based systems have been developed that complement manual annotation and pathology grading of liver biopsy samples.[18] Existing systems for detecting and segmenting PT have been reported and are under evaluation for potential use in MASH clinical trials.[19] Such automated systems can show high degrees of correlation between manual annotations and automatic measures[20–22] and are increasingly able to distinguish and segment anatomical features such as interface hepatitis and bile ducts.[23] However, current systems have mainly been developed for fibrosis quantification rather than for quantification of the pathological features for MASH for assessment of progression/resolution. Existing efforts to detect steatosis and inflammation with AI lack an anatomical framework to assist with interpretation and grading of pathology.[21]

To further demonstrate the method's ability to quantify inflammation in different regions of the liver, we have also assessed a cohort of patients with autoimmune hepatitis (AIH). In AIH, the degree and localisation of the inflammation is of diagnostic, predictive, and prognostic significance.

In view of the above, we have created an AI system capable of detecting PT and quantifying pathological features of fat and inflammation relative to PT and other landmarks from biopsy samples stained with H&E. Using this new method in a large dataset of MASLD/MASH and AIH biopsies, we aimed to comprehensively evaluate the burden and regional distribution of (1) steatosis and (2) inflammation. We calculated the correlations with manual pathologist scoring and the associations to disease severity.

MATERIALS AND METHODS

Study design and populations

This study aimed to quantify liver disease features using our new AI system in samples pooled from four clinical cohorts that were fully independent at the patient level, with no overlap, as detailed previously.[24–26] Inclusion of digitised datasets was based on staining and quality of biopsy samples,[27 28] and on confirmation of AIH[29] and MASLD/MASH based on clinical criteria and pathology scoring independent of this work[27 28 30] (online supplemental figure 1). Participants with MASLD/MASH or AIH were enrolled at routine patient visits in secondary or tertiary care, at sites in the United Kingdom, the USA, Poland, and Japan. Patients with any other known chronic liver disease were excluded from this analysis. Where biopsies from multiple timepoints were available, only the baseline timepoint was selected to avoid duplication. All patients and volunteers provided written informed consent (details in online supplemental file 1).

Pathology protocol

Histological scoring

Histological whole slide images (WSI) were split between four pathologists with 14–40+ years of individual experience, blinded to all clinical data (online supplemental table 1). Biopsy sample adequacy was assessed using the definition outlined by the Royal College of Pathologists.[31]

Initial consensus scoring for inflammation, steatosis, ballooning and fibrosis was obtained from the original trial on all cases using H&E and trichrome (TRC)-stained slides to determine diagnosis. All MASLD/MASH and AIH WSIs were further scored by the expert pathologists using the Ishak scoring system[32] for portal, interface and lobular inflammation and the Non-Alcoholic Steatohepatitis Clinical Research Network (NASH CRN) scoring system[33] for lobular inflammation and steatosis in addition to separate fibrosis grading.

AI system for PT detection and quantification of steatosis and inflammation

We developed a proprietary suite of inter-related AI systems from H&E images that could identify landmark features, detect PT, perform regional demarcation and spatial analysis to quantify steatotic vesicles and inflammatory cells (figure 1, online supplemental methods). The AI system is built on convolutional neural network architectures (U-Net and U-Net++ variants) trained to detect key histological features, including PTs, steatotic vesicles and inflammatory cells. Models were trained on 31 liver biopsy WSIs, with an 80:20 train:validation split, and evaluated on an independent test set of 12 WSIs. Further technical details of the architecture, training and validation procedures are provided in online supplemental methods and online supplemental table 2.

AI measurements were derived within portal, interface, periportal, and lobular regions.

Steatosis was quantified using two complementary metrics:

- Steatosis fraction (SF) was calculated as the area occupied by steatotic vesicles divided by the total tissue area, providing an overall measure of fat burden.
- Vesicle-hepatocyte fraction (VHF) was calculated as the total number of steatotic vesicles divided by the total number of hepatocytes, approximating the proportion of hepatocytes containing lipid droplets.

Inflammation was quantified using two complementary metrics:

- Inflammatory cell number (ICN) represents the total number of inflammatory cells.
- Inflammatory burden (IB) represents the number of inflammatory cells per regional area (cells/mm^2), providing a density-based measure of inflammation.

Statistical analyses

Descriptive statistics were used to summarise cohort characteristics. Median and IQR were used to describe all non-normally distributed variables. Frequency and percentage were used to describe categorical variables.

Differences between AI measurements across disease states were assessed using Wilcoxon rank-sum test. Correlations between AI measurements and pathologist scoring were assessed using Spearman's rank correlation. Grouped scores between pathologists were either compared using Kruskal-Wallis tests or, if there were fewer than three scores within each category, a Wilcoxon rank-sum test was performed. Correlations between AI measurements in different regions were assessed using Pearson's correlation coefficient.

Multiple comparisons were addressed using per-family Bonferroni correction ($\alpha=0.05$). Given the small number of comparisons within each family (maximum of 12), this approach provided strong control of type 1 error without substantial loss of power.

All statistical analyses were performed in R Studio V.4.2.2. Values of $p<0.05$ were considered statistically significant.

RESULTS

Study populations

A total of 390 individuals were recruited from secondary care and paediatric clinics between February 2014 and March 2019 (online supplemental figure 1). Their mean age was 51 years, 48% were male, and their mean body mass index was 29.5 kg/m^2 (table 1). Overall, 89/390 (23%, all adults) had been

