Title

Artificial intelligence-assisted digital pathology for non-alcoholic steatohepatitis: current status and future directions.

Journal

Journal of Hepatology, 80(2)

Authors

Ratziu, Vlad

Hompesch, Marcus

Petitjean, Mathieu

et al.

Publication Date

2024-02-01

DOI

10.1016/j.jhep.2023.10.015

Peer reviewed

HHS Public Access

Author manuscript

J Hepatol. Author manuscript; available in PMC 2025 February 13.

Published in final edited form as:

J Hepatol. 2024 February ; 80(2): 335–351. doi:10.1016/j.jhep.2023.10.015.

Artificial intelligence-assisted digital pathology for non-alcoholic steatohepatitis: current status and future directions

Vlad Ratziu[1,*], **Marcus Hompesch**[2], **Mathieu Petitjean**[3], **Cindy Serdjebi**[4], **Janani S. Iyer**[5], **Anil V. Parwani**[6], **Dean Tai**[7], **Elisabetta Bugianesi**[8], **Kenneth Cusi**[9], **Scott L. Friedman**[10], **Eric Lawitz**[11], **Manuel Romero-Gómez**[12], **Detlef Schuppan**[13,14], **Rohit Loomba**[15], **Valérie Paradis**[16], **Cynthia Behling**[17], **Arun J. Sanyal**[18]

[1]Sorbonne Université, ICAN Institute for Cardiometabolism and Nutrition, Hospital Pitié-Salpêtrière, INSERM UMRS 1138 CRC, Paris, France

[2]ProSciento, San Diego, CA, USA

[3]PharmaNest, Inc., Princeton, NJ, USA

[4]Biocellvia, Marseille, France

[5]PathAI, Inc., Boston, MA, USA

[6]Department of Pathology, The Ohio State University, Columbus, OH, USA

[7]HistoIndex Pte Ltd, Singapore

[8]Department of Medical Sciences, University of Turin, Turin, Italy

[9]Division of Endocrinology, Diabetes and Metabolism, University of Florida, Gainesville, FL, USA

[10]Division of Liver Diseases, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[11]Texas Liver Institute, University of Texas Health San Antonio, San Antonio, TX, USA

[*]Corresponding author. Address: Sorbonne Université, Institute for Cardiometabolism and Nutrition, Hospital Pitié-Salpêtrière, Paris, France. vlad.ratziu@inserm.fr (V. Ratziu).

Authors' contributions

VR conceived and drafted the manuscript. J.S.I. contributed specifically to writing and critical revisions to discussion relating to AIM-NASH and NASH explore; M.P. contributed specifically to writing and critical revisions to discussion relating to FibroNest™; C.S. contributed specifically to writing and critical revisions to discussion relating to MorphoQuant®; D.T. contributed specifically to writing and critical revisions to discussion relating to qFibrosis, qInflammation, qBallooning, and qSteatosis. All authors reviewed all subsequent drafts, and approved the final version of the manuscript for submission.

Conflict of interest

Mathieu Petitjean is an employee of PharmaNest, Inc., Princeton, NJ, USA; Cindy Serdjebi is an employee of Biocellvia, Marseille, France; Janani S. Iyer is an employee of PathAI, Inc., Boston, MA, USA; Dean Tai is an employee of HistoIndex Pte Ltd, Singapore. VR: received consulting fees from Novo-Nordisk, Sagimet, Madrigal, Enyo, Poxel, Northsea, Intercept Pharmaceuticals, Prosciento and research grants (to institution) from Gilead Sciences and Intercept Pharmaceuticals. SLF: Active Consulting: Alnylam, Axcella Health, Cargene, Cellarity, ChemomAb, Fate Therapeutics, Galmed, Gordian Biotechnology, Glycotest, Hepgene, In sitro, Korro Bio, Laekna, Laronde, Ochre Bio, Merck, Morphic Therapeutics, Overtone Therapeutics, North Sea Therapeutics, Pfizer Pharmaceuticals, Pliant, Prosciento Research, Resolution Therapeutics, Sagimet, Satellite Bio, Surrozen, Takeda Pharmaceuticals, Yaqrit; Stock options: Escient, Galectin Galmed, Genfit, Glympse, Gordian Biotechnology, Hepgene, Laekna, Lifemax, Metacrine, Morphic Therapeutics, Nimbus, North Sea, Therapeutics, Sagimet, Satellite Bio, Scholar Rock, Surrozen; Research Collaborations/ Research with Commercial Entities: Morphic Therapeutics; Novo Nordisk; Abalone Bio (SBIR Grant); Espervita, Galmed, Pionyr. Please refer to the accompanying ICMJE disclosure forms for further details.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2023.10.015.

[12]Hospital Universitario Virgen del Rocío, CiberEHD, Insituto de Biomedicina de Sevilla (HUVR/CSIC/US), Universidad de Sevilla, Seville, Spain

[13]Institute of Translational Immunology and Department of Medicine, University Medical Center, Mainz, Germany

[14]Department of Hepatology and Gastroenterology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

[15]NAFLD Research Center, University of California at San Diego, San Diego, CA, USA

[16]Université Paris Cité, Service d'Anatomie Pathologique, Hôpital Beaujon, Paris, France

[17]Pacific Rim Pathology, San Diego, California, USA

[18]Division of Gastroenterology, Hepatology and Nutrition, Virginia Commonwealth University, Richmond, VA, USA

Summary

The worldwide prevalence of non-alcoholic steatohepatitis (NASH) is increasing, causing a significant medical burden, but no approved therapeutics are currently available. NASH drug development requires histological analysis of liver biopsies by expert pathologists for trial enrolment and efficacy assessment, which can be hindered by multiple issues including sample heterogeneity, inter-reader and intra-reader variability, and ordinal scoring systems. Consequently, there is a high unmet need for accurate, reproducible, quantitative, and automated methods to assist pathologists with histological analysis to improve the precision around treatment and efficacy assessment. Digital pathology (DP) workflows in combination with artificial intelligence (AI) have been established in other areas of medicine and are being actively investigated in NASH to assist pathologists in the evaluation and scoring of NASH histology. DP/AI models can be used to automatically detect, localise, quantify, and score histological parameters and have the potential to reduce the impact of scoring variability in NASH clinical trials. This narrative review provides an overview of DP/AI tools in development for NASH, highlights key regulatory considerations, and discusses how these advances may impact the future of NASH clinical management and drug development. This should be a high priority in the NASH field, particularly to improve the development of safe and effective therapeutics.

Keywords

NASH; NAFLD; histology; machine learning; digital pathology; artificial intelligence; liver biopsy; fibrosis; steatosis; inflammation; ballooning; clinical trials

Introduction

Despite being associated with a considerable health and economic burden, there are no approved therapies for nonalcoholic steatohepatitis (NASH), the progressive form of nonalcoholic fatty liver disease (NAFLD), which is now a leading cause of chronic liver disease worldwide.[1–3] The NASH drug development process currently requires the demonstration of improvement in histological endpoints as a prerequisite for conditional

approval, followed by the demonstration of clinical benefit in long-term extension trials.[4–6] The assumption underlying this sequence of events is that NASH resolution and fibrosis regression, the two surrogate endpoints defining histological improvement, are likely predictors of improved outcomes.[7] Therefore, histological examination of liver biopsies is mandatory for the drug registration process in NASH.

While currently considered the reference standard,[4,5] conventional histological analysis has major shortcomings. In particular, imprecise definitions and subjective interpretation of lesions have caused high rates of intra- and inter-observer variability, while semi-quantitative grading and staging systems have limited sensitivity to change.[8,9] Consequently, most NASH therapeutic trials are plagued by a high screening failure rate at enrolment, high placebo response rates, and a reduced ability to uncover early treatment effects at follow-up biopsies.[10] Therefore, automated, objective, quantitative, precise, and reproducible methods to analyse liver histology with higher sensitivity are needed.

Digital pathology (DP) in combination with artificial intelligence (AI) is an area of active investigation with the potential to revolutionise histological analysis of liver biopsies in both clinical practice and therapeutic trials.[11] DP- and AI-based workflows are already established methods in other areas of medicine, supporting the potential utilisation of these tools in NASH.[12,13] Some DP/AI platforms bring translational capabilities that can be used for drug candidate screening and validation. They may also assist in uncovering biological mechanisms underpinning NASH development, including the co-existence of both progressive and regressive features,[14] as well as in the development of new NASH preclinical models and non-invasive tests (NITs) for diagnosis and monitoring of patients.[15–17] This review article will provide an overview of DP and AI, introduce DP/AI tools in development for NASH, summarise regulatory considerations for these tools, and discuss how these advances may impact the future of NASH drug development and clinical management. We will not discuss the large body of data generated by morphometry analysis of collagen content as the methodology used, while computer assisted and quantitative, does not involve AI and machine learning (ML) methodologies and therefore cannot provide information beyond what is already visible to the human eye.

Methods

This article was initially based on the proceedings of the 5th NASH Roundtable Virtual Forum on Thursday, June 16, 2022, titled “Is digital pathology a game-changer for histology in drug development?”, with input from four companies developing DP/AI tools for NASH, and was updated thereafter. This information was supplemented by literature searches performed using PubMed and the search terms “digital pathology AND artificial intelligence AND liver” and “non-alcoholic steatohepatitis AND digital pathology”. Published abstracts from international liver congresses (International Liver Congress and The Liver Meeting) in 2020–2023 were also searched using “NASH” or “digital”, or “artificial”. Abstracts were manually reviewed to identify those of relevance.

Limitations and challenges of conventional histopathological assessment for NASH

To diagnose NASH and assess disease progression or regression, pathologists must perform histological analysis of liver biopsies. Standard practice includes staining of liver sections with H&E followed by conventional (*i.e.* light microscopy) pathologist review to assess lobular inflammation, steatosis, and hepatocyte ballooning and with Masson's trichrome or picrosirius red to assess fibrosis and collagen proportional area (for the latter).[18] The NASH Clinical Research Network (NASH-CRN) classification is the most widely used histological scoring system for NAFLD[19] but other classifications have been proposed and are being used.[20,21] The NASH-CRN staging system includes disease activity assessment using the NAFLD activity score, which combines steatosis, lobular inflammation, and ballooning.[19] It also stages fibrosis on a 5-point scale, from stage 0 (no fibrosis) to stage 4 (cirrhosis).

Historically, conventional pathological interpretation has been key in diagnosing NAFLD, and in particular its progressive form (steatohepatitis), and establishing the fibrosis stage, a major prognostic determinant.[22] Advanced late-phase therapeutic trials have used conventional pathology to assess efficacy endpoints and define treatment success. In clinical practice, conventional pathology can identify complex diagnoses when other lesions or liver diseases coexist with NAFLD and help solve discordant results obtained with fibrosis biomarkers. Therefore, conventional pathology is and will continue to be essential for clinical research and will remain useful for the clinical management of patients with NAFLD.

Yet, there are multiple challenges associated with histological analysis of liver biopsies. First, there are issues inherent to the collection of a limited core fragment of the liver which is not necessarily representative of the entire liver. Landmark studies have shown variability of fibrosis distribution and fibrosis amount within the same core fragment or different sections of the same fragment[23] or within adjacent tissue fragments both for fibrosis, inflammatory foci and hepatocyte ballooning.[24] It is unclear if the variability in distribution is random or if it follows anatomical patterns related to, for instance, differential vascularisation of segments or lobes of the liver,[25,26] and, importantly, if these differences are magnified during the resolution of injury.[27] While the confidence in the pathological interpretation increases with fragment size, not all biopsies collected in clinical trials are of optimal length and width. A second issue relates to reader variability, either intraobserver variability or inter-observer variability. This is partly due to subjectivity in the interpretation of insufficiently precise histological definitions[28] (in particular for the quantitative aspects) and partly to lack of reproducibility at the observer level.[8] Importantly, observer disagreement is magnified in the middle of the spectrum of severity of injury rather than at the most severe or the almost normal ends. Regardless, lack of reproducibility may cause patients who meet study eligibility criteria to fail screening, lead to misclassification of disease activity or fibrosis stage and result in incorrect interpretation of drug treatment effects. A particular concern has been raised around hepatocyte ballooning, as its presence is a prerequisite for NASH clinical trial enrolment and its absence is established as a surrogate endpoint for NASH resolution.[28] Finally, the categorical nature of fibrosis staging systems,

mostly based on architectural disruption has two caveats: the amount of fibrotic matrix does not increase linearly with increasing fibrosis stages[23] and genuine changes in the amount of fibrosis may be missed because histological stages are semiquantitative. The first caveat implies that a one-stage reduction (or increase) in fibrosis stage may not represent the same amount of collagen removed (or synthesised) between two patients at different baseline stages. For instance, fibrosis stage 3, as a single stage, covers a large and variable amount of fibrotic tissue. The second caveat implies that early signals of antifibrotic potency of a candidate drug may be largely missed.