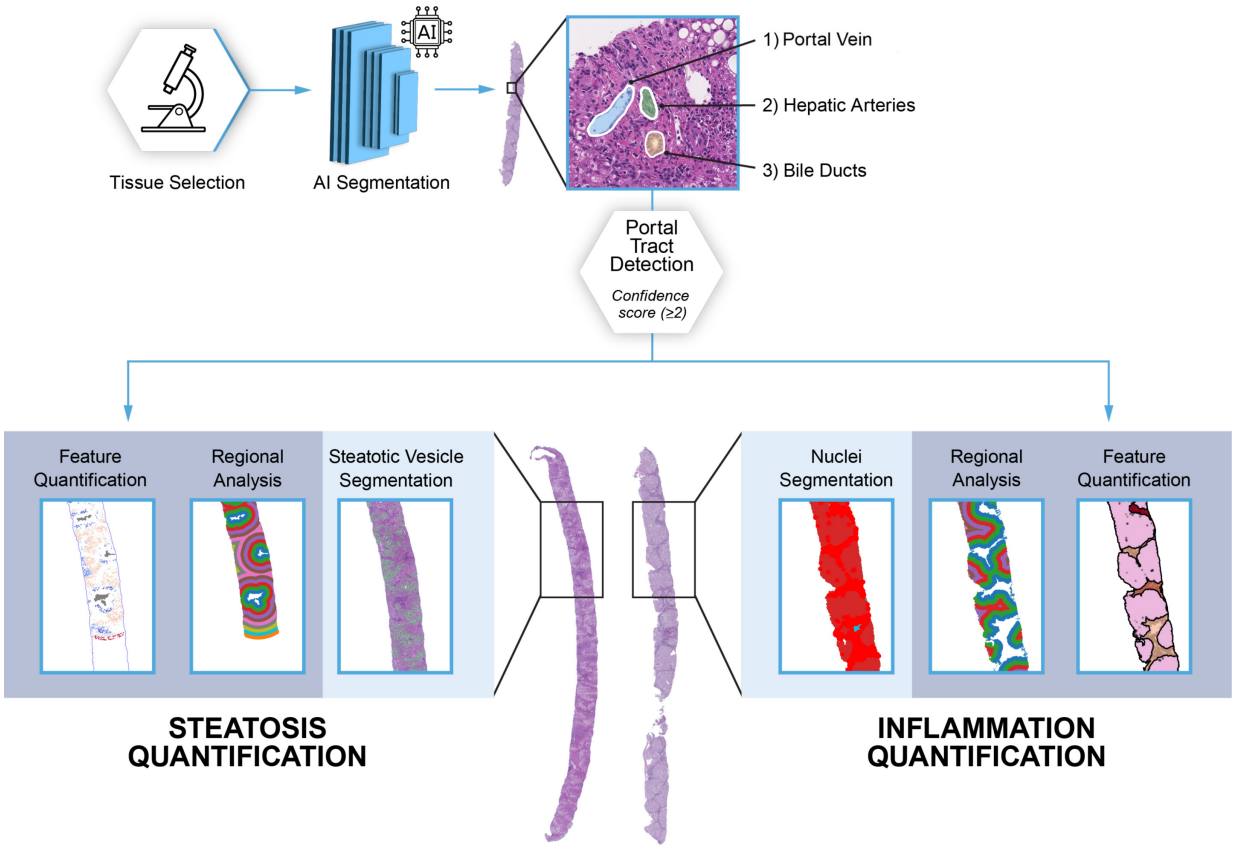

Figure 1 AI pipeline to segment inflammation and steatosis in relation to portal tract location. Equidistant regions (100 µm) from PT allow for detailed quantification of steatosis and inflammation in portal, periportal, and lobular regions of liver samples. AI, artificial intelligence; PT, portal tracts.

diagnosed with MASLD, 238/390 (61%, all adults) with MASH and 63/390 (16%, of which 3% were adults) had AIH. In those with MASLD/MASH, known comorbidities were obesity (51%), diabetes (35%), hypertension (21%) and dyslipidaemia (33%). These comorbidities were rare in those with AIH. Treatment for AIH (56%) was considered as ongoing if participants were receiving either steroids or azathioprine.

In those with MASH, 95 had 'at-risk MASH' (NAS≥4 and ≥2 fibrosis grade). Similarly, for AIH disease, 7 had moderate-severe disease (≥2 lobular and portal inflammation score).

PT quantification using AI

Every individual underwent a liver biopsy. Biopsy samples were scored manually by 1–4 pathologists and quantified with the AI-based quantification system of PT detection and liver tissue characterisation described previously (figure 1).

The AI system identified a total of 7359 PT across 390 digitised WSI, with a mean number of 19 PT per WSI and 7 PT per 10 mm of liver sampled (online supplemental table 3). Though not required as an inclusion criterion in our analyses, most samples met adequacy criteria for length and PT number recommended for pathological assessment.[14] 378/390 (97%) WSI met the minimum 6–11 PT/sample and 325/390 (83%) WSI met the minimum 10–12 PT/sample. The PT count correlated with length of sample (Spearman's $\rho=0.71$, $p<0.0001$) (online supplemental figure 2).

PT detection by AI was comparable in performance with manual pathologist detection (online supplemental table 4). Agreement on PT detection between different pathologists ranged in F1 score from 0.58±0.16 to 0.72±0.11. Comparison of AI to the pathologists' consensus on PT detection in F1 score was 0.66±0.18. The level of agreement improved in high fibrosis cases (F1 Score—0.74±0.12, fibrosis stage 3–4 vs 0.59±0.20, fibrosis stage 1–2; online supplemental table 4).

Quantification of steatosis in MASLD and MASH

The AI system delineated 1 658 463 steatotic vesicles from 327 WSI with MASLD or MASH, with a mean 3732 vesicles per sample for MASLD cases and 5572 vesicles per sample for MASH (online supplemental table 3).

The two AI metrics correlated with each other: the SF and the VHF (Spearman's $\rho=0.9$; $p<0.0001$). The SF was lower in MASLD cases compared with MASH (3.2% (IQR 5.6) vs 7.5% (IQR 7.1), $p<0.0001$) (figure 2, online supplemental figure 3). Similarly, the VHF was lower in MASLD (10.2%; IQR 12.1) than in MASH (20.1% IQR 17.2). The size of steatotic vesicles also differed. Median area of steatotic vesicles was 190 μm^2 (IQR 110) in those with MASLD and 260 μm^2 (IQR 100) in MASH ($p<0.0001$) (figure 3). These results indicate that the overall fat burden, the number of hepatocytes containing lipid droplets and the droplet size, are all greater in MASH than in MASLD.

Regional distribution of steatosis

Significant differences in SF between MASLD and MASH were observed in all regions up to 1 mm from each PT. In the immediate vicinity of the PT, up to 100 μm away, SF was lower compared with lobular regions at 500 μm from the PT (figure 3) in both MASLD and MASH. The biggest difference in SF was approximately 900 μm from the PT, with SF of 2.7% (IQR 5.9)

Table 1 Study populations

	All (n=390)	MASLD (n=89)	MASH (n=238)	AIH (n=63)
Demographics**				
Age (year)	51±19	55±15	56±12	21±17
BMI (kg/m^2)	29.5±6.8 Unknown (n=69)	27.9±5.1 Unknown (n=3)	32.6±5.7 Unknown (n=62)	22.6±6.1 Unknown (n=4)
Sex††				
Male	186 (48%)	61 (69%)	99 (42%)	26 (41%)
Female	201 (52%)	28 (21%)	139 (58%)	34 (54%)
Not specified	3 (<1%)	–	–	3 (5%)
Hypertension	68 (17%)	27 (30%)	41 (17%)	BP<133/85 mm Hg
Treatment††				
Yes—steroids/azathioprine	35 (9%)	–	–	35 (56%)
No	342 (88%)	89 (100%)	238 (100%)	15 (24%)
Unknown	13 (3%)	–	–	13 (20%)
Biopsy sampling and digitisation**				
Number of biopsy samples and WSI	390	89	238	63
Time between clinical visit and biopsy (days)	3 (n=50)	–	–	3 (n=50)
Length of biopsy (mm)	22±9 (n=108)	25±9 (n=34)	26±5 (n=24)	20±9 (n=50)
Manual pathology scoring**				
Number of PT per biopsy	19±12 (n=108)	12±2 (n=34)	11±4 (n=24)	29±17 (n=50)
PT per length of sample (/10 mm)	7±3 (n=108)	5±1 (n=34)	5±1 (n=24)	7±3 (n=50)
Interface Inflammation Score††				
Grade 0	238 (61%)	66 (74%)	151 (63%)	21 (33%)
Grade 1	106 (27%)	20 (23%)	69 (29%)	17 (27%)
Grade 2	28 (7%)	2 (2%)	16 (7%)	10 (16%)
Grade 3	16 (4%)	–	1 (1%)	15 (24%)
Grade 4	1 (1%)	1 (1%)	–	–
Lobular inflammation†† (Ishak)				
Grade 0	117 (30%)	57 (64%)	55 (23%)	5 (8%)
Grade 1	207 (53%)	27 (30%)	135 (57%)	45 (71%)
Grade 2	54 (14%)	4 (5%)	42 (18%)	8 (13%)
Grade 3	11 (3%)	–	6 (2%)	5 (8%)
Grade 4	1 (<1%)	1 (1%)	–	–
Lobular inflammation†† (NASH CRN)				
Grade 0	122 (31%)	56 (63%)	55 (23%)	11 (17%)
Grade 1	202 (52%)	28 (32%)	135 (57%)	39 (62%)
Grade 2	53 (14%)	4 (4%)	42 (18%)	7 (11%)
Grade 3	13 (3%)	1 (1%)	6 (2%)	6 (10%)
Grade 4	–	–	–	–
Portal Inflammation Score†† (Ishak)				
Grade 0	81 (21%)	28 (32%)	42 (18%)	11 (17%)
Grade 1	200 (51%)	43 (48%)	132 (55%)	25 (40%)
Grade 2	90 (23%)	17 (19%)	54 (23%)	19 (30%)
Grade 3	16 (4%)	–	10 (4%)	6 (10%)
Grade 4	3 (1%)	1 (1%)	–	2 (3%)
Steatosis grade†† (NASH CRN)				
Grade 0	91 (23%)	30 (34%)	3 (1%)	58 (92%)
Grade 1	121 (31%)	38 (43%)	79 (33%)	4 (6%)
Grade 2	90 (23%)	9 (10%)	80 (34%)	1 (2%)
Grade 3	88 (23%)	12 (13%)	76 (32%)	–
Fibrosis stage†† (NASH CRN)				
Stage 0	78 (20%)	21 (24%)	4 (2%)	53 (85%)
Stage 1	56 (14%)	17 (19%)	37 (15%)	2 (3%)
Stage 2	74 (19%)	13 (14%)	61 (26%)	–
Stage 3	130 (34%)	24 (27%)	102 (43%)	4 (6%)
Stage 4	51 (13%)	14 (16%)	33 (14%)	4 (6%)
Fibrosis stage†† (Ishak)				
Stage 0	57 (15%)	25 (28%)	22 (9%)	10 (16%)

Continued

Table 1 Continued

	All (n=390)	MASLD (n=89)	MASH (n=238)	AIH (n=63)
Stage 1	46 (12%)	10 (11%)	28 (12%)	8 (13%)
Stage 2	72 (18%)	13 (15%)	50 (21%)	9 (14%)
Stage 3	86 (22%)	18 (20%)	58 (24%)	10 (16%)
Stage 4	55 (14%)	7 (8%)	38 (16%)	10 (16%)
Stage 5	52 (13%)	6 (7%)	15 (6%)	10 (16%)
Stage 6	38 (10%)	9 (10%)	23 (10%)	6 (9%)

Shown as median (IQR).
*Shown as mean±SD.
†Shown as n (%).
AIH, autoimmune hepatitis; BMI, body mass index; BP, blood pressure; MASH, metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; NASH CRN, NASH Clinical Research Network; PT, portal tracts; WSI, whole slide image.

for MASLD and 7.9% (IQR 4.2) for MASH (p<0.0001). Vesicle size also differed between MASLD and MASH at all distances from PT.

Comparison of AI steatosis grading with manual pathology reads

Expert pathologist grading for steatosis did not differ significantly between MASLD and MASH (online supplemental figure 3). Both AI-derived metrics for steatosis quantification showed significant correlation with individual and consensus pathologist steatosis grading in lobular regions in both MASLD (SF r_s=0.61 and VHF r_s=0.64; p<0.0001) and MASH (SF r_s=0.70 and VHF r_s=0.72; p<0.0001).

Quantification of inflammation in MASLD/MASH and AIH

The AI system identified 1.5 million inflammatory cells at PT, 0.4 million inflammatory cells in lobular regions and 0.11 million cells in interface regions across all 390 WSI.

Inflammatory cells were detected in biopsies from MASLD or MASH individuals (mean ICN: 4300 per WSI) and AIH cases (mean ICN: 12 000 per WSI) (online supplemental table 3). The IB was higher in AIH compared with MASLD/MASH at PT (2700 cells/mm^2 (IQR 2800) vs 1400 cells/mm^2 (IQR 1800), p<0.0001) and in lobular regions (60 cells/mm^2; (IQR 90) vs 40 cells/mm^2; (IQR 50), p<0.001) (figure 4).

Regional distribution of inflammation

There was significantly more inflammation in all regions in AIH cases compared with MASLD/MASH.

The regional distribution was similar: maximal ICN and IB were in the PT for both AIH and MASLD/MASH (online supplemental table 3, figure 5). In the immediate vicinity of the PT, up to 100μm away, IB in AIH was 150 cells/mm^2 (IQR 130) (figure 5). Lower levels of IB were measured in MASLD/MASH cases on the same region (IB: 80 cells/mm^2 (IQR 90), p<0.0001).

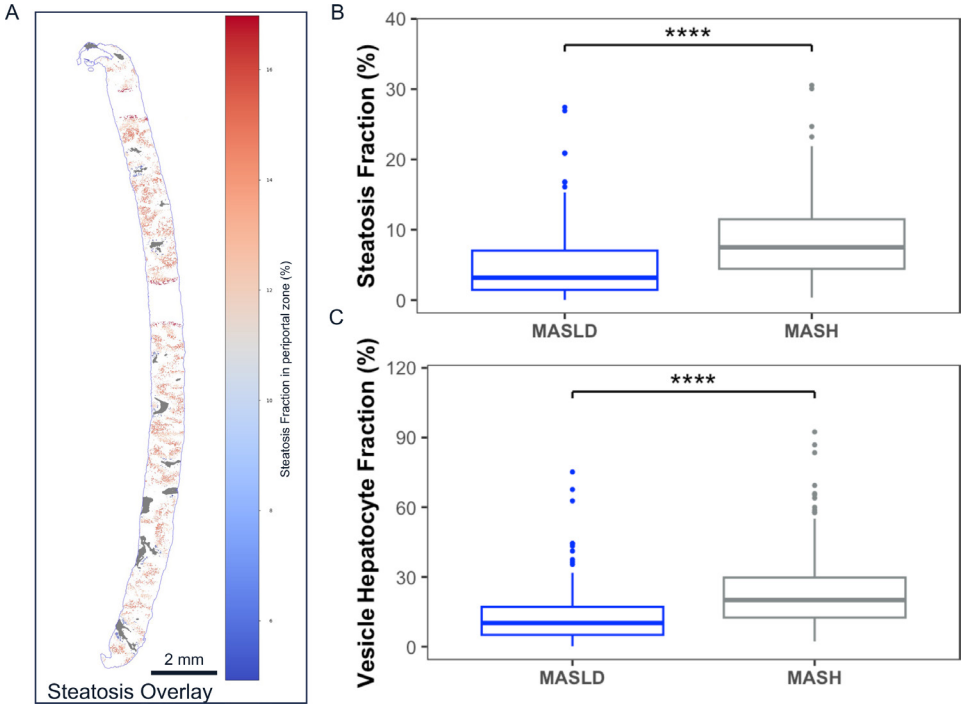

Figure 2 Steatosis in MASLD (n=89) and MASH (n=238). (A) Example case with an interactive overlay of steatosis fraction represented on a colourimetric scale (blue-red) at specific regions from PT (grey). (B) Steatosis fraction (%) in lobular regions in MASLD and MASH. (C) Vesicle to hepatocyte fraction (%) in lobular regions in MASLD and MASH. MASH metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; PT, portal tracts.