There is an unmet need for tools that can assist pathologists in reliable and reproducible identification, scoring, and quantification of NASH histological features in liver biopsies. These methods should also help pathologists to identify complex features of progression or regression of liver injury and allow for quantification of multiple histological lesions in an enhanced dynamic range for a more sensitive assessment of drug effects. They should be able to identify intra stage or intra grade improvement which is currently missed by the semiquantitative classification systems. This is important since early studies are typically of short duration and conducted in small study populations which limits the ability to detect significant changes unless measurements are more precise and truly quantitative. DP in combination with AI is an evolving discipline that may address some of these unmet needs. Moreover, by identifying patterns of progression or regression, aspects of fibrotic matrix disposition and by computing combined scores between different histological lesions, DP/AI could help pathologists refine the histological diagnosis and disease staging of NASH.

Digital pathology and artificial intelligence

DP refers to the tools and systems used to digitise pathology slides and associated meta-data and their review, storage, analysis, and related infrastructure (Table S1).[29] An example of DP is whole-slide imaging (WSI), which involves the production of virtual images of stained or unstained pathology slides, created through robotic scanning and digital reconstruction at very high resolution. WSI allows for the online sharing and digital storage of pathology slide images, and/or integration into telemedicine.[30,31] Another example is the stain-free imaging approach using second harmonic generation/two-photon excitation (SHG/TPE) fluorescence, where staining-related variations are removed, allowing for analysis of fine structural detail.[32,33] Digital image analysis algorithms can segment tissue regions, cells and relevant structures (*e.g.*, blood vessels, portal tracts), quantify positive staining (or other sources of contrast), define regions of interest, and/or allow users to capture or annotate images of selected regions for future examination.[31]

AI is a branch of computer science dealing with the simulation of intelligent behaviour in computers. AI can be used in combination with digital image analysis techniques to accurately classify or segment images and can be trained in supervised (leveraging annotations provided by expert pathologists) or unsupervised (without pathologist input) manners.[29] Expert systemsare an older branch of AI based on a deterministic approach,[34] unlike ML where a computer learns how to perform a task after being exposed to representative data using a statistical approach. Deep learning is where a computer trains itself to perform tasks by exposing artificial neural networks to large amounts of data and

does not require annotations.[29] By incorporating AI into DP, image analysis can be made quantitative and potentially more efficient and accurate.[35–37] Potential benefits of DP/AI approaches include improvements in fidelity and reproducibility, but results need to be interpreted with caution due to limited clinical experience.[37,38]

There are several potential challenges to introducing DP/AI tools into clinical practice, including optimal sample collection and preparation, digitisation, pathologist involvement and training, and implementation of suitable IT infrastructure.[37,39] Standardised sample collection and preparation is essential for both conventional reading and DP. Many DP/AI approaches rely on WSI of stained samples where the inherent issue of staining variability remains. AI can also be used to interpret SHG images of unstained samples to further optimise histological assessment of tissue samples.[40]

DP/AI models can be used to automatically detect, localise, quantify, and score various histological parameters.[29,41] Moreover, they offer the possibility to study quantitatively the relative proportion of different morphological features including those that are not part of the histological classifications (such as portal inflammation or bile duct changes). In the future, these data may independently, and in combination with other medical data (*e.g.*, genetic, radiological, liquid biomarkers, novel immunohistological markers), enable prediction of patient outcomes.[42–44] DP/AI approaches may also allow pathologists and clinicians to obtain insights into histological features or changes that may be associated with disease progression and regression, which are not included or readily distinguishable using conventional scoring systems.[43,45] DP/AI-based methods have been developed and approved for use in oncology diagnostics.[12] For example, image analysis software has been approved in the US and EU for immunohistochemical assays in breast cancer and the US Food and Drug Administration (FDA) has approved an AI-based pathology product for prostate cancer diagnosis.[46] DP/AI is therefore becoming a mainstream technology with significant prospects for growth with technological improvements.

DP/AI tools for the assessment of NASH histological features

Several academic research groups have developed DP/AI tools to evaluate the histological features of liver biopsies from patients with NASH or preclinical models of the disease. Some DP/AI tools analysed specific features of NASH liver histology, such as fibrosis[47–49] steatosis[50–53] inflammation[54,55] and ballooning.[56,57] Several studies demonstrated good correlation between results obtained by novel DP/AI tools and experienced pathologists.[47,51,58,59] However, these approaches have yet to become widely adopted.

For widespread implementation of DP/AI approaches in NASH, commercial development and approval by regulatory authorities is necessary. To that end, several companies are developing DP/AI tools for NASH diagnosis and monitoring,[15,45,60–64] but this review focuses only on the four that are most developed (Table 1). This section provides an overview of the different methodologies, including technical insights. The following sections will discuss how these DP/AI tools are being developed and validated in NASH. A potential bias in the published literature summarised in these sections may be the over-representation of successful results.

AIM-NASH and NASH explore (PathAI Inc.)

AI-based histologic measurement of NASH (AIM-NASH) is an AI-based DP tool designed to assist pathologists in achieving accurate, reproducible grading and staging of NASH histology in clinical trials (Table 1 and Fig. 1)[45]. AIM-NASH algorithms were trained using WSI from liver biopsies across several NASH clinical trials and annotations provided by expert pathologists to predict NASH-CRN steatosis, ballooning, and inflammation grades, as well as fibrosis stage. The AIM-NASH AI-assist clinical trial workflow is conducted via a Good Clinical Practice-compliant WSI viewer and displays both the model's predictions for NASH-CRN scores and the original WSI with corresponding colourised heatmap overlays that spotlight histological features relevant to scoring. This approach allows the pathologist to accept or reject the model-provided scores for each histological feature. AIM-NASH is currently undergoing qualification for clinical trial enrolment and primary endpoint assessment with the FDA and European Medicines Agency.

NASH explore comprises a suite of AI-based algorithms developed to enable quantitative, continuous evaluation of NASH biopsy tissue to facilitate precise measurement of therapeutic response. NASH explore uses both H&E- and Masson's trichrome-stained images to enable continuous quantification of tissue features (*e.g.*, proportionate area of lobular and portal inflammation and subtypes of fibrosis) and cell features (*e.g.*, counts and densities of different types of hepatocytes and immune cells), in addition to spatial relationships amongst these features. Both AIM-NASH and NASH explore tools have integrated AI-based artifact models that detect and exclude image and tissue artifacts (*e.g.*, out-of-focus areas and tissue folds) prior to downstream feature quantitation and scoring.

Analytical validation of the algorithms in terms of repeatability and reproducibility between scanners at different locations has shown high agreement rates for fibrosis, hepatocyte ballooning, lobular inflammation and steatosis ranging between 0.93 and 0.96[65], which is better than the repeatability of conventional pathologist reads.[8] Clinical validation studies tested agreement between a consensus read of three expert hepatopathologists, designated as "ground truth", and, on the one hand, the AIM NASH algorithms and on the other hand, the conventional reads of three individual hepatopathologists. In a study including 1,500 biopsies from three randomised-controlled trials, AI-assisted reads were superior to conventional reads for hepatocyte ballooning, lobular inflammation, NASH resolution and at-risk NASH (NAS ≥4 with stage 2 or 3 fibrosis) and non-inferior for fibrosis and steatosis.[65]

FibroNest™ (PharmaNest Inc.)

FibroNest™ is a drug development tool designed to assist pathologists in the assessment of fibrotic tissues (Table 1 and Fig. 2). FibroNest™ is compatible with conventional pathology workflows, stains, and WSI scanners, and is enabled by the BISQUE (Bio-Image Semantic Query User Environment) database[66] to provide next-generation digital image management, visualisation, annotation, and sharing. The approach is based on several steps, including: (1) digital image colour normalisation, standardisation, and segmentation pre-processing to reduce colour and intensity variations present in stained images, including additional preprocessing steps to eliminate scanning and other artifacts;[67–69] (2) single-fibre,

single-nucleus, single-object, high-content, quantitative image analysis of DP images to generate large quantitative and relevant data lakes at the levels of collagen content, fibre morphometry, and fibrosis architecture;[69,70] and (3) supervised AI for the automated and continuous quantification of fibrosis phenotypes and associated histological features.[71] This approach is independent of fibrosis aetiology and the type of tissues studied and can quantify and classify fibrosis phenotypes independently of staining or imaging methodology[72] and liver disease aetiology.[73] The high signal-to-noise ratio generates outputs that avoid staining variability.[74] When FibroNest™ was compared with SHG imaging of unstained tissue sections, no significant differences in fibrosis severity scoring of liver biopsies from patients with NASH or rodent NASH models were identified.[72,75] Results from a retrospective study of patients with NASH diagnosed by conventional pathology showed that FibroNest™ scores correlated with NASH-CRN fibrosis scores at baseline.[74] FibroNest™ can classify type 1 and 2 fibrosis phenotypes in paediatric patients with NASH[76] and has demonstrated that fibrosis phenotypes are identical in lean and obese patients with moderate-to-severe NASH.[77] At the extremes of fibrosis severity ranges, FibroNest™ can classify substages of F1 and F4[78]. In addition to fibrosis assessment, FibroNest™ can quantify other histological features including steatosis, inflammation, and hepatocyte ballooning.[79]

MorphoQuant® (Biocellvia)

MorphoQuant® is a fully automated AI expert system designed to identify all histological features required for NASH activity scoring and fibrosis grading, and to perform quantitative measurement (Table 1 and Fig. 3)[60]. It uses the deterministic approach, made of rules and statements to recognise, describe, and characterise histological features, based on the principles of morphometric analysis. MorphoQuant® works from histological slides stained with classic stains alone or in combination with specific labels to emphasise information contained in images. This tool does not require expert pathologists' annotations to train a machine. Use of MorphoQuant® is compatible with the general workflow of clinical studies, where whole slide images are submitted to MorphoQuant®. After selecting organ, pathology, species, and image magnification, no further human intervention is required to generate quantitative data and pathology images, ensuring reproducibility and objectivity in the process. The software also automatically discards very small, fragmented sections and considers the size of fragments to provide a global quantity. Generated data are submitted to one or several pathologists for review and used for scoring. Prospects for future development include investigation of new morphological definitions of histological features, such as inflammation or hepatocyte ballooning.

qFibrosis, qInflammation, qBallooning, and qSteatosis (HistoIndex Pte Inc.)

HistoIndex developed a tissue imaging system based on nonlinear optical microscopy that enables observation of endogenous tissue signals using SHG/TPE in unstained tissue samples (Table 1 and Fig. 4). This avoids both the need for staining procedures and degradation artifacts. TPE fluorescence microscopy provides visualisation of the background liver architecture (inflammation, hepatocyte ballooning, steatosis) while the SHG signal provides accurate identification and quantification of fibrillar collagen.[80,81] A large number of individual collagen fibre parameters termed quantitative fibrosis parameters (qFPs)[48] have been identified and quantified, including number, length, diameter, orientation,

contour, and cross-linkage. Interestingly, four of these biophysical parameters with the strongest association with fibrosis stages (number of collagen strands, strand length, strand eccentricity, and strand solidity) allowed for discrimination of adjacent fibrosis stages (except for stage 1 *vs.* stage 2, which both include a peri-sinusoidal fibrotic contingent). These parameters were used to model a continuous linear fibrosis score (qFibrosis) that is strongly related to fibrosis stage in NAFLD.[48,49,82] The algorithm was developed with pathologists to ensure that output interpretation is relevant for clinical and pathological assessment.[40,61,83] Conversely, when assisted by qFibrosis, the concordance rate between pathologists for fibrosis staging improved substantially,[38] which indicates its potential value as a diagnosis-assistive tool in clinical practice or therapeutic trials. The AI-based algorithm recognises and segregates the biopsy area into histological regions: central vein, peri-central, portal tract, peri-portal, peri-sinusoidal, and transitional.[84] These are then classified into zone 1 (portal tract and peri-portal), zone 2 (peri-sinusoidal), and zone 3 (per-central and central vein). The definition of these regions is based on the NASH-CRN staging system such that specific fibrosis morphological features (*e.g.*, area, number, and dimensions of collagen strings) can be quantified in each of these regions independently. In addition, the AI algorithms allow for the accurate detection of septa and nodules in patients with NASH cirrhosis, which can be used for disease monitoring and treatment response evaluation.[85]

QFibrosis has been combined with quantitative analysis of steatosis (qSteatosis), inflammation (qInflammation), and hepatocyte ballooning (qBallooning) in a Qfibs tool.[55] Correlation with hepatocyte ballooning is only moderate,[54] which is largely explained by inter-observer variability,[28] limiting the fidelity of the reference standard. Also, the score has only moderate discrimination for severe inflammation and cannot distinguish between different inflammatory cell types.[55] While awaiting further validation, these limitations require that Qfibs be used once pathological diagnosis of NASH has been established as an adjunct to pathologist grading of disease activity and staging. Despite this, the greatest potential of Qfibs may be for the robust quantification of changes in disease severity before and after therapeutic interventions, for example in clinical trials. Other SHG-based algorithms have been developed, allowing for automated quantification of fibrosis and prediction of fibrosis stage in chronic viral hepatitis.[86,87]