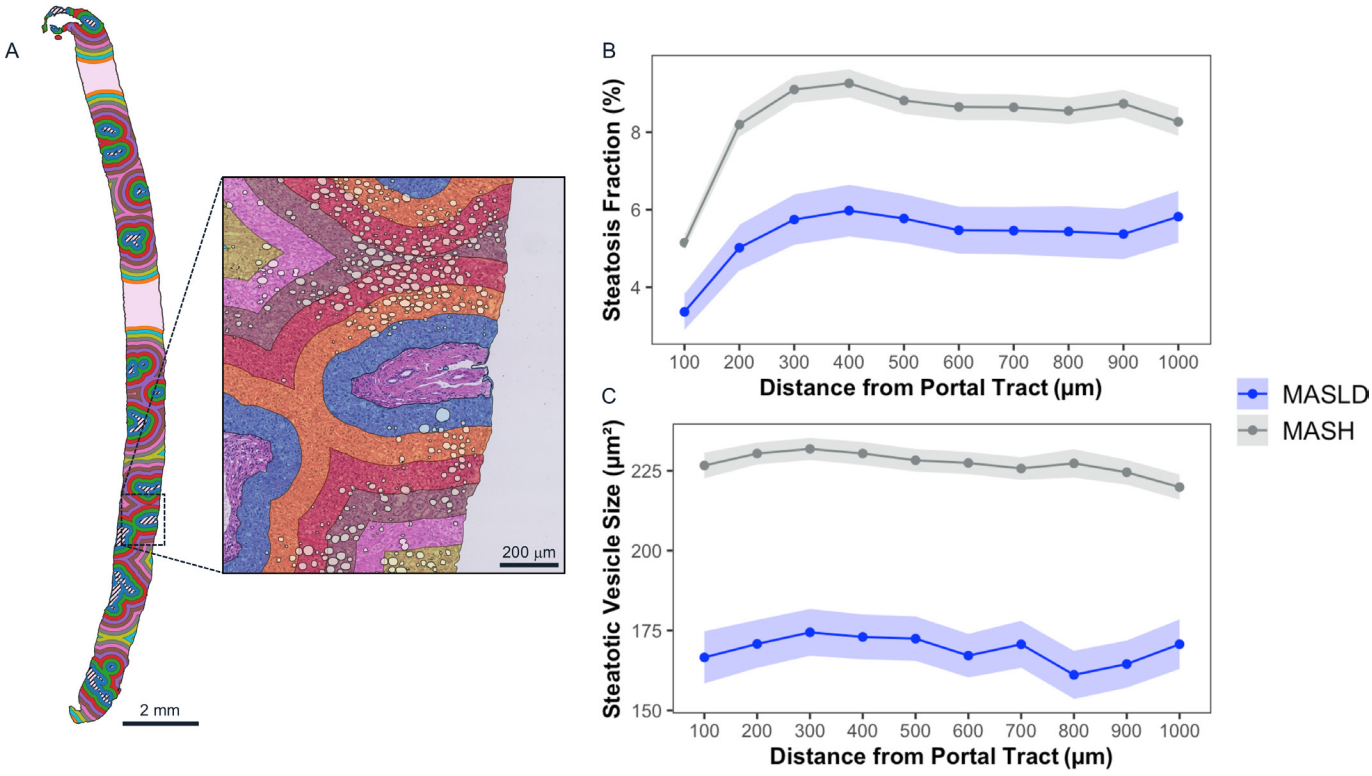

Figure 3 Regional analysis of steatosis in MASLD (n=89) and MASH (n=238). (A) Steatotic vesicles were quantified within areas defined by regions positioned equidistant from the PT, at intervals of 100 µm, in a case with MASLD. (B) Steatosis fraction (%) in MASLD and MASH. (C) Size (µm^2) of steatotic vesicles in MASLD and MASH. Median values are represented as a solid line, and the upper/lower IQR is represented as a shaded error band. MASH metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; PT, portal tracts.

The ICN was also higher in AIH (390 cells (IQR 580) compared with MASLD/MASH (140 cells (IQR 220), p<0.0001) in this region.

In both AIH and MASLD/MASH, the IB was lower at distances of 200–600 μm from the PT (2.2-fold lower relative to the portal region for AIH, and 1.8-fold for MASLD/MASH). However, the IB then increased again further away in lobular regions in AIH (700–900 μm; 1.3-fold higher relative to the IB at 200 μm from PT). In contrast, the ICN decreased gradually across lobular regions in both AIH and MASLD/MASH (figure 5).

Association with scoring of inflammation from manual pathology reads

Grading of inflammation showed good agreement between pathologists for the PT and periportal regions but was less aligned for lobular regions (online supplemental figure 4). Severity of portal inflammation from manual scoring correlated well with ICN count at PT (r_s=0.46–0.73; p=0.0067-p<0.0001), but only weakly with IB (r_s=0.16–0.28; all non-significant). Pathologist scorings for lobular inflammation (both Ishak and NASH CRN) correlated weakly with both IB and ICN (r_s=0.15–0.35; all non-significant).

Manual grading from individual pathologists showed similar trends. Increasing inflammation grade correlated with increasing ICN and IB at the PT and interface regions. For all regions, severity of inflammation from manual pathology correlated better with ICN than with IB (figure 6)—ICN at PT correlated well with severity of interface hepatitis in AIH (r_s=0.68, p<0.0001) and MASLD/MASH (r_s=0.56, p<0.0001) (online supplemental figure 5). For equivalent grades of portal inflammation, the ICN was up to 3.1-fold higher in AIH than in MASLD/MASH, and this was seen at portal, interface, and lobular regions (figure 6, online supplemental table 5).

DISCUSSION

Our AI system to identify and quantify PT, steatosis, and inflammation in 390 MASLD/MASH and AIH cases revealed three key findings. First, MASH showed a 2.3-fold increase in quantity in steatosis, with larger vesicles than MASLD. Second, inflammation was higher in all AIH cases compared with MASLD/MASH, and although regional distribution was similar between the liver diseases, inflammation was higher at the PT and interface. Third, overall, greater degrees of inflammation were detected in the AIH cases, even for equivalent manual grades.

The quantification of both steatotic vesicles and inflammatory cells in relation to the PT anatomical feature differentiates this AI system from previous tools that mainly calculate NASH CRN scores.[21] This regional differentiation has the potential to support pathologists in the differentiation of liver disease and in monitoring disease progression and response to treatment in clinical trials more consistently and less subjectively.

PT detection

Manual PT identification is challenging due to their variable appearance. This is particularly relevant for highly fibrotic cases where vascular structures merge together and PT expand, resulting in an unclear boundary.[34] As well as helping the pathologist recognise PT in this situation, our system can perform automatic quality control on liver biopsy samples by quantifying PT number. Similarly, as biopsy length is often used as an alternative