DP/AI tools and longitudinal changes in liver histology

Several DP/AI tools discussed above have been evaluated in clinical studies to assess longitudinal changes in liver histology and to assist pathologists in achieving reproducible scoring in clinical trials. In a study of Chinese patients with NAFLD and serial biopsies, the sensitivity of selected qFPs for fibrosis reduction was low, while they showed higher sensitivity and acceptable specificity for fibrosis progression.[81] Importantly, the diagnostic accuracy of selected qFPs at baseline biopsy was insufficient to predict fibrosis changes on follow-up biopsy.[55] In a subgroup of patients from a 1-year phase IIb trial of tropifexor *vs.* placebo, SHG/TPE analysis revealed a treatment-associated reduction of liver fibrosis otherwise not discernible by conventional pathology.[84] Importantly, patterns of regression in fibrosis septa morphology and reduction in septa parameters (septa area, septa width, fibre interactions and aggregated septa) were documented for different fibrosis locations specifically (*i.e.* portal, periportal, pericentral, perisinusoidal and septal), thus providing

tools for future studies of fibrosis regression in relation to the mechanism of action of different drugs.[88]

Moreover, these changes can be quantitatively assessed (through radar maps), a key requirement in the context of therapeutic trials.[88] Conventional pathology assessed fibrotic septa as “unchanged”, highlighting again the potential of DP/AI to fully document anti-fibrotic effects in the context of therapeutic trials.[84,88] Similarly, preliminary data from the phase IIb resmetirom trial has documented a significant reduction in fibrosis assessed by SHG/TPE compared with placebo, while no difference was seen between groups when fibrosis regression was assessed using conventional pathology.[89,90] These data were confirmed in the phase III MAESTRO-NASH trial of resmetirom *vs.* placebo: the qSteatosis score reproduced the reduction in steatosis grade observed by conventional pathology and correlated with changes in MRI-estimated proton density fat fraction. Changes in categorical qFibrosis stage were aligned with NASH-CRN pathologist scoring and identified 90% of responders as per conventional pathology. Changes from baseline in the continuous qFibrosis score confirmed the antifibrotic response of resmetirom, while demonstrating less worsening of fibrosis compared to placebo. These data further reinforce the validity of digital image quantification in therapeutic trials.

AIM-NASH has recapitulated endpoints met by central pathologists in several studies, including the phase II EMINENCE trial and phase II study of resmetirom.[91,92] In the phase IIb ATLAS trial, analysis of liver biopsies using the NASH explore tool showed that cilofexor and firsocostat combination treatment caused a significant decrease in NASH-CRN fibrosis score compared with placebo.[93] While a phase II trial of semaglutide did not demonstrate a significant improvement in fibrosis stage compared with placebo using conventional histological analysis,[94] NASH explore continuous fibrosis scoring showed a significant improvement in fibrosis score in patients treated with semaglutide *vs.* placebo.[95] Similar documentation of fibrosis improvement was possible at the cirrhotic stage with the same agent.[96] AIM-NASH has also detected successful primary endpoint achievement (fibrosis regression) where the central pathologist did not in a phase II study of pegbelfermin.[97]

In a phase IIb trial of aramchol, the use of a fibrosis composite score by the FibroNest™ tool enabled identification of patients with fibrosis improvement with high sensitivity.[98] These results were corroborated by conventional staging, which confirmed fibrosis stage reduction and improvement by ranked assessment.[98] Similar results were observed in the phase II LPCN 1144 LiFT trial.[78] These results were confirmed prospectively in the phase II pegbelfermin trial programme, which demonstrated correlation with changes in Ishak fibrosis stages before and after 24 weeks of treatment.[99] Distinct single fibre quantitative traits combined in a single score demonstrated a dose-response relationship between placebo and three increasing doses of active drug,[100] while no antifibrotic response was visible by conventional pathology. These data illustrate how quantitative image analysis can help uncover treatment effects, but more studies are necessary for the optimal selection of the fibrosis parameters to be included in the best performing scores. Since these scores are based on different underlying reading methodologies, they will not be interchangeable. This will raise the prospect of direct face to face comparisons within a trial dataset, although the

outcome of these comparisons will need to be agreed upon: sensitivity for early changes? specificity for robust major changes? accuracy for prediction of liver-related events?

DP/AI tools and correlation with clinical outcomes

While it is unquestionable that DP/AI provides enhanced granularity and higher sensitivity when assessing temporal changes in liver histology, the clinical relevance of DP/AI changes remains to be established. Therefore, the next big challenge is to assess to what extent changes in DP/AI parameters predict the occurrence of clinical outcomes. In the aforementioned study of serial liver biopsies in Chinese patients with NAFLD, after a mean follow-up period of 5.6 years, patients with high baseline values of selected qFPs had an increased incidence of liver-related events.[82] It is unclear if changes in qFPs were also predictive of clinical outcomes.

In a series of 294 Scottish patients with NAFLD, with long-term follow-up, a baseline index based on 25 fibrosis parameters derived from SHG-TPEF imaging was able to predict all-cause mortality, cirrhosis decompensation and occurrence of hepatocellular carcinoma with greater accuracy than the NASH-CRN stage. Interestingly these predictive indices were not identical to and actually performed better than the qFibrosis score itself.[101] The association between fibrosis parameters and clinical outcomes was also demonstrated using single fibre quantitative and high resolution technology (PharmaNest) in a multicentric series of 304 patients with NAFLD and a median follow-up of 11.4 years.[102] In this study, a liver event predictive score was developed and achieved an AUC of 0.78 for prediction of the occurrence of a composite endpoint of liver-related events and hepatocellular carcinoma. In contrast to the previous series based on SHG-TPEF readings[101] the predictive value of the liver event predictive score was not different from that of the baseline quantitative fibrosis score.[102] These studies are continuing in order to increase the accuracy of prediction and statistical power.

In the phase III STELLAR trials of selonsertib, AIM-NASH continuous fibrosis scoring approaches predicted progression to cirrhosis in patients with stage 3 fibrosis and the occurrence of clinical events (cirrhosis decompensation) in patients with cirrhosis.[42] Similarly, an ML-based approach that used deep convolutional neural networks to produce pixel-level predictions of fibrosis and nodularity was shown to be well correlated with haemodynamically measured hepatic venous pressure gradient (HPVG), only modestly correlated with reductions in HPVG, and strongly correlated with fibrosis changes.[103] Importantly, increases from baseline in ML-HPVG were associated with an increased risk of clinical events. A different ML-based algorithm was used in patients with NASH and compensated cirrhosis to identify and quantify fibrosis changes within the F4 stage as well as portal hypertension.[85] Across all these studies, ML-based histological scores accurately predicted HPVG changes, clinically significant portal hypertension, and development of varices in patients with cirrhotic NASH.[42,85,103] Finally, in patients who underwent liver transplantation, unsupervised AI-identified quantitative fibrosis traits in non-tumoral tissue associated with the presence of hepatocellular carcinoma.[104] This exciting area of investigation may reveal the full potential of the different DP/AI methodologies in

determining disease severity, predicting disease course, and identifying therapeutic benefits of pharmacotherapies in development.

DP/AI tools in preclinical models

DP/AI tools are also being investigated in preclinical models to improve understanding of the biological mechanisms underlying NASH and to screen potential drug candidates. AIM-NASH, as part of a multidimensional analysis of hepatic and splenic immune cells, revealed distinct immunological patterns that inhibit hepatic carcinogenesis and that could provide valuable insights on progression of liver disease.[105] Results from the FAT-NASH mouse model showed that quantitative fibrosis and steatosis scoring with FibroNest™ was superior to conventional NASH staging when describing progression and treatment responses and their effect on specific phenotypic fibrosis traits.[106] FibroNest™ can detect and quantify the effects of multiple drug treatments, as well as the effect of gene therapies and knockout approaches.[107] FibroNest™ has also been used to assess changes in liver histology in response to drug treatments in studies of human 3D NASH models.[15] The MorphoQuant® tool can quantify fibrosis, steatosis, and inflammation in preclinical rodent models of NASH,[60,108,109] with good correlation between MorphoQuant® assessment of fibrosis and conventional pathologist scoring when used to assess liver samples from two mouse models of liver disease.[17] In particular, inflammation was assessed through the identification of hepatic crown-like structures, a NASH-specific feature, involved in inflammation and fibrosis development.[60,110] The qFibrosis tool has been used to systematically evaluate liver fibrosis progression patterns in different animal models.[14,111] This approach can be used to quantify specific fibrosis progression and regression patterns in different liver regions, which can assist in improving understanding of therapeutic agents' mechanisms of action. Furthermore, qFibrosis has been used to estimate changes in fibrosis patterns in preclinical NASH treatment models with either pharmacological or metabolic interventions.[112,113]

Other potential uses for DP/AI tools in NASH

A major advantage to DP/AI tools is the quantitative and continuous scoring of histological features *vs.* conventional categorical scoring systems. NIT development for NASH diagnosis and management is an active area of investigation. The potential added value of integrating DP/AI tools with NITs is providing a bridge between quantitative and continuous histology scores and continuous data from NITs. A *post hoc* analysis of the phase IIb FALCON1 study evaluated the association between continuous histological scores generated by FibroNest™ and imaging-based scores using magnetic resonance enterography and MRI-estimated proton density fat fraction measurements.[114] The results showed agreement between quantitative DP scores and imaging-based measurements. Additional studies have shown correlations between changes in imaging-based biomarkers and NASH explore-derived quantitative histological features.[115] These preliminary data suggest that DP/AI tools can be used to benchmark histological changes with imaging-based measurements, and likely novel serum markers, thereby assisting NIT development and validation, while also improving the ability to demonstrate treatment-induced histological benefit.

Pathomic fusion is a recently proposed approach to integrate histopathology data with molecular features for improved diagnosis and prognosis.[116] It combines large quantitative DP datasets with molecular datasets using AI. ML-based analyses of liver histology have been integrated with transcriptomic data from the same tissue to obtain mechanistic insights into NASH disease progression and to identify genes associated with portal inflammation and bile duct area to predict patient prognosis.[117] Moreover, an ML-based histological analysis from patients with NASH and stage 3 fibrosis or cirrhosis, when combined with phenotypes derived from transcriptomic and genetic data, enabled improved prediction of molecular endpoints relative to standard histology. It also identified a histological phenotype associated with the *PNPLA3* I148M genotype in the NASH cohort.[62] These results highlight the potential of combining large DP datasets with genomic data to promote our understanding of NASH pathogenesis, but this research is still at an early stage.

A proposal for clinical validation

A proposal for the clinical validation of DP/AI is outlined in Fig. 5. At a minimum DP/AI should be able to reproduce fibrosis staging, as currently assessed by common histological classifications, such as the NASH-CRN classification, using conventional pathology reading as a reference standard. However, the ability of DP/AI to go beyond the mere reproduction of the traditional histological classifications creates the opportunity for a quantitative and sensitive assessment of disease progression (and possibly regression) which could have significant implications for therapeutic trials. Continuous quantitative scores should provide information on disease course, both during the natural course and as a result of specific interventions. A crucial requirement is that both fibrosis progression and fibrosis regression be reflected in the directionality of score changes. If measurable, structural collagen fibre properties should provide information about patterns of fibrosis progression or regression. This would help determine, for a given fibrosis stage, the trajectory of the fibrotic process and whether this trajectory has been impacted by a particular intervention.

Ultimately, the relevance of the continuous scores or the fibrosis fibre structures will be demonstrated by their ability to predict clinical, liver-related events. This should be tested for both baseline biopsy values and for changes between serial biopsies resulting from an intervention. A convincing demonstration of the clinical value of DP/AI findings will position this technology as a surrogate for drug efficacy in clinical trials, with the main advantages being the reproducibility, lack of variability and robustness of an automated technique, while the continuing reliance on liver biopsy is a disadvantage.

Regulatory considerations for DP/AI tools

Regulatory processes vary across countries and evolve over time. Therefore, this section discusses some general points that should be considered when developing, validating, and selecting DP/AI tools. The Digital Pathology Association (DPA) has released a white paper that provides recommendations and guidance on best practice, proposing that companies and regulatory authorities should collaborate to set formal standards for DP/AI components and software to ensure that tools are safe, effective, and beneficial for patients.[29] The DPA released another white paper detailing the FDA regulatory framework of quantitative image

analysis (QIA) tests for biomarker use and related guidelines. With more than 33 QIA applications approved by the FDA, these regulatory pathways and guidelines can also be applied to DP/AI tools.[13]

DP/AI tools consist of both hardware (*e.g.*, microscope, slide scanner) and software (*e.g.*, AI algorithm). Each component can be approved individually or as a complete system.[118] Two WSI systems have been approved by the FDA for histopathological analysis and these include the slide scanner, image management system, and display.[39] The first AI product to be approved by the FDA is designed to identify regions of interest on digital prostate biopsy images for further analysis by pathologists. This software is compatible for use with available digital WSI systems.[46] DP/AI tools are regulated as *in vitro* diagnostics (IVDs) or medical devices. Device classifications are based on the intended use and the risk to human health.[119] In the US, devices (including software) that are developed for use in an individual pathology laboratory are defined as laboratory-developed tests, regulated under the Clinical Laboratory Improvement Amendments, and require validation and self-certification by the individual laboratory. Devices that are developed for commercial distribution are regulated by the FDA based on patient risk (classes I–III; Table S2). FDA-approved WSI systems have been classified as class II devices with special controls.[119,120] DP image management and viewing platforms can be included as Medical Device Data Systems, Medical Image Storage Devices, and Medical Image Communication Devices, and are covered by their own FDA guidelines.[121] In the EU, WSI and image analysis systems (including software specifically intended for medical uses) are considered IVDs and are regulated by IVD device regulations that came into effect in 2022. WSI systems are considered to be class C devices based on their intended use and risk (Table S2). DP/AI tools developed for use in NASH will likely be classified as class C devices.[118]

The FDA has approved several QIA algorithms. However, these algorithms do not interpret data but provide quantification scores that must be approved by a pathologist.[13] ML-based algorithms that are trained with pathologist involvement are considered more likely to gain FDA approval compared with unsupervised AI approaches.[53] However, a deep learning-based AI algorithm has been approved by the FDA for use in detection of diabetic retinopathy.[31] The FDA released an AI/ML Software as a Medical Device (SaMD) Action Plan that states that SaMD must not reproduce human bias and must demonstrate greater certainty and agreement compared w*ith conventional ap*proaches.[122] The SaMD includes AI-based algorithms intended to be used to diagnose, treat, cure, mitigate, or prevent disease and can be locked or open for continuous learning.[13,122] For algorithms that are open for continuous learning, the DPA recommends use of Good Machine Learning Practices to anticipate modifications and thereby define what the algorithm should become when it is learning and how it will remain safe and effective.[13]

Prospects for future utilisation and limitations of DP/AI in NASH

There is enthusiasm in the NASH community about the potential of DP/AI so it is important to consider where these tools will be most useful in NASH diagnosis and management (Table 2 and Fig. 6). In the short term, DP/AI tools are being developed to assist pathologists with grading and staging of liver biopsies in an accurate and reproducible manner,[38] while

reducing the impact of limitations associated with manual scoring. Histological assessment of liver biopsies is required for patient enrolment and assessment of drug activity in clinical trials, and this is one area where DP/AI tools are beginning to demonstrate their potential. Several of the DP/AI tools described in this review are supporting clinical trials, both prospectively and retrospectively. The use of continuous scoring outputs generated by DP/AI tools may allow for the more accurate assessment of disease progression and regression than can be determined using standard categorical scoring systems, including subtle changes that may be missed or impossible or impractical to categorise by light microscopy.[38,45,84,95] This could enable more accurate identification of potential biological activity of investigational therapeutic agents within the short timeframe of standard clinical trials and will hopefully allow for more successful drug development and the use of DP/AI tools as surrogate endpoints. Other short-term goals of DP/AI in NASH include relating digitally captured changes in cell characteristics to fibrosis stage and disease progression over time, correlating early changes in fibrosis patterns with risk of cirrhosis and clinical outcomes, and linking short-term changes in histological activity to disease progression and regression. For example, novel immunohistochemistry-based parameters could be quantified using DP/AI to reflect changes in fibrosis dynamics. These could include markers of fibrosis progression (*e.g.,* platelet-derived growth factor and transforming growth factor-β) and regression (*e.g.*, matrix metalloproteins). In the long term, DP/AI could improve knowledge of the biology underpinning NASH, thereby improving disease definitions, characterisation of patient subgroups (such as paediatric patients),[123] and ultimately prediction of clinical outcomes. The identification of AI-derived continuous or exploratory histological features may also assist with the development of novel clinical endpoints, thereby improving the drug development process, and identification of novel biomarkers of both NASH and fibrosis progression/regression. DP/AI tools may also have an impact on day-to-day clinical practice. DP images could be used to communicate with patients about a disease that may have minimal symptoms, thus allowing patients to visualise and better understand the changes that are happening to their liver. In turn, this may improve patient adherence to lifestyle modifications and treatments.

Current limitations of DP/AI for use in NASH trials

There are limitations associated with DP/AI tools under development in NASH. Identification of subtle changes in histological processes, while potentially beneficial in drug development, currently has unknown clinical relevance for predicting patient outcomes. Additionally, most ML-based algorithms rely on training by multiple experienced pathologists who may not agree on individual features.[45] However, this would not be an issue for tools using expert system-based AI algorithms. It is essential for developers of DP/AI tools to demonstrate that they are more reliable and reproducible than conventional histological approaches, requiring transparency when publishing results. However, even with improvements in reliability and reproducibility, DP/AI tools are still subject to sampling variability and thus provide results unrepresentative of the entirety of the liver. Also, pre-analytical steps in biopsy management, such as quality of staining or thickness of cut sections, need to be carefully standardised given the impact they have on the quality of the readings. Moreover, the formal demonstration of the lack of intrareader variability of

DP/AI methods is still lacking when addressing, for instance, stains performed on different days, slides scanned on different scanners, different section thicknesses or diameter of the core fragments and any other technical differences between commercial vendors. Also, with multiple companies developing different DP/AI tools, there must be alignment on the optimal utilisation of these tools, *i.e.* in which patients or scenarios the different tools are most effective and how they can be combined for analysis of clinical data sets. It is likely that the NASH community will be able to select from a portfolio of tools based on their needs and available clinical data. As with other new technologies, the end users will ultimately decide which tools are the most useful, but approval by regulatory authorities may influence these decisions.

It is valid to argue that conventional histology readings may be sufficient to detect changes in NASH pathological features for compounds with strong efficacy thus rendering DP/AI analyses irrelevant. However, phase IIb trials, which can have a significant impact on the fate of drug candidates, are usually of small sample size and of rather short duration. Even potent drugs may not demonstrate noticeable effects in this setting, while positive results from small trials may be chance findings that are not subsequently confirmed. Therefore, it is important to either corroborate anti-fibrotic effects detected by conventional histology or document early histological changes that indicate (the dynamics of) anti-fibrotic effects in a quantifiable manner. This could help establish dose-response relationships or understand kinetics or zonal determinants of histological changes, which would aid in the assessment of the potential of a drug candidate.

As NASH prevalence increases, DP/AI approaches may help to manage increasing liver biopsy assessments that place escalating demands on pathologists. It is important to remember that DP/AI tools will not replace pathologists and the impact on pathologists' workload will depend on how the tools are designed, for example, whether a tool requires a pathologist to use the software themselves or only to receive data and outputs generated by the software. Additionally, it is essential that DP/AI tools can be adapted for use in other settings, such as for diagnosis and staging of paediatric NASH.[123,124] Ultimately, histological analysis of NASH may be superseded by NITs, but in the meantime, we strongly believe that developing and validating DP/AI tools should be of the highest priority in the NASH field. This is particularly important for drug development as the limitations of current histological assessments may have played a major role in the failure of some trials to achieve their histological endpoints, even when drug candidates have demonstrated clear biological efficacy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Partial support for writing and editing provided by Karen Beckett (Science Collected), James Olivera (Science Collected), Kutbuddin Akbary (HistoIndex), and Michelle Mazuranic (ProSciento). Funded by ProSciento.

Financial support

The authors received no financial support to produce this manuscript.

Abbreviations

AI	artificial intelligence
AIM-NASH	AI-based histologic measurement non-alcoholic steatohepatitis
CRN	Clinical Research Network
DP	digital pathology
DPA	Digital Pathology Association
FDA	US Food and Drug Administration
HPVG	hepatic venous pressure gradient
IVD	in vitro diagnostic
ML	machine learning
NAFLD	non-alcoholic fatty liver disease
NASH	non-alcoholic steatohepatitis
NIT	non-invasive test
qFPs	quantitative fibrosis parameters
QIA	quantitative image analysis
SaMD	Software as a Medical Device
SHG	second harmonic generation
TPE	two-photon excitation
WSI	whole-slide imaging

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;75:284–291. [PubMed: 33746083]

[2]. Le MH, Le DM, Baez TC, et al. Global incidence of non-alcoholic fatty liver disease: a systematic review and meta-analysis of 63 studies and 1,201,807 persons. J Hepatol 2023;79:287–295. [PubMed: 37040843]

[3]. Golabi P, Paik JM, AlQahtani S, et al. Burden of non-alcoholic fatty liver disease in asia, the Middle East and North africa: data from global burden of disease 2009–2019. J Hepatol 2021;75:795–809. [PubMed: 34081959]

[4]. Food Usa and Administration Drug. Noncirrhotic nonalcoholic steatohepatitis with liver fibrosis: developing drugs for treatment guidance for industry. Draft Guidance; 2018.

[5]. European Medicines Agency. Reflection paper on regulatory requirements for the development of medicinal products for chronic non 6 infectious liver diseases. PBC, PSC, NASH); 2018.

[6]. Loomba R, Ratziu V, Harrison SA. Expert panel review to compare FDA and EMA guidance on drug development and endpoints in nonalcoholic steatohepatitis. Gastroenterology 2022;162:680–688. [PubMed: 34822801]

[7]. Ratziu V, Francque S, Sanyal A. Breakthroughs in therapies for NASH and remaining challenges. J Hepatol 2022;76:1263–1278. [PubMed: 35589249]

[8]. Davison BA, Harrison SA, Cotter G, et al. Suboptimal reliability of liver biopsy evaluation has implications for randomized clinical trials. J Hepatol 2020;73:1322–1332. [PubMed: 32610115]

[9]. Younossi ZM, Loomba R, Anstee QM, et al. Diagnostic modalities for nonalcoholic fatty liver disease, nonalcoholic steatohepatitis, and associated fibrosis. Hepatology 2018;68:349–360. [PubMed: 29222917]

[10]. Ratziu V, Friedman SL. Why do so many nonalcoholic steatohepatitis trials fail? Gastroenterology 2023;165:5–10. [PubMed: 32439497]

[11]. Nam D, Chapiro J, Paradis V, et al. Artificial intelligence in liver diseases: improving diagnostics, prognostics and response prediction. JHEP Rep 2022;4:100443. [PubMed: 35243281]

[12]. Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. J Intern Med 2020;288:62–81. [PubMed: 32128929]

[13]. Lara H, Li Z, Abels E, et al. Quantitative image analysis for tissue biomarker use: a white paper from the digital pathology association. Appl Immunohistochem Mol Morphol 2021;29:479–493. [PubMed: 33734106]

[14]. Ng N, Tai D, Ren Y, et al. Second-harmonic generated quantifiable fibrosis parameters provide signatures for disease progression and regression in nonalcoholic fatty liver disease. Clin Pathol 2023;16.

[15]. Petitjean L, Strobel S, Thoma C, et al. Quantitative digital pathology of 3D human NASH models establish continuous scores to evaluate the antifibrotic effects of selonsertib, firsocostat and resmetiron. J Hepatol 2022;76:S481.

[16]. Martic M, Tai D, Naoumov N. Magnetic resonance elastography (MRE) demonstrate the strongest correlation with digital pathology and NASH CRN fibrosis assessments, compared to transient elastography and other assessed non-invasive tests (NITs). J Hepatol 2022;77:S493.

[17]. Serdjebi C, Bertotti K, Huang P, et al. Automated whole slide image analysis for a translational quantification of liver fibrosis. Sci Rep 2022;12:17935. [PubMed: 36333365]

[18]. Goodman ZD, Becker RL Jr, Pockros PJ, et al. Progression of fibrosis in advanced chronic hepatitis C: evaluation by morphometric image analysis. Hepatology 2007;45:886–894. [PubMed: 17393526]

[19]. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–1321. [PubMed: 15915461]

[20]. Bedossa P. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology 2014;60:565–575. [PubMed: 24753132]

[21]. Pai RK, Jairath V, Hogan M, et al. Reliability of histologic assessment for NAFLD and development of an expanded NAFLD activity score. Hepatology 2022;76:1150–1163. [PubMed: 35332569]

[22]. 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 2020;158:1611–1625 e12. [PubMed: 32027911]

[23]. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449–1457. [PubMed: 14647056]

[24]. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005;128:1898–1906. [PubMed: 15940625]

[25]. Larson SP, Bowers SP, Palekar NA, et al. Histopathologic variability between the right and left lobes of the liver in morbidly obese patients undergoing Roux-en-Y bypass. Clin Gastroenterol Hepatol 2007;5:1329–1332. [PubMed: 17702661]

[26]. Jensen VS, Tveden-Nyborg P, Zacho-Rasmussen C, et al. Variation in diagnostic NAFLD/NASH read-outs in paired liver samples from rodent models. J Pharmacol Toxicol Methods 2020;101:106651. [PubMed: 31733366]

[27]. Wanless IR, Nakashima E, Sherman M. Regression of human cirrhosis. Morphologic features and the genesis of incomplete septal cirrhosis. Arch Pathol Lab Med 2000;124:1599–1607. [PubMed: 11079009]

[28]. 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;76:1030–1041. [PubMed: 35090960]

[29]. Abels E, Pantanowitz L, Aeffner F, et al. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. J Pathol 2019;249:286–294. [PubMed: 31355445]

[30]. Thrall MJ, Wimmer JL, Schwartz MR. Validation of multiple whole slide imaging scanners based on the guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Arch Pathol Lab Med 2015;139:656–664. [PubMed: 25927149]