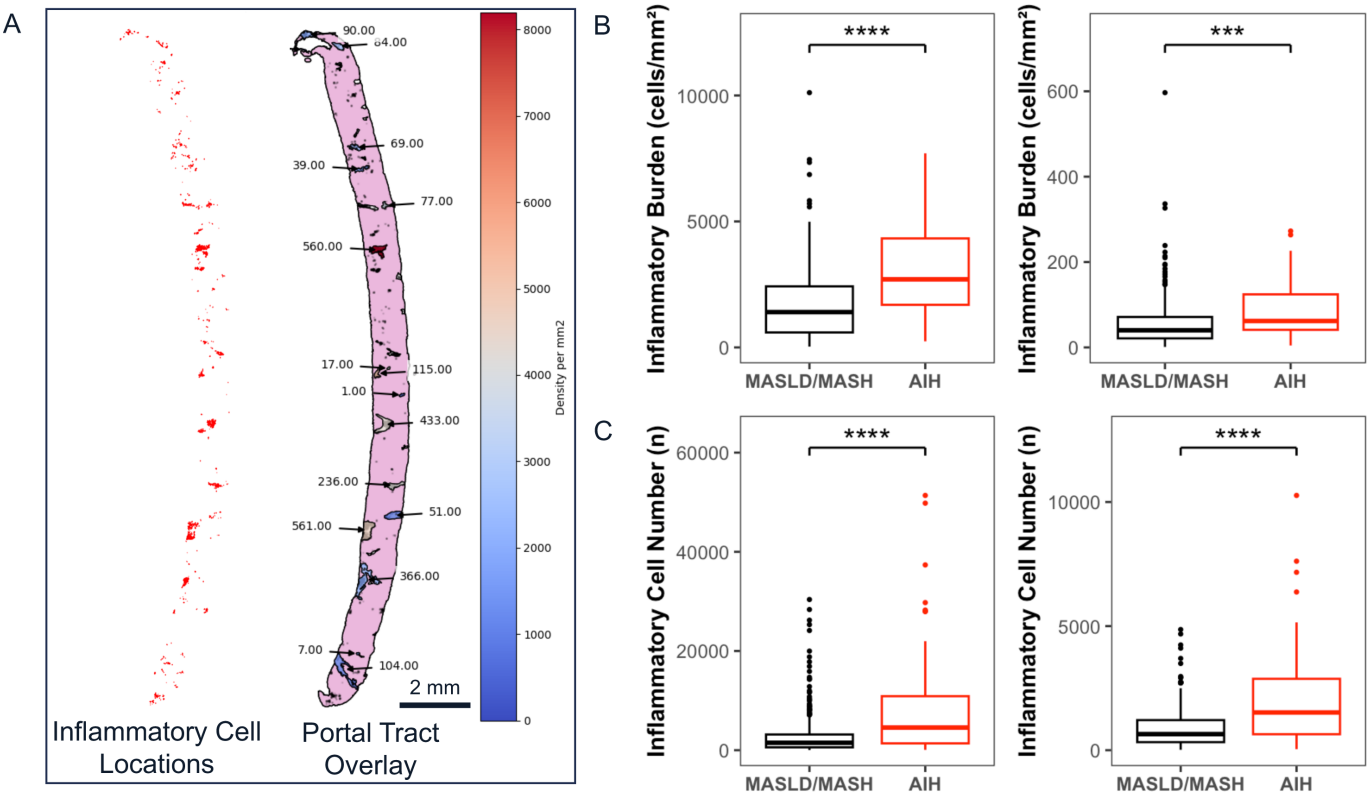

Figure 4 Comparison of inflammation in MASLD/MASH (n=327) versus AIH (n=63). (A) Example of a case with MASLD; both segmented inflammatory cells (ICN) and inflammatory burden (IB) were quantified and overlayed with the anatomical location of the PT and represented with a colourimetric scale. (B) The IB (cells/mm^2) and (C) ICN in portal (left) and lobular (right) regions in MASLD/MASH and AIH. AIH, autoimmune hepatitis; MASH metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; PT, portal tracts.

to PT count for sample adequacy, this AI system could assist by automatically calculating biopsy length and PT per linear cm.

Steatosis quantification

In the NASH CRN scoring system, steatosis is scored by assessing percentage of hepatocytes that contain lipid droplets, equivalent to our VHF. However, microvesicular steatosis is excluded which is in line with the NAS grading system.[35] Future work aimed at distinguishing macro from microvesicular steatosis will be of particular interest, as we may have underestimated the burden of steatosis. Overall, SF was higher in MASH than MASLD across periportal and lobular regions. Our system also revealed a difference in the steatotic vesicle size between MASLD and MASH in the periportal and lobular areas—a novel finding.[36 37]

AI metrics and pathologist steatosis grading broadly aligned for MASLD and MASH cases. Average SF was under 33%, consistent with the low liver fat content by MRI (MRI-Proton Density Fat Fraction (<10%)) that was recorded for this population,[38] but was inconsistent with manual steatosis scores >1, which is classified as >33%. Overestimation of steatosis is generally reported in manual scoring, especially in severe cases[39] and as liver steatosis resolution is relevant for MASH drug efficacy,[40 41] accurate quantification of fat is a potential application for this AI system.

Inflammation quantification

With respect to inflammation, as expected, total levels were higher in AIH compared with MASLD/MASH and correlated well with manual grades in cases of high inflammatory cell burden. Despite equivalent Ishak/NASH CRN scores, in both AIH and MASLD/MASH, IB was greatest in interface regions and at 700–900μm from the PT. This regional distribution may reflect the biology of acinar zones, with higher levels of inflammation in the immune-privileged zone 1 and increased necroinflammation in zone 3.[42] In MASLD/MASH, increasing ICN and IB in portal and interface regions correlated with higher pathologist scores, but low lobular IB did not correlate with pathologist scores. This latter result reflects the low inter-rater reliability for lobular inflammation found in NASH CRN scoring.[43]

Clinical implications

Several digital pathology models have been developed to quantify NASH CRN histological criteria.[20] The qFIBS tool provides clinically useful measures of septal parameters (qFibrosis) in unstained slides; these can distinguish regressive from progressive fibrosis across five lobular regions.[22 44] AIM-MASH, trained at large scale on H&E/TRC-stained slides from clinical trials, provides NASH CRN scores and AI overlays for steatosis, ballooning, lobular inflammation and fibrosis, but without structured regional breakdown.[21] FibroNest[45] and MorphoQuant[46] also quantify all NASH CRN features without structured regional breakdown. The MUSA-UNet model, trained on a similar number of WSI (30) as herein (31), quantified PT and Scheuer fibrosis stage in transplant biopsies.[19] Overall, these solutions represent important advances but do not account for PT detection and characterisation that determines the location, boundaries, and granular regional quantification of steatosis and inflammation across liver biopsies.

An application of our AI system would be in accurate assessment of histological endpoints in clinical trials. Evidence of