[31]. Zarella MD, Bowman D, Aeffner F, et al. A practical guide to whole slide imaging: a white paper from the digital pathology association. Arch Pathol Lab Med 2019;143:222–234. [PubMed: 30307746]

[32]. Astbury S, Grove JI, Dorward DA, et al. Reliable computational quantification of liver fibrosis is compromised by inherent staining variation. J Pathol Clin Res 2021;7:471–481. [PubMed: 34076968]

[33]. Mostaço-Guidolin L, Rosin NL, Hackett TL. Imaging collagen in scar tissue: developments in second harmonic generation microscopy for biomedical applications. Int J Mol Sci 2017;18. [PubMed: 29267212]

[34]. Elyashberg ME, Blinov KA, Williams AJ, et al. Are deterministic expert systems for computer-assisted structure elucidation obsolete? J Chem Inf Model 2006;46:1643–1656. [PubMed: 16859296]

[35]. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20:e253–e261. [PubMed: 31044723]

[36]. Border SP, Sarder P. From what to why, the growing need for a focus shift toward explainability of AI in digital pathology. Front Physiol 2021;12:821217. [PubMed: 35087427]

[37]. Cheng JY, Abel JT, Balis UGJ, et al. Challenges in the development, deployment, and regulation of artificial intelligence in anatomic pathology. Am J Pathol 2021;191:1684–1692. [PubMed: 33245914]

[38]. Soon GST, Liu F, Leow WQ, et al. Artificial intelligence improves pathologist agreement for fibrosis scores in nonalcoholic steatohepatitis patients. Clin Gastroenterol Hepatol 2023;21:1940–1942 e3. [PubMed: 35697267]

[39]. Jahn SW, Plass M, Moinfar F. Digital pathology: advantages, limitations and emerging perspectives. J Clin Med 2020;9. [PubMed: 33374507]

[40]. Wanless IR. Quantitative SHG-microscopy: unraveling the nano-architecture of the cirrhotic liver. Clin Res Hepatol Gastroenterol 2020;44:1–3. [PubMed: 31416788]

[41]. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 2016;7:29. [PubMed: 27563488]

[42]. Iyer JS, Biddle-Snead C, Le Q. Machine learning-enabled continuous scoring of histologic features facilitates prediction of clinical disease progression in patients with non-alcoholic steatohepatitis. J Hepatol 2022:A3788.

[43]. Pokkalla H, Pethia K, Taylor-Weiner A. Machine learning models identify novel histologic features predictive of clinical disease progression in patients with advanced fibrosis due to non-alcoholic steatohepatitis. J Hepatol 2020;73:S402.

[44]. Travis N, Billaut V, Pokkalla H. Machine learning models accurately interpret liver histology and are associated with disease progression in patients with primary sclerosing cholangitis. J Hepatol 2020;73:S485.

[45]. Taylor-Weiner A, Pokkalla H, Han L, et al. A machine learning approach enables quantitative measurement of liver histology and disease monitoring in NASH. Hepatology 2021;74:133–147. [PubMed: 33570776]

[46]. Us Food and Drug Administration. FDA authorizes software that can identify prostate cancer. September 2021.

[47]. Ramot Y, Deshpande A, Morello V, et al. Microscope-based automated quantification of liver fibrosis in mice using a deep learning algorithm. Toxicol Pathol 2021;49:1126–1133. [PubMed: 33769147]

[48]. Wang Y, Vincent R, Yang J, et al. Dual-photon microscopy-based quantitation of fibrosis-related parameters (q-FP) to model disease progression in steatohepatitis. Hepatology 2017;65:1891–1903. [PubMed: 28133774]

[49]. Chang PE, Goh GBB, Leow WQ, et al. Second harmonic generation microscopy provides accurate automated staging of liver fibrosis in patients with non-alcoholic fatty liver disease. PLoS One 2018;13:e0199166. [PubMed: 29924825]

[50]. Vanderbeck S, Bockhorst J, Komorowski R, et al. Automatic classification of white regions in liver biopsies by supervised machine learning. Hum Pathol 2014;45:785–792. [PubMed: 24565203]

[51]. Sethunath D, Morusu S, Tuceryan M, et al. Automated assessment of steatosis in murine fatty liver. PLoS One 2018;13:e0197242. [PubMed: 29746543]

[52]. Munsterman ID, van Erp M, Weijers G, et al. A novel automatic digital algorithm that accurately quantifies steatosis in NAFLD on histopathological whole-slide images. Cytometry B Clin Cytom 2019;96:521–528. [PubMed: 31173462]

[53]. Goh GB, Leow WQ, Liang S, et al. Quantification of hepatic steatosis in chronic liver disease using novel automated method of second harmonic generation and two-photon excited fluorescence. Sci Rep 2019;9:2975. [PubMed: 30814650]

[54]. Marti-Aguado D, Fernandez-Paton M, Alfaro-Cervello C, et al. Digital pathology enables automated and quantitative assessment of inflammatory activity in patients with chronic liver disease. Biomolecules 2021:11. [PubMed: 35053159]

[55]. Liu F, Goh GB, Tiniakos D, et al. qFIBS: an automated technique for quantitative evaluation of fibrosis, inflammation, ballooning, and steatosis in patients with nonalcoholic steatohepatitis. Hepatology 2020;71:1953–1966. [PubMed: 31600834]

[56]. Vanderbeck S, Bockhorst J, Kleiner D, et al. Automatic quantification of lobular inflammation and hepatocyte ballooning in nonalcoholic fatty liver disease liver biopsies. Hum Pathol 2015;46:767–775. [PubMed: 25776030]

[57]. Teramoto T, Shinohara T, Takiyama A. Computer-aided classification of hepatocellular ballooning in liver biopsies from patients with NASH using persistent homology. Comput Methods Programs Biomed 2020;195:105614. [PubMed: 32650090]

[58]. Gawrieh S, Sethunath D, Cummings OW, et al. Automated quantification and architectural pattern detection of hepatic fibrosis in NAFLD. Ann Diagn Pathol 2020;47:151518. [PubMed: 32531442]

[59]. Marti-Aguado D, Rodríguez-Ortega A, Mestre-Alagarda C, et al. Digital pathology: accurate technique for quantitative assessment of histological features in metabolic-associated fatty liver disease. Aliment Pharmacol Ther 2021;53:160–171. [PubMed: 32981113]

[60]. 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;100:147–160. [PubMed: 31506634]

[61]. Leow WQ, Bedossa P, Liu F, et al. An improved qFibrosis algorithm for precise screening and enrollment into non-alcoholic steatohepatitis (NASH) clinical trials. Diagnostics (Basel) 2020;10. [PubMed: 33374618]

[62]. Bereket M, Casale FP, Loomba R. Machine learning encoding of liver biopsy images enables prediction of molecular measurements in a NASH F3/F4 fibrosis clinical cohort. J Hepatol 2022;77:S409.

[63]. Murugesan K Evidence for a robust artificial intelligence-based diagnostic support system in a clinical non-alcoholic fatty liver disease (NAFLD) scoring workflow. Hepatology 2021;74:954A.

[64]. Overgaard A, Salinas C, Eschen C, et al. Evaluation of pharmacological intervention using a deeplearning approach for histopathological scoring in a diet-induced obese and biopsy-confirmed mouse model of NASH with fibrosis. J Hepatol 2020;73:S682.

[65]. Harrison SA, Pulaski H, Vitali MC, et al. Analytical and clinical validation of AIM-NASH: a digital pathology tool for artificial intelligence-based measurement of nonalcoholic steatohepatitis histology. J Hepatol 2023;78:S32.

[66]. Kvilekval K, Fedorov D, Obara B, et al. Bisque: a platform for bioimage analysis and management. Bioinformatics 2010;26:544–552. [PubMed: 20031971]

[67]. Pérez-Bueno F, Vega M, Sales MA, et al. Blind color deconvolution, normalization, and classification of histological images using general super Gaussian priors and Bayesian inference. Comput Methods Programs Biomed 2021;211:106453. [PubMed: 34649072]

[68]. Azevedo Tosta TA, de Faria PR, Neves LA, et al. Computational normalization of H&E-stained histological images: progress, challenges and future potential. Artif Intell Med 2019;95:118–132. [PubMed: 30420242]

[69]. Zheng Y, Jiang Z, Zhang H, et al. Adaptive color deconvolution for histological WSI normalization. Comput Methods Programs Biomed 2019;170:107–120. [PubMed: 30712599]

[70]. Amin A, Mahmoud-Ghoneim D. Texture analysis of liver fibrosis microscopic images: a study on the effect of biomarkers. Acta Biochim Biophys Sin (Shanghai) 2011;43:193–203. [PubMed: 21258076]

[71]. FibroNest website.

[72]. Chen L, Lung M, Behling C. Evaluation of the multivendor performance of a novel histology-based fibrosis phenotypic composite score and its correlation with NASH-CRN fibrosis scores in patients with NASH. Hepatology 2021;74:953A.

[73]. Watson A, Petitjean L, Petitjean M. Etiology-independent fibrosis severity scoring by quantitative digital pathology image analysis. J Hepatol 2022;77:S482.

[74]. L C, Lung M, Petitjean M, et al. Evaluation of a novel histology-based fibrosis phenotypic composite score and its correlation with NASH-CRN fibrosis scores in patients with NASH. J Hepatol 2020;73:S421.

[75]. L C, Eerikson M, Adorini L, et al. Digital pathology image analysis accurately quantifies anti-fibrotic and anti-steatotic effects of FXR agonists using multiple histological methods. Hepatology 2021;74:1135A. [PubMed: 33811378]

[76]. Reynoso E, Chen L, Petitjean M, et al. Development of an optimal continuous pediatric fibrosis score to assess severity and progression of fibrosis in non-alcoholic fatty liver disease (NAFLD). Hepatology 2019;70:470A.

[77]. Iwaki M L C, Petitjean M, et al. Is the histological phenotype of fibrosis different between lean and obese NASH patients? Hepatology 2021;74:1021A.

[78]. L C, Fedorov D, Petitjean M, et al. Continuous staging of NASH patients at low (F1) fibrosis severity: evaluation of the performance of a novel histology-based fibrosis phenotypic composite score and predictive AI tools. Hepatology 2021;74:945A.

[79]. Petitjean L, Chen L, Mattis A. 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.

[80]. Tai DC, Tan N, Xu S, et al. Fibro-C-Index: comprehensive, morphology-based quantification of liver fibrosis using second harmonic generation and two-photon microscopy. J Biomed Opt 2009;14:044013. [PubMed: 19725725]

[81]. Sun W, Chang S, Tai DC, et al. Nonlinear optical microscopy: use of second harmonic generation and two-photon microscopy for automated quantitative liver fibrosis studies. J Biomed Opt 2008;13:064010. [PubMed: 19123657]

[82]. Wang Y, Wong GL, He FP, et al. Quantifying and monitoring fibrosis in nonalcoholic fatty liver disease using dual-photon microscopy. Gut 2020;69:1116–1126. [PubMed: 31563875]

[83]. Soon G, Wee A. Updates in the quantitative assessment of liver fibrosis for nonalcoholic fatty liver disease: histological perspective. Clin Mol Hepatol 2021;27:44–57. [PubMed: 33207115]

[84]. Naoumov NV, Brees D, Loeffler J, et al. Digital pathology with artificial intelligence analyses provides greater insights into treatment-induced fibrosis regression in NASH. J Hepatol 2022;77:1399–1409. [PubMed: 35779659]

[85]. Noureddin M, Goodman Z, Tai D, et al. Machine learning liver histology scores correlate with portal hypertension assessments in nonalcoholic steatohepatitis cirrhosis. Aliment Pharmacol Ther 2023;57:409–417. [PubMed: 36647687]

[86]. Xu S, Wang Y, Tai DCS, et al. qFibrosis: a fully-quantitative innovative method incorporating histological features to facilitate accurate fibrosis scoring in animal model and chronic hepatitis B patients. J Hepatol 2014;61:260–269. [PubMed: 24583249]

[87]. Gailhouste L, Le Grand Y, Odin C, et al. Fibrillar collagen scoring by second harmonic microscopy: a new tool in the assessment of liver fibrosis. J Hepatol 2010;52:398–406. [PubMed: 20149472]

[88]. Naoumov NV, Kleiner D, Chang E, et al. A deep exploration of bridging fibrosis evolution and individual septa parameters in nonalcoholic steatohepatitis using quantitative second harmonic generation imaging reveals fibrosis changes in natural history and treatment-induced not seen with conventional histology. J Hepatol 2023:2023.

[89]. Harrison SA, Bashir MR, Guy CD, et al. Resmetirom (MGL-3196) for the treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet 2019;394:2012–2024. [PubMed: 31727409]

[90]. Tai D, Bashir M, Taub R, et al. Impact of resmetirom-mediated reductions in liver volume and steatosis compared with placebo on the quantification of fibrosis using second harmonic generation in a serial liver biopsy study. J Hepatol 2022;77:S32.

[91]. Wang J, Pouryahya M, Leidal K. Liver biopsy graph neural networks for automated histologic scoring using the NASH CRN system. J Hepatol 2021;75:S602.