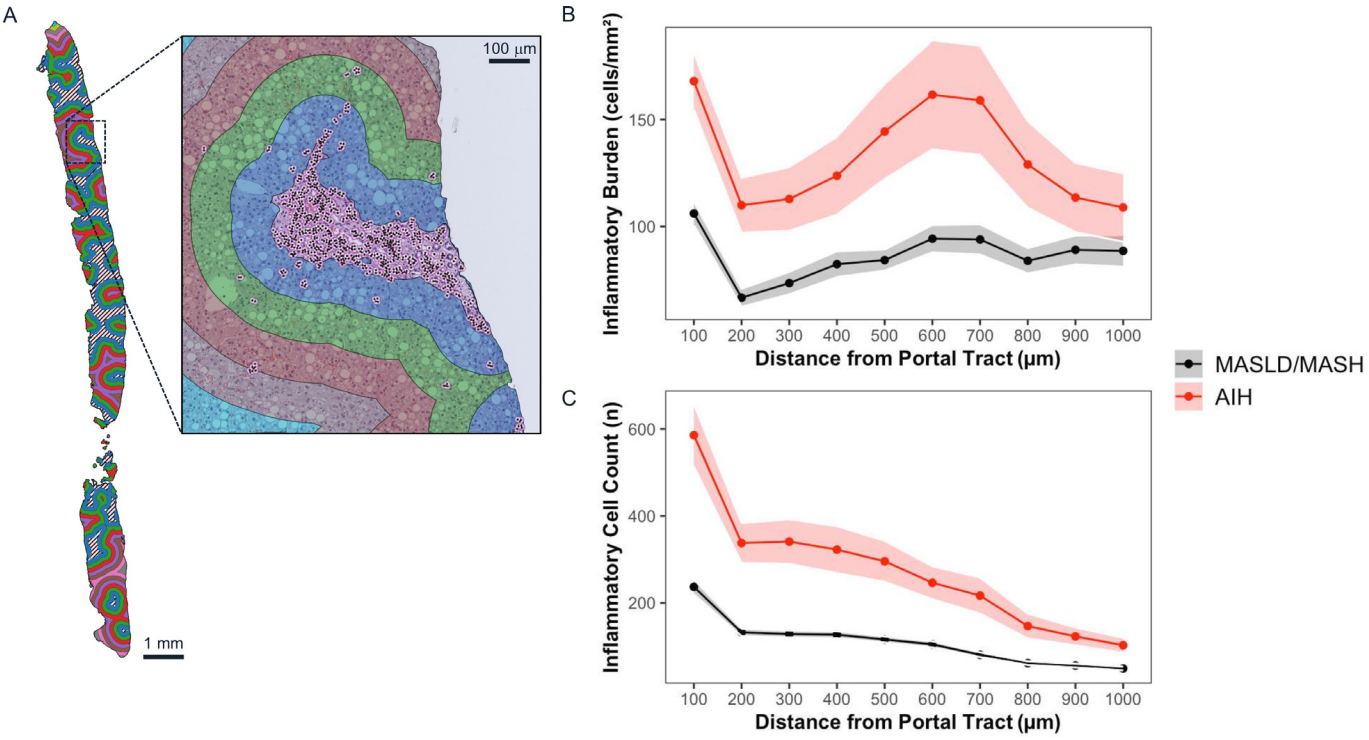

Figure 5 Regional analysis performed across WSI for inflammatory burden and inflammatory cell numbers in MASLD/MASH (n=327) and AIH (n=63). (A) Inflammatory metrics were extracted from individual regions positioned equidistant from the PT, at intervals of 100 µm, in an AIH liver biopsy. (B) Inflammatory burden and (C) inflammatory cell number in MASLD/MASH and AIH, across equidistant regions. Median values are represented as a solid line and the upper/lower IQR is represented as a shaded error band. AIH, autoimmune hepatitis; MASH, metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; PT, portal tracts; WSI, whole slide image.

asymmetric improvements between lobular and portal inflammation in MASH was observed in the phase 2 *FLINT* trial with obeticholic acid; only lobular inflammation improved.[47] Similarly, in the *PIVENS* trial, with Vitamin E/pioglitazone, a significant improvement in steatosis and lobular inflammation was reported while no significant change was observed in portal inflammation.[9] Detailed analysis of inflammatory changes across zonal regions may enhance such work; for example, the

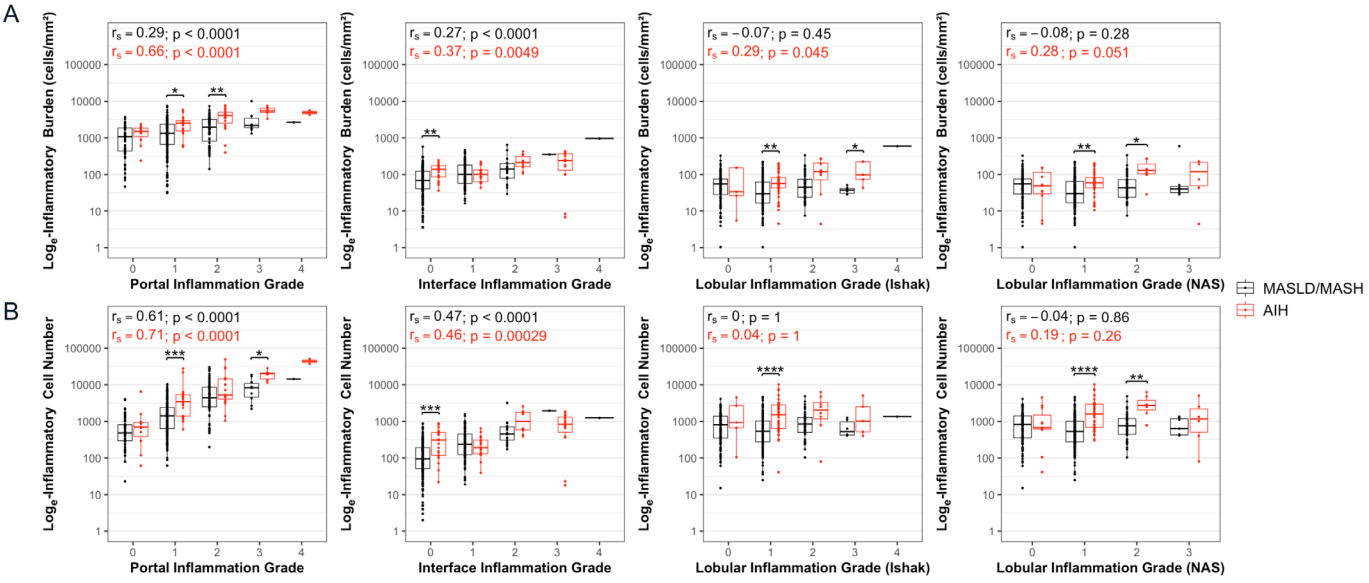

Figure 6 Inflammation quantified by AI in relation to manual grading by a single pathologist for MASLD/MASH (n=327) and for AIH (n=63). (A) Grades relative to IB. (B) Shows Ishak grades relative to ICN. Grades shown are based on Ishak scoring at portal and interface regions, and by either Ishak or NASH CRN scoring at the lobules. AI, artificial intelligence; AIH, autoimmune hepatitis; IB, inflammatory burden; ICN, inflammatory cell number; MASH metabolic dysfunction-associated steatohepatitis; MASLD, metabolic dysfunction-associated steatotic liver disease; NASH CRN, NASH Clinical Research Network.

antioxidant effect of vitamin E may be more pronounced in periportal regions, where oxidative stress is highest in progressive disease. The therapies in current clinical trials in MASH have antisteatotic effects and mechanisms of action[48] but anti-inflammatory modes of action are also being considered.[49]

Portal inflammation is associated with disease progression in MASLD and is a consideration in disease resolution endpoints in MASH trials[1–3 50–52] . In AIH, recent consensus recommendations have suggested that interface hepatitis is one feature in the composite diagnosis.[53] AIH features, such as interface hepatitis, can coexist with features of both MASLD and MASH. Using our system, the accurate quantification of inflammation at the interface region and graphical identification of inflammation will support pathologist scoring and identify interface hepatitis, avoid misclassification and flag potential incidental findings (eg, PBC/PSC).

Limitations and future directions

The quantification of both steatotic vesicles and inflammatory cells, in relation to the PT anatomical feature, differentiates this AI system from previous tools and its validation in four independent cohorts is a strength of this work. However, there remain several limitations. First, our work was cross-sectional and would have benefited from longitudinal samples and prospective validation against outcomes. Future work prospectively evaluating disease progression or treatment response is warranted. Second, while the curated clinical trial datasets herein allowed us to develop an accurate and robust model, future iterations incorporating datasets with lower quality slides (eg, with tissue folding, out-of-focus regions, variable staining) will be important to validate our findings. Such real-world cases could include complex cases of MASLD/MASH with concurrent AIH to validate our findings. Third, in our predominantly paediatric AIH cohort, we consistently found elevated levels of inflammation compared with MASLD/MASH. Paediatric AIH is considered to have fewer disease features compared with adult AIH,[54] so this finding requires validation in prospective studies. Finally, the present study focuses on H&E-stained slides to reflect the scoring of steatosis and inflammation. Incorporating ballooning and fibrosis detection and other histological stains would strengthen the overall outputs and can be incorporated into future pipelines.