[92]. Harrison S, Iyer JS, Biddle-Snead C. Retrospective AI-based measurement of NASH histology (AIM-NASH) analysis of biopsies from Phase 2 study of resmetirom confirms significant treatment-induced changes in histologic features of non-alcoholic steatohepatitis. J Hepatol 2022;77:S711.

[93]. Loomba R, Noureddin M, Kowdley KV, et al. Combination therapies including cilofexor and firsocostat for bridging fibrosis and cirrhosis attributable to NASH. Hepatology 2021;73:625–643. [PubMed: 33169409]

[94]. Newsome PN, Buchholtz K, Cusi K, et al. A placebo-controlled trial of subcutaneous semaglutide in nonalcoholic steatohepatitis. N Engl J Med 2021;384:1113–1124. [PubMed: 33185364]

[95]. Harrison S, Francque S, Carrasco-Zevallos O. Artificial intelligence-powered digital pathology model supports that fibrosis is reduced by semaglutide in patients with NASH. Hepatology 2021;74:1131A. [PubMed: 33559205]

[96]. Loomba R, Cejvanovic V, Iyer JS. Comparison of the effects of semaglutide on liver histology in patients with non-alcoholic steatohepatitis cirrhosis between machine learning model assessment and pathologist evaluation. Hepatology 2022:76.

[97]. Shevell D, Brown E, Minnich A. Comparison of manual vs machine learning approaches to liver biopsy scoring for NASH and fibrosis: a post hoc analysis of the FALCON 1 study. Hepatology 2021:74. [PubMed: 34387904]

[98]. Ratziu V, Yilmaz Y, Lazas D. Multimodality assessment of hepatic fibrosis: ranked paired reading and artificial intelligence identifies fibrosis improvement with aramchol missed by conventional staging. J Hepatol 2022;77:S714.

[99]. L C, Brown C, Minnich A. Novel digital pathology quantitative image analysis and AI method detects the treatment effects of NASH drug candidates with a performance that benchmarks imaging-based measurements. J Hepatol 2022;77:S446.

[100]. Chen L, Minnich A, Baxi V, et al. Novel Digital Pathology quantitative image analysis and AI method detects traits of fibrosis treatment response. Hepatology 2023:78.

[101]. Kendall TJ, Tai D, Ho G, et al. Digital pathology using stain-free imaging indices allows direct prediction of all-cause mortality, hepatic decompensation and hepatocellular carcinoma development in patients with nonalcoholic fatty liver disease. J Hepatol 2023;78:S70.

[102]. Chen L, Petitjean L, Ampuero J, et al. Novel artificial intelligence-assisted digital pathology quantitative image analysis predicts the occurrence of liver-related clinical events in the multicentric, European, Hepatic OuTcomes and SURvival Fatty Liver Registry (HOTSURFR) study. J Hepatol 2023;78:S651.

[103]. Bosch J, Chung C, Carrasco-Zevallos OM, et al. A machine learning approach to liver histological evaluation predicts clinically significant portal hypertension in NASH cirrhosis. Hepatology 2021;74:3146–3160. [PubMed: 34333790]

[104]. Nakamura Y, Miyaaki H, Miuma S, et al. Automated fibrosis phenotyping of liver tissue from non-tumor lesions of patients with and without hepatocellular carcinoma after liver transplantation for non-alcoholic fatty liver disease. Hepatol Int 2022;16:555–561. [PubMed: 35553006]

[105]. Mirshahi F, Aqbi HF, Isbell M, et al. Distinct hepatic immunological patterns are associated with the progression or inhibition of hepatocellular carcinoma. Cell Rep 2022;38:110454. [PubMed: 35235789]

[106]. Chen L, Bhattacharya D, Friedman S, et al. Advanced quantitative phenotypic fibrosis and steatosis scoring is more superior to histology-based conventional staging in NASH animal models. J Hepatol 2022;77:S688.

[107]. Wang S, Li K, Pickholz E, et al. An autocrine signaling circuit in hepatic stellate cells underlies advanced fibrosis in nonalcoholic steatohepatitis. Sci Transl Med 2023;15:eadd3949. [PubMed: 36599008]

[108]. Albadrani M, Seth RK, Sarkar S, et al. Exogenous PP2A inhibitor exacerbates the progression of nonalcoholic fatty liver disease via NOX2-dependent activation of miR21. Am J Physiol Gastrointest Liver Physiol 2019;317:G408–g428. [PubMed: 31393787]

[109]. Van Herck MA, Vonghia L, Kwanten WJ, et al. Diet reversal and immune modulation show key role for liver and adipose tissue T cells in murine nonalcoholic steatohepatitis. Cell Mol Gastroenterol Hepatol 2020;10:467–490. [PubMed: 32360637]

[110]. Itoh M, Kato H, Suganami T, et al. Hepatic crown-like structure: a unique histological feature in non-alcoholic steatohepatitis in mice and humans. PLoS One 2013;8:e82163. [PubMed: 24349208]

[111]. Liu F, Chen L, Rao HY, et al. Automated evaluation of liver fibrosis in thioacetamide, carbon tetrachloride, and bile duct ligation rodent models using second-harmonic generation/two-photon excited fluorescence microscopy. Lab Invest 2017;97:84–92. [PubMed: 27918557]

[112]. Morrison MC, Verschuren L, Salic K, et al. Obeticholic acid modulates serum metabolites and gene signatures characteristic of human NASH and attenuates inflammation and fibrosis progression in Ldlr−/−.Leiden mice. Hepatol Commun 2018;2:1513–1532. [PubMed: 30556039]

[113]. van den Hoek AM, de Jong J, Worms N, et al. Diet and exercise reduce pre-existing NASH and fibrosis and have additional beneficial effects on the vasculature, adipose tissue and skeletal muscle via organ-crosstalk. Metabolism 2021;124:154873. [PubMed: 34478753]

[114]. Chen L, Brown E, Minnich A, et al. Novel Digital Pathology quantitative image analysis and AI method detects the treatment effect of NASH drug candidates with a performance that benchmarks Imaging based measurements. J Hepatol 2022;77:S446.

[115]. Abel JT, Dwivedi C, Nofallah S. Exploratory analyses of NASH histology using CRN scores derived from a multi-stain machine learning method. Hepatology 2022;76.

[116]. Chen RJ, Lu MY, Wang J, et al. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. IEEE Trans Med Imaging 2022;41:757–770. [PubMed: 32881682]

[117]. Pourhyaya M, Taylor-Weiner A, Pokkalla H. Integration of machine learning-based histopathology and hepatic transcriptomic data identifies genes associated with portal inflammation and ductular proliferation as predictors of disease progression in advanced fibrosis due to NASH. Hepatology 2020;72:358A.

[118]. García-Rojo M, De Mena D, Muriel-Cueto P, et al. New European union regulations related to whole slide image scanners and image analysis software. J Pathol Inform 2019;10:2. [PubMed: 30783546]

[119]. Kearney SJ, Lowe A, Lennerz JK, et al. Bridging the gap: the critical role of regulatory affairs and clinical affairs in the total product life cycle of pathology imaging devices and software. Front Med (Lausanne) 2021;8:765385. [PubMed: 34869473]

[120]. Colling R, Pitman H, Oien K, et al. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. J Pathol 2019;249:143–150. [PubMed: 31144302]

[121]. Us Food and Drug Administration. Medical device data systems, medical image storage devices, and medical image communications devices. Guidance for industry and Food nd drug administration staff. 2019. Published online September 2019, https://www.fda.gov/media/88572/download.

[122]. Us Food and Drug Administration. Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. 2021.

[123]. Liu F, Wei L, Leow WQ, et al. Developing a new qFIBS model assessing histological features in pediatric patients with non-alcoholic steatohepatitis. Front Med (Lausanne) 2022;9:925357. [PubMed: 35833109]

[124]. Vos MB, Abrams SH, Barlow SE, et al. NASPGHAN clinical practice guideline for the diagnosis and treatment of nonalcoholic fatty liver disease in children: recommendations from the expert committee on NAFLD (ECON) and the North American society of pediatric Gastroenterology, Hepatology and nutrition (NASPGHAN). J Pediatr Gastroenterol Nutr 2017;64:319–334. [PubMed: 28107283]

Keypoints

- Safe and effective therapies are urgently needed to treat patients with non-alcoholic steatohepatitis (NASH), but a major challenge is the requirement for conventional (*i.e.* light microscopic) histological analysis of liver biopsies in clinical trials to diagnose NASH for enrolment and efficacy assessment.
- Conventional assessment of liver biopsies is associated with inherent limitations, including pathologist subjectivity, lack of reproducibility, the co-existence of progressive and regressive fibrosis features, and categorical scoring systems that lack the sensitivity to detect drug-induced changes.
- Digital pathology (DP) workflows are established in other areas of medicine and can be used to automatically detect, localise, and quantify histological features, and in combination with artificial intelligence (AI), can score various histological parameters.
- Several DP/AI tools in development for NASH can quantitatively analyse the key histological features of NASH (*i.e.*, steatosis, inflammation, hepatocyte ballooning, and fibrosis) and could assist pathologists with grading and staging of liver biopsies in an accurate and reproducible manner and, importantly, provide continuous metrics for measuring treatment-induced changes.
- Regulatory processes for DP/AI tools vary between countries and are rapidly evolving, but currently, specific guidance is complex and limited.
- Development, validation, and standardisation of DP/AI tools is an area of high priority in the NASH field, particularly with respect to drug development, where the limitations of conventional histological assessments may have contributed to the failure of some trials to achieve their histological endpoints despite evidence of biological activity.

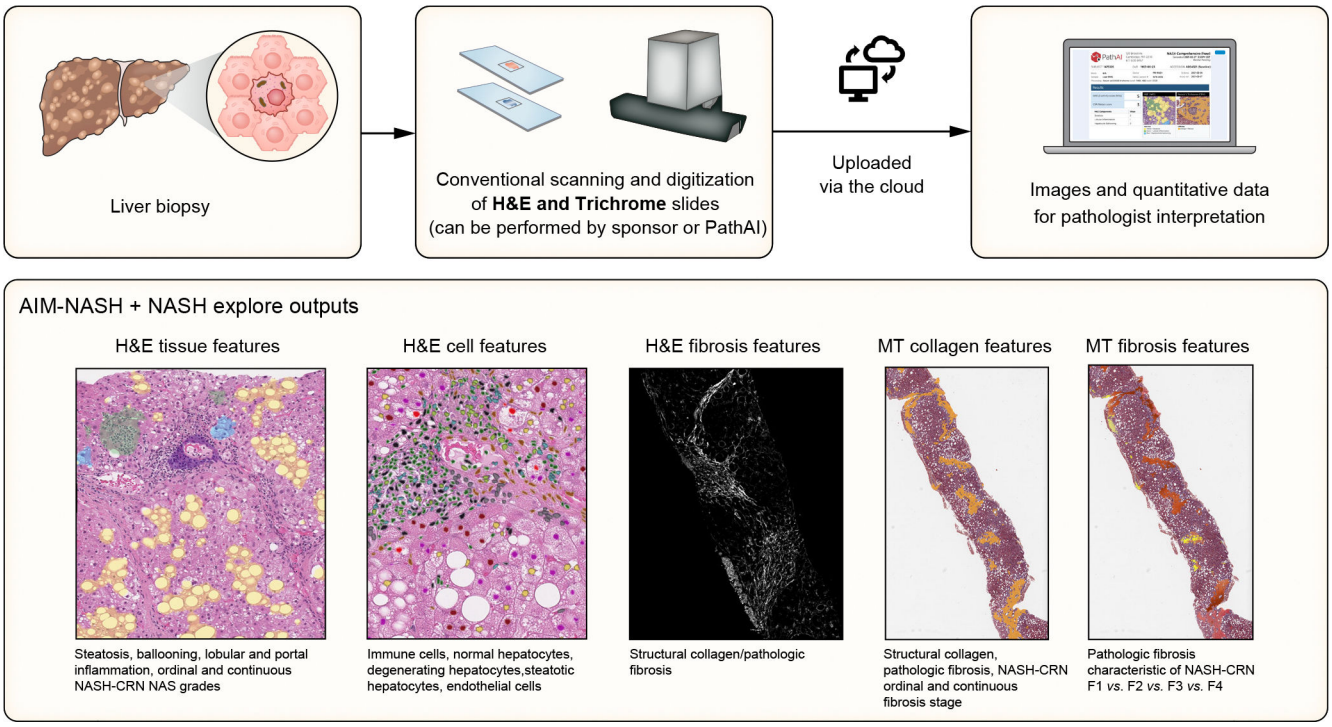

Fig. 1. AIM-NASH and NASH explore workflow and outputs.

Liver biopsy samples are stained with H&E and MT prior to conventional scanning and digitisation performed by either the user or at PathAI Diagnostics. WSIs are uploaded via the cloud and are evaluated first by artifact detection and exclusion algorithms, and subsequently by AIM-NASH and/or NASH-explore machine learning-based algorithms for tissue characterisation. WSI, heatmap overlays, and quantitative human interpretable features are either returned directly to the user or are reviewed by a pathologist who accepts or rejects the model-derived ordinal scores prior to delivery. Examples of AIM-NASH output images illustrating key NASH histological features are shown in the lower panel and were provided by Dr Janini Iyer of PathAI, Inc. AIM-NASH, AI-based histologic measurement non-alcoholic steatohepatitis; CRN, Clinical Research Network; MT, Masson's trichrome; NAS, non-Alcoholic fatty liver disease activity score; NASH, non-alcoholic steatohepatitis; WSI, whole-slide imaging.