CONCLUSION

PT characterisation is fundamental to establishing sample adequacy of liver biopsies and to assessing the pathological abnormalities such as staging fibrosis, identification of interface hepatitis and zonal necrosis. This AI system generated detailed visual overlays to assist in the above, as well as producing automated quantification of PT, inflammation and steatosis across portal, interface, and lobular regions. Such AI systems could improve the reproducibility and accuracy of liver biopsy evaluation and consensus decisions in clinical trials.

Furthermore, while the use of non-invasive approaches to quantify liver inflammation is becoming more prevalent, the low levels of IB and ICN that are present in MASLD and MASH will require the high degree of sensitivity afforded by AI to measure the small effect sizes that are likely when assessing the impact of anti-inflammatory treatments.

Permission to reproduce material from other sources

Anonymised individual patient data can be shared upon request or as required by law and/or regulation with qualified external researchers. Approval of such requests is at the discretion of the study sponsors and is dependent on the nature of the request, the merit of the research proposed, the availability of the data and the intended use of the data.

Handling editor Vikram Deshpande.

Acknowledgements We would like to thank the patients recruited into the studies used in this article for their participation. Study 1 was registered with registered with clinicaltrials.gov (NCT03551522). Study 2 was registered with UMIN clinical trials registry (UMIN000026145). Study 3 was registered with the ISRCTN registry (ISCRTN39463479) and the National Institute for Health Research portfolio (15912). The collection of study 3 pathology data was funded by Innovate UK (project number: 101679). Study 4 was registered with clinicaltrials.gov (NCT03198104) and funded by the Eureka Eurostars 2 Grant (E!10124).

Contributors Guarantor—CRL. Conceptualisation—DW/AM (equal). Writing (original draft)—DW (lead), AM (supporting), SL (supporting). Data curation—AM (lead), CB/KHDW/KF/EF/TJK/RG (supporting). Investigation—DW (lead), AM/CB (supporting). Formal analysis—AM/CB (equal). Preparation—DW (lead), AM/CB/SL (supporting), visualisation—AM (lead), CB/DW (supporting). Review and editing—DW/HTB (equal), AM/HTB/SL/PA/KF/EF/TJK/RK/PW/CRL/PB/RG (supporting).

Funding This paper presents data from independent research funded by Innovate UK (Project number: 101679) and the Eureka Eurostars 2 Grant (E!10124).

Competing interests DW/AM/CB/HTB/SL/KH/PA/RK/PW and CRL are employees at Perspectum Ltd. KF/EF/TJK and RG are consultants for Perspectum Ltd. All other coauthors have no conflicts of interest to declare relevant to this work. Manuscript includes use of data generated via collaboration with; Yokohama City University Hospital, CymaBay Therapeutics Inc (acquired by Gilead Sciences), University of Birmingham and Children's Memorial Health Institute in Warsaw (IPCZD). Two trials of which data were utilised for this study were sponsored by Perspectum Ltd (Study 3 and 4).

Patient consent for publication Not applicable.

Ethics approval All parent studies were conducted in accordance with the ethical principles of the Declaration of Helsinki 2013 and the Good Clinical Practice guidelines and all studies were approved by ethical committees (institutional review at trial sites for NCT03551522, Ethics committee at Yokohama City Hospital for UMIN000026145, West Midlands—Black Country Research Ethics Committee 14/WM/0010 for ISRCTN39463479, Komisja Bioetyczna przy Instytucie Pomnik-Centrum Zdrowia Dziecka 11/KBE/2016 for NCT03198104). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer-reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. The portal tract identification system described for the first time in this study is the intellectual property of Perspectum Ltd and the subject of the Patent Cooperation Treaty PCT/IB2025/0530307. Summary data are included in the manuscript or uploaded as online supplemental information.

ORCID iDs
Dylan Windell https://orcid.org/0000-0002-8188-6573
Timothy James Kendall https://orcid.org/0000-0002-4174-2786
Caitlin Rose Langford https://orcid.org/0000-0003-0730-7091

REFERENCES

1 Harrison SA, Gawrieh S, Roberts K, *et al*. Prospective evaluation of the prevalence of non-alcoholic fatty liver disease and steatohepatitis in a large middle-aged US cohort. *J Hepatol* 2021;75:284–91.

2 Rinella ME, Lazarus JV, Ratziu V, *et al*. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology* 2023;78:1966–86.
3 Younossi ZM, Golabi P, Paik JM, *et al*. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review. *Hepatology* 2023;77:1335–47.
4 Le P, Tatar M, Dasarathy S, *et al*. Estimated Burden of Metabolic Dysfunction–Associated Steatotic Liver Disease in US Adults, 2020 to 2050. *JAMA Netw Open* 2025;8:e2454707.
5 Dalekos GN, Gatselis NK, Zachou K, *et al*. NAFLD and autoimmune hepatitis: Do not judge a book by its cover. *Eur J Intern Med* 2020;75:1–9.
6 Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, *et al*. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology* 2023;77:1797–835.
7 Harrison SA, Bedossa P, Guy CD, *et al*. A Phase 3, Randomized, Controlled Trial of Resmetirom in NASH with Liver Fibrosis. *N Engl J Med* 2024;390:497–509.
8 Francque SM, Bedossa P, Ratziu V, *et al*. A Randomized, Controlled Trial of the Pan-PPAR Agonist Lanifibranor in NASH. *N Engl J Med* 2021;385:1547–58.
9 Sanyal AJ, Chalasani N, Kowdley KV, *et al*. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. *N Engl J Med* 2010;362:1675–85.
10 Sanyal AJ, Brunt EM, Kleiner DE, *et al*. Endpoints and clinical trial design for nonalcoholic steatohepatitis. *Hepatology* 2011;54:344–53.
11 Harrison SA, Bashir MR, Guy CD, *et al*. Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet* 2019;394:2012–24.
12 Brunt EM, Clouston AD, Goodman Z, *et al*. Complexity of ballooned hepatocyte feature recognition: Defining a training atlas for artificial intelligence-based imaging in NAFLD. *J Hepatol* 2022;76:1030–41.
13 Fryer E, Wang LM, Verrill C, *et al*. How often do our liver core biopsies reach current definitions of adequacy? *J Clin Pathol* 2013;66:1087–9.
14 Langford CR, Goldinger MH, Treanor D, *et al*. Improved pathology reporting in NAFLD/NASH for clinical trials. *J Clin Pathol* 2022;75:73–5.
15 Sporea I, Gherhardt D, Popescu A, *et al*. Does the size of the needle influence the number of portal tracts obtained through percutaneous liver biopsy? *Ann Hepatol* 2012;11:691–5.
16 Brunt EM. Liver biopsy reliability in clinical trials: Thoughts from a liver pathologist. *J Hepatol* 2020;73:1310–2.
17 Davison BA, Harrison SA, Cotter G, *et al*. Suboptimal reliability of liver biopsy evaluation has implications for randomized clinical trials. *J Hepatol* 2020;73:1322–32.
18 Perveen S, Shahbaz M, Keshavjee K, *et al*. A Systematic Machine Learning Based Approach for the Diagnosis of Non-Alcoholic Fatty Liver Disease Risk and Progression. *Sci Rep* 2018;8:2112.
19 Yu H, Sharifai N, Jiang K, *et al*. Artificial intelligence based liver portal tract region identification and quantification with transplant biopsy whole-slide images. *Comput Biol Med* 2022;150:106089.
20 Ratziu V, Hompesch M, Petitjean M, *et al*. Artificial intelligence-assisted digital pathology for non-alcoholic steatohepatitis: current status and future directions. *J Hepatol* 2024;80:335–51.
21 Pulaski H, Harrison SA, Mehta SS, *et al*. Clinical validation of an AI-based pathology tool for scoring of metabolic dysfunction-associated steatohepatitis. *Nat Med* 2025;31:315–22.
22 Abdurrachim D, Lek S, Ong CZL, *et al*. Utility of AI digital pathology as an aid for pathologists scoring fibrosis in MASH. *J Hepatol* 2025;82:898–908.
23 Tsai H-W, Chiou C-Y, Yang W-J, *et al*. Lymphocyte-Infiltrated Periportal Region Detection With Structurally-Refined Deep Portal Segmentation and Heterogeneous Infiltration Features. *IEEE Open J Eng Med Biol* 2024;5:261–70.
24 Andersson A, Kelly M, Imajo K, *et al*. Clinical Utility of Magnetic Resonance Imaging Biomarkers for Identifying Nonalcoholic Steatohepatitis Patients at High Risk of Progression: A Multicenter Pooled Data and Meta-Analysis. *Clin Gastroenterol Hepatol* 2022;20:2451–61.
25 Socha P, Shumbayawonda E, Roy A, *et al*. Quantitative digital pathology enables automated and quantitative assessment of inflammatory activity in patients with autoimmune hepatitis. *J Pathol Inform* 2024;15:100372.
26 Dennis A, Kelly MD, Fernandes C, *et al*. Correlations Between MRI Biomarkers PDFF and cT1 With Histopathological Features of Non-Alcoholic Steatohepatitis. *Front Endocrinol (Lausanne)* 2020;11:575843.
27 Imajo K, Tetlow L, Dennis A, *et al*. Quantitative multiparametric magnetic resonance imaging can aid non-alcoholic steatohepatitis diagnosis in a Japanese cohort. *World J Gastroenterol* 2021;27:609–23.
28 McDonald N, Eddowes PJ, Hodson J, *et al*. Multiparametric magnetic resonance imaging for quantitation of liver disease: a two-centre cross-sectional observational study. *Sci Rep* 2018;8:9189.
29 Janowski K, Shumbayawonda E, Cheng L, *et al*. Quantitative multiparametric MRI as a non-invasive stratification tool in children and adolescents with autoimmune liver disease. *Sci Rep* 2021;11:15261.
30 Alkhouri N, Beyer C, Shumbayawonda E, *et al*. Decreases in cT1 and liver fat content reflect treatment-induced histological improvements in MASH. *J Hepatol* 2025;82:438–45.
31 Neuberger J, Patel J, Caldwell H, *et al*. Guidelines on the use of liver biopsy in clinical practice from the British Society of Gastroenterology, the Royal College of Radiologists and the Royal College of Pathology. *Gut* 2020;69:1382–403.
32 Knodell RG, Ishak KG, Black WC, *et al*. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. *Hepatology* 1981;1:431–5.
33 Brunt EM, Kleiner DE, Wilson LA, *et al*. Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology* 2011;53:810–20.
34 Vollmar B, Wolf B, Siegmund S, *et al*. Lymph vessel expansion and function in the development of hepatic fibrosis and cirrhosis. *Am J Pathol* 1997;151:169–75.
35 Tandra S, Yeh MM, Brunt EM, *et al*. Presence and significance of microvesicular steatosis in nonalcoholic fatty liver disease. *J Hepatol* 2011;55:654–9.
36 Dempsey JL, Ioannou GN, Carr RM. Mechanisms of Lipid Droplet Accumulation in Steatotic Liver Diseases. *Semin Liver Dis* 2023;43:367–82.
37 Gluchowski NL, Becuwe M, Walther TC, *et al*. Lipid droplets and liver disease: from basic biology to clinical implications. *Nat Rev Gastroenterol Hepatol* 2017;14:343–55.
38 Caussy C, Reeder SB, Sirlin CB, *et al*. Noninvasive, Quantitative Assessment of Liver Fat by MRI-PDFF as an Endpoint in NASH Trials. *Hepatology* 2018;68:763–72.
39 Hall AR, Dhillon AP, Green AC, *et al*. Hepatic steatosis estimated microscopically versus digital image analysis. *Liver Int* 2013;33:926–35.
40 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* 2015;149:389–97.
41 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* 2024;8:e0423.
42 Krishna M. Patterns of necrosis in liver disease. *Clin Liver Dis (Hoboken)* 2017;10:53–6.
43 Fukusato T, Fukushima J, Shiga J, *et al*. Interobserver variation in the histopathological assessment of nonalcoholic steatohepatitis. *Hepatol Res* 2005;33:122–7.
44 Naoumov NV, Kleiner DE, Chng E, *et al*. Digital quantitation of bridging fibrosis and septa reveals changes in natural history and treatment not seen with conventional histology. *Liver Int* 2024;44:3214–28.
45 Petitjean L, Chen L, Mattis A, *et al*. Evaluation of the performance of a novel digital pathology method for the continuous quantification of steatosis, ballooning and inflammation in liver biopsies and its correlation with NASH-CRN scores in patients with NASH. *J Hepatol* 2022;77:S439.
46 De Rudder M, Bouzin C, Nachit M, *et al*. Automated computerized image analysis for the user-independent evaluation of disease severity in preclinical models of NAFLD/NASH. *Lab Invest* 2020;100:147–60.
47 Neuschwander-Tetri BA, Loomba R, Sanyal AJ, *et al*. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet* 2015;385:956–65.
48 Hartman ML, Loomba R, Lawitz EJ, *et al*. Consistent improvements in liver histology across subgroups in a post hoc analysis of the SYNERGY-NASH trial with tirzepatide. *JHEP Rep* 2025;7:101472.
49 Harrison SA, Alkhouri N, Ortiz-Lasanta G, *et al*. A phase IIb randomised-controlled trial of the FFAR1/FFAR4 agonist icosabutate in MASH. *J Hepatol* 2025;83:293–303.
50 Brunt EM, Kleiner DE, Wilson LA, *et al*. Portal chronic inflammation in nonalcoholic fatty liver disease (NAFLD): a histologic marker of advanced NAFLD-Clinicopathologic correlations from the nonalcoholic steatohepatitis clinical research network. *Hepatology* 2009;49:809–20.
51 Dhingra S, Mahadik JD, Tarabishy Y, *et al*. Prevalence and clinical significance of portal inflammation, portal plasma cells, interface hepatitis and biliary injury in liver biopsies from patients with non-alcoholic steatohepatitis. *Pathology (Phila)* 2022;54:686–93.
52 Rakha EA, Adamson L, Bell E, *et al*. Portal inflammation is associated with advanced histological changes in alcoholic and non-alcoholic fatty liver disease. *J Clin Pathol* 2010;63:790–5.
53 Lohse AW, Sebode M, Bhathal PS, *et al*. Consensus recommendations for histological criteria of autoimmune hepatitis from the International AIH Pathology Group: Results of a workshop on AIH histology hosted by the European Reference Network on Hepatological Diseases and the European Society of Pathology. *Liver Int* 2022;42:1058–69.
54 Mau B, Hakar M, Lin HC, *et al*. A review of histopathologic features of pediatric autoimmune liver disease. *Clin Liver Dis (Hoboken)* 2022;20:116–9.