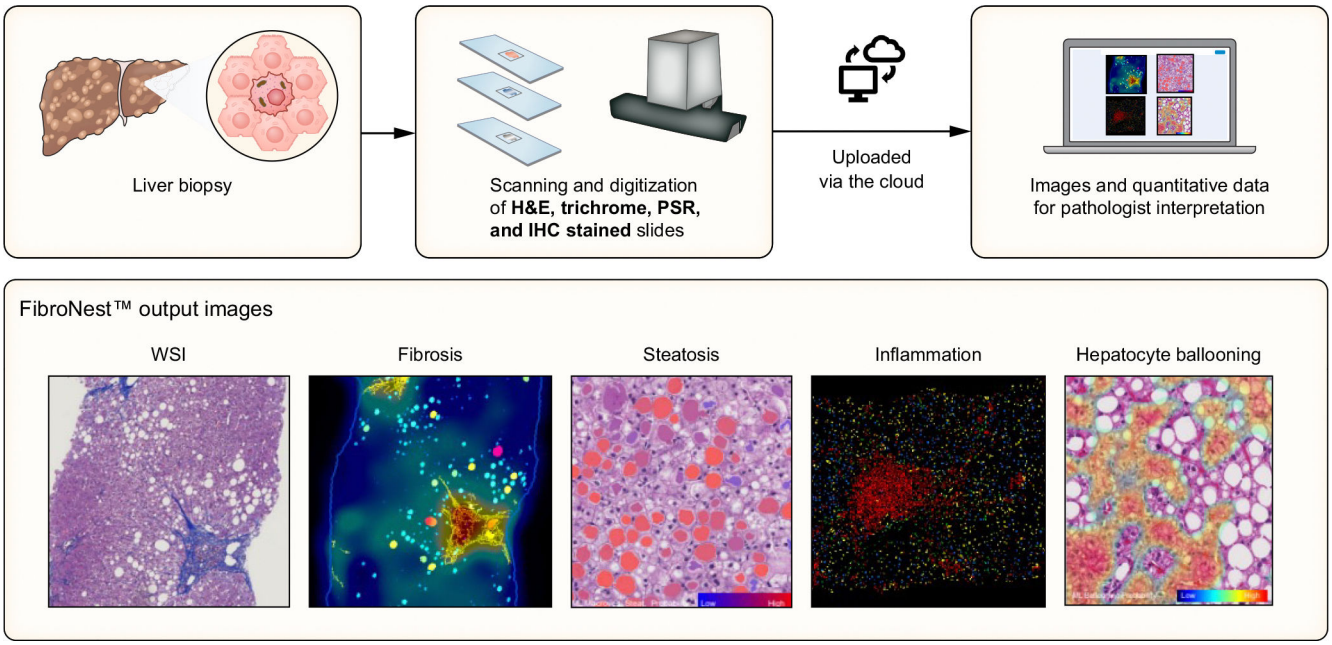

Fig. 2. FibroNest™ workflow and outputs.

H&E, Masson's trichrome, PSR, and IHC staining are performed on liver biopsy samples prior to scanning and digitisation. WSIs are uploaded via the cloud and analysed by FibroNest™ machine learning-based algorithms. Images and quantitative data are returned for pathologist review and interpretation. Examples of FibroNest™ output images illustrating key NASH histological features are shown in the lower panel and provided by Dr Mathieu Petitjean of PharmaNest Inc. IHC, immunohistochemistry; NASH, non-alcoholic steatohepatitis; PSR, picrosirius red; WSI, whole-slide images.

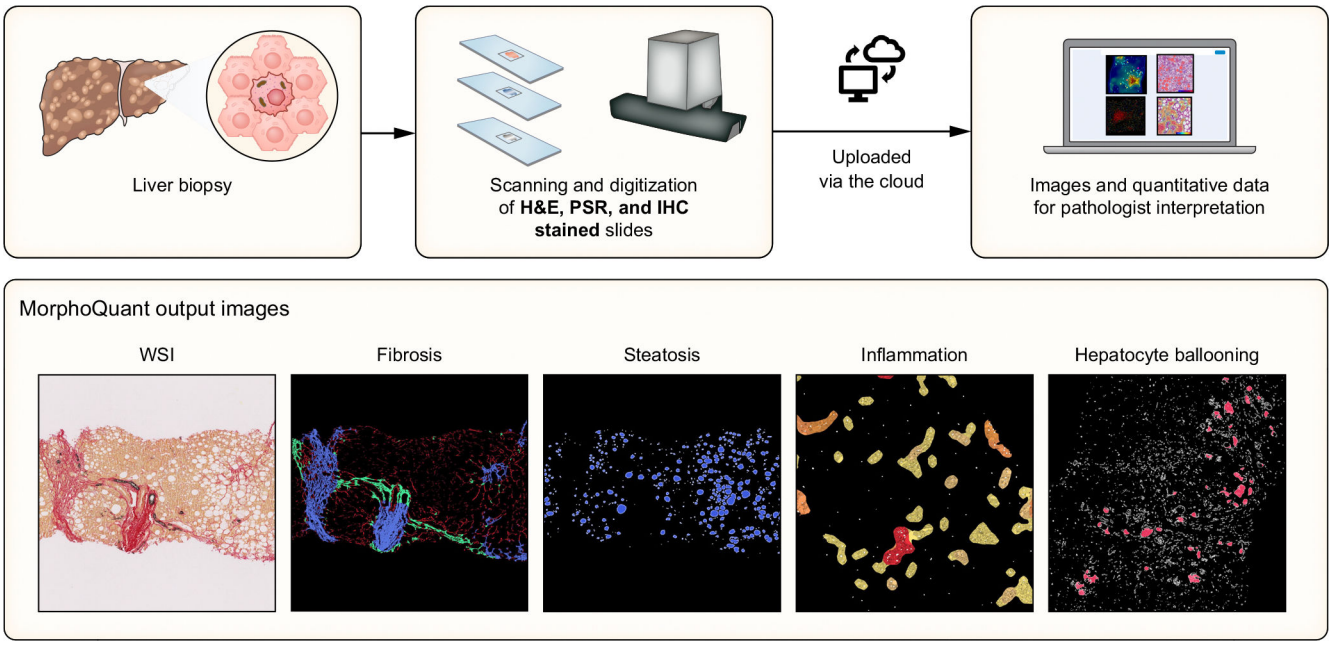

Fig. 3. MorphoQuant® workflow and outputs.

H&E, PSR, and IHC staining are performed on liver biopsy samples prior to scanning and digitisation. WSI are uploaded via the cloud and analysed by MorphoQuant® automated algorithms. Images and quantitative data are returned for pathologist review and interpretation. Examples of MorphoQuant® output images illustrating key NASH histological features are shown in the lower panel and provided by Dr Cindy Serdjebi of Biocellvia. IHC, immunohistochemistry; NASH, non-alcoholic steatohepatitis; PSR, picrosirius red; WSI, whole-slide imaging.

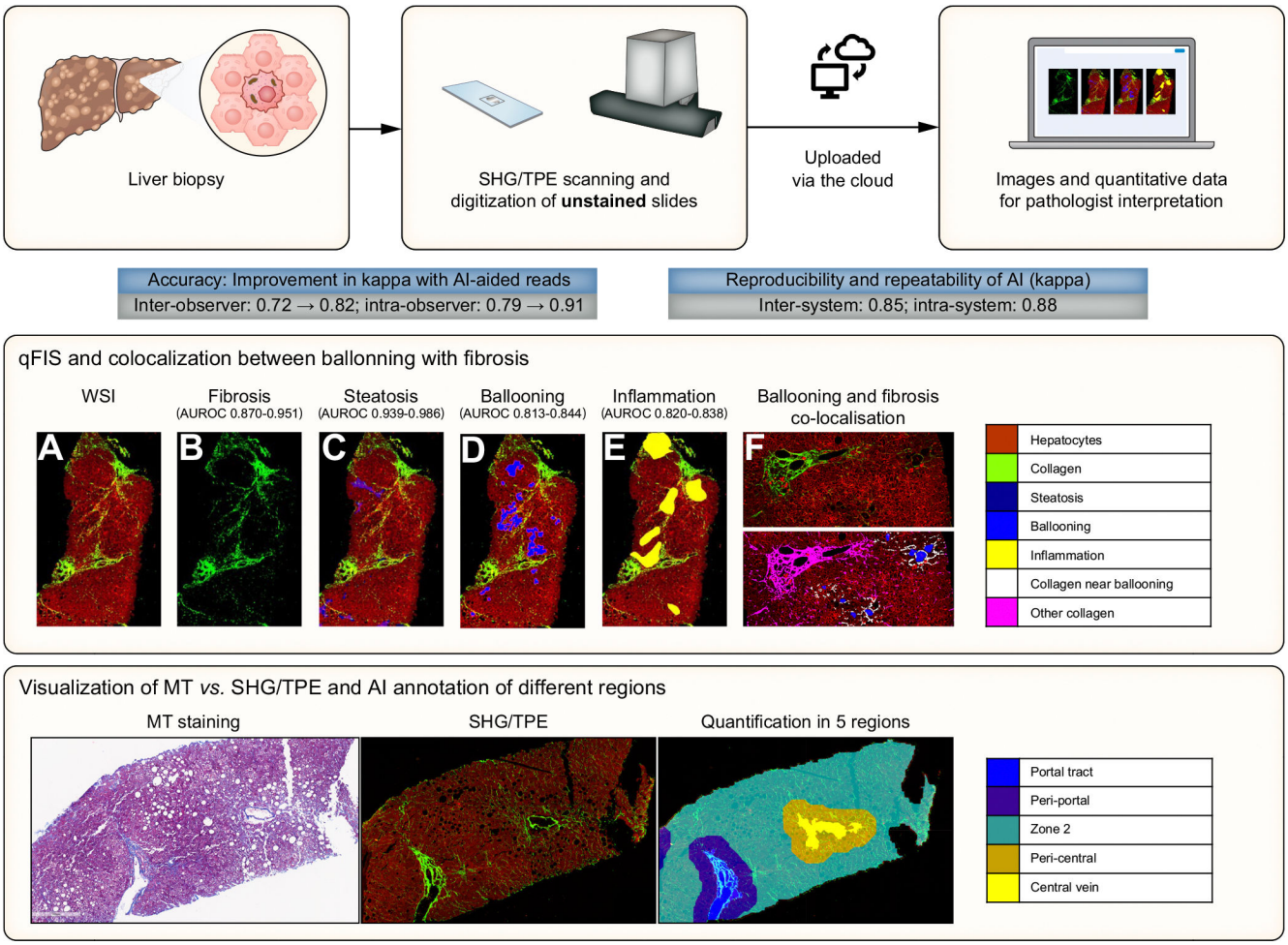

Fig. 4. qFibrosis, qInflammation, qBallooning, and qSteatosis workflow and outputs.

Unstained liver biopsy samples are imaged by SHG/TPE fluorescence. WSI are uploaded via the cloud and analysed by qFibrosis knowledge-based algorithms. Images and quantitative data are returned for pathologist review and interpretation. Examples of qFibrosis output images illustrating key NASH histological features are shown in the lower panel and provided by Dr Dean Tai of HistoIndex Pte Inc., as further described in Liu *et al.* and Naoumov *et al.*[62,91] AI, artificial intelligence; AUROC, area under the receiver operating characteristic; MT, Masson's trichrome; NASH, non-alcoholic steatohepatitis; SHG/TPE, second harmonic generation/two-photon excitation; WSI, whole-slide imaging.

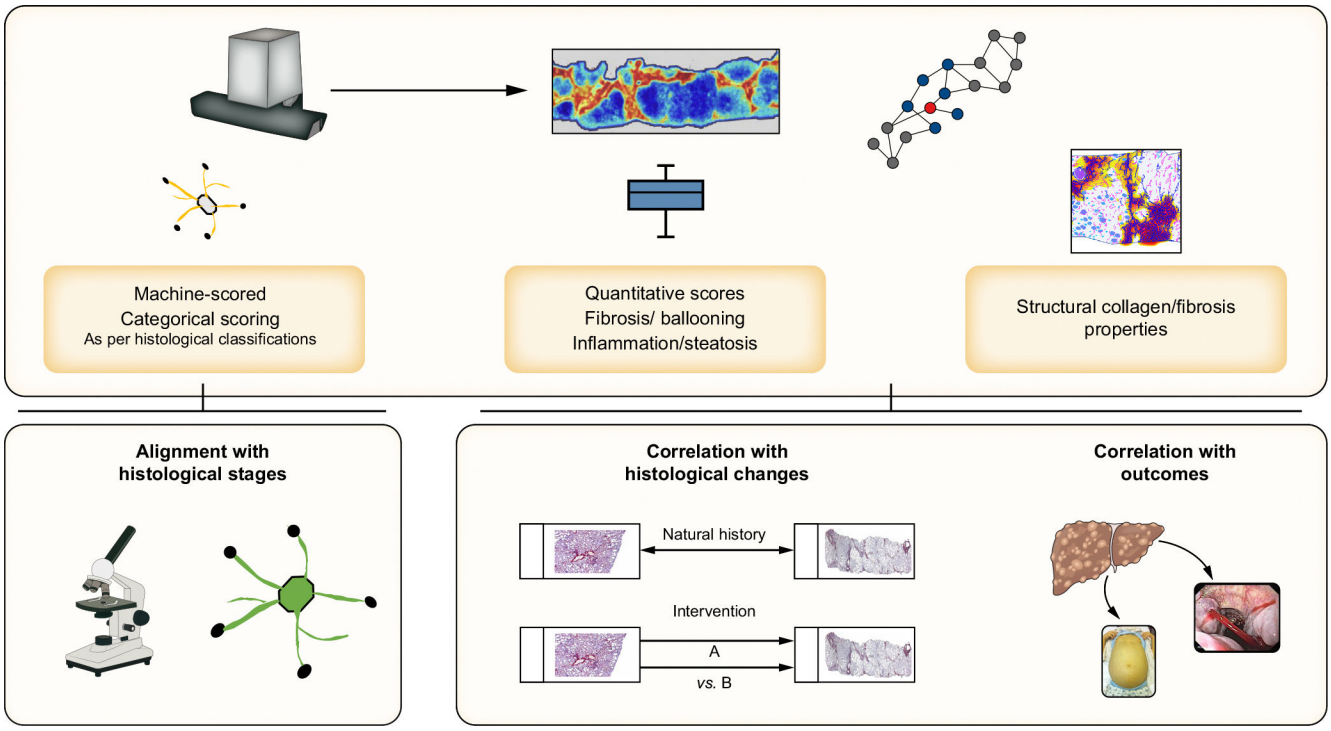

Fig. 5. A proposal for clinical validation of DP/AI diagnostic procedures for NASH.
AI, artificial intelligence; DP, digital pathology; NASH, nonalcoholic steatohepatitis.

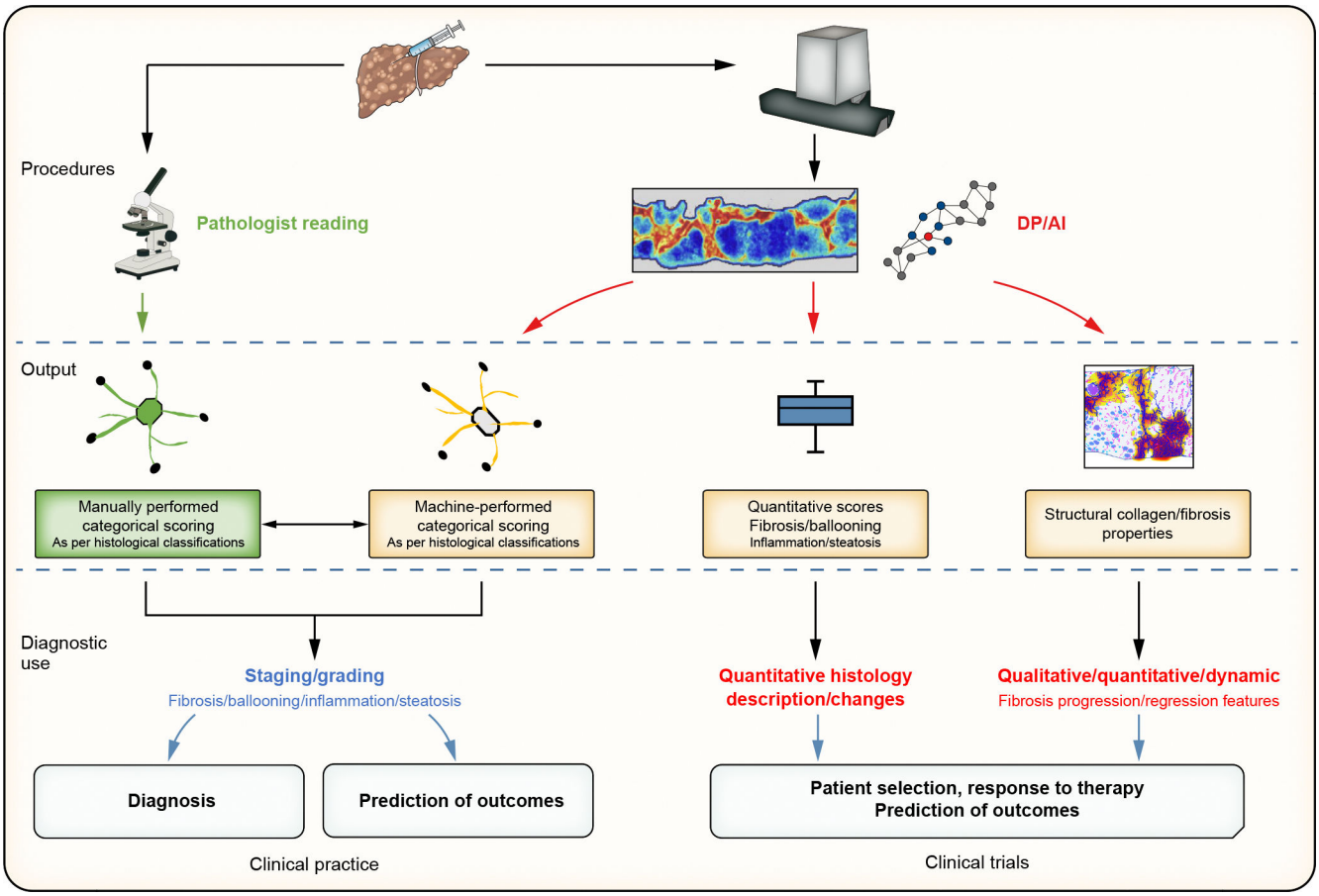

Fig. 6. Diagnostic capabilities of DP/AI methodologies applied to liver histology.
AI, artificial intelligence; DP, digital pathology.

J Hepatol. Author manuscript; available in PMC 2025 February 13.

Table 1.

Key features of NASH digital pathology/AI tools.

	AIM-NASH and NASH explore	**FibroNest™**	**MorphoQuant®**	**Histoindex®**
Input	• Digital WSI from slides stained with H&E (steatosis, inflammation, ballooning, fibrosis) and MT (fibrosis) • 40x recommended, 20x accepted • All digital formats accepted, Leica Aperio AT2 recommended • Also accepts tissue blocks and/or physical slides for slide preparation and scanning	• Digital WSI from slides stained with H&E or IHC (steatosis, inflammation, ballooning, cell injury) and MT, PSR, PSR-PL, or IHC (fibrosis, steatosis) • 40x preferred, 0.15–0.25 micron/pixel, 20x feasible • Compatible with all digital formats	• Digital WSI from slides stained with H&E (inflammation), PSR (steatosis and fibrosis-related endpoints), and IHC (ballooning, disease activity, hepatic crown-like structures, and any other label) • From 10x to 40x (at least 20x recommended) • Compatible with all digital formats	• Digital unstained WSI using SHG/TPE fluorescence, image tiles are acquired at 20x magnification with 512x512 pixel resolution with a dimension of 200x200 μm^2 • Multiple adjacent image tiles are captured to encompass the whole tissue in each slide • From 10x to 40x (20x recommended) • Compatible with all digital formats
High level workflow	• User shares WSIs that undergo AI-based artifact quality control to detect and exclude tissue and image artifacts from downstream analysis • H&E and MT AI algorithms are deployed to generate scores for all histological features • WSIs, corresponding heatmap overlays, and quantitative HIFs are made accessible to users via PathAI platforms	• Delivered via the cloud • Users upload WSI images to their accounts • Augmented pathology images are delivered back, with related quantitative scores	• Delivered via the cloud • Users upload WSI images to their accounts • Quantitative values for readouts and augmented pathology images delivered	• Delivered via the cloud • Users upload WSI (SHG/TPE fluorescence images) to their accounts • Pathology images in the form of SHG/TPE fluorescence images are delivered, along with quantitative scores
Methodology	• ML-based algorithms trained using annotations collected from liver biopsies across multiple clinical trials. Annotations provided by expert pathologists to facilitate interpretable model learning and outputs. Algorithms underwent rigorous analytical validation, demonstrating model performance accuracy, repeatability, and reproducibility • Convolutional neural networks used to train tissue segmentation and cell detection models to generate quantitative HIFs. All tissue and cell model outputs are displayed as heatmap overlays so pathologists can leverage these features in making grading/staging decisions • Parenchymal normalisation for presence of specific features such as fat performed at user's request • Graph neural networks used to train ordinal and continuous NASH-CRN scoring models	• High-content, single-fibre, and single-nucleus image analysis to quantify fibrosis, steatosis, and inflammation • ML-based algorithm trained by pathologists for ballooning • Quantitative AI to establish composite scores [1 to 10] to quantify the phenotypes of fibrosis, steatosis, inflammation, and ballooning • Parenchymal normalisation to decouple the cross-talk between steatosis and fibrosis severity/regression • Large quantitative data lakes are generated and can be exploited for investigational purposes	• Automated delineation of histological sections. Automated and quantitative detection and measurements of NASH features based on morphometric definitions and relevant standardised histology techniques • No ML or training of system, provides raw quantitative data for steatosis, fibrosis, inflammation, and ballooning with no interpretation • Parenchymal normalisation for features of interest	• Knowledge-based algorithm developed with more than 10 expert pathologists with validation in over 20 NASH clinical trials • Collagen is detected from the SHG channel, and collagen fibril characteristics are measured in liver tissue regions, including portal tract, peri-portal, peri-sinusoidal, central vein, and peri-central; this interpretation is logical for pathologist assessment • Parenchymal normalisation of these collagen characteristics
Output	• Tissue segmentation models detect and quantify tissue-based features such as steatosis, ballooning, lobular and portal inflammation, and fibrosis patterns relevant to NASH disease severity. Fibrosis analysis is performed after model-derived distinction between normal and pathologic fibrosis • Cell detection models identify lymphocytes, neutrophils, normal hepatocytes, degenerating hepatocytes, and endothelial cells amongst several other cell types encountered in the NASH tissue environment. Count proportions, densities, and proximity features are computed • All tissue and cell features are displayed to the pathologist as heat-map overlays such that the pathologist	• Digital biopsy adequacy score • Continuous, parenchymal scores for fibrosis, steatosis, inflammation, and ballooning • Quantitative data for the morphometric, architectural, and phenotypic description of collagen fibres, macrosteatosis vacuoles, and inflammatory clusters • Tissue panel (normalised quantity and % of each cell's category) • Augmented pathology images with image analysis features	• Quantitative data: o Steatosis area, mean vesicle size for steatosis, total inflammation area, number of inflammatory foci, hepatic crown-like structures • Total collagen, perivascular, septal peri-sinusoidal, periductular collagens • Ductular reaction, Shh expression, area of active injury • Supportive augmented pathology images with overlayed detected features	• Composite index made up of more than 100 parameters for fibrosis and takes into consideration septa and nodule parameters in a cirrhotic cohort • AI can also provide zonal fibrosis annotations • qBallooning and qSteatosis indexes are also available, which are made up of more than 40 and 60 ballooning- and steatosis-related parameters, respectively

	AIM-NASH and NASH explore	FibroNest™	MorphoQuant®	Histoindex®
	can leverage these features in making grading/staging decisions • Scoring models predict NASH-CRN grades for steatosis, ballooning, and lobular inflammation, as well as fibrosis stage, in addition to generating interpretable continuous scores that correspond to each feature and ordinal category			
Regulatory framework	• Letter of Intent accepted, Qualification Plan approved, analytical validation studies completed • PathAI Clinical Trial Services Platform is GCP compliant • PathAI AI Sight DX Platform is cleared by the FDA and EMA for primary diagnosis in clinical settings • PathAI Translational Science Platform is for research use only	• GCP compliant • Research use only • FDA BQP Letter of Intent in Q1 2023	• Working towards GCP compliance	• Research use only • FDA IVD and BQP application in process

AI, artificial intelligence; AIM-NASH, AI-based histologic measurement non-alcoholic steatohepatitis; BQP, Biomarker Qualification Program; CRN, Clinical Research Network; EMA, European Medicines Agency; FDA, US Food and Drug Administration; GCP, Good Clinical Practice; HIF, human interpretable feature; IHC, immunohistochemistry; IVD, *in vitro* diagnostic; ML, machine learning; MT, Masson's trichrome; NASH, non-alcoholic steatohepatitis; PSR, picrosirius red; SHG/TPE, second harmonic generation/two-photon excitation; Shh, Sonic hedgehog; WSI, whole-slide image.

J Hepatol. Author manuscript; available in PMC 2025 February 13.

Table 2.

Diagnostic and clinical relevance of DP/AI methodologies applied to human NASH.

	AIM-NASH®	Histoindex®	FibroNest™	MorphoQuant®
Requirement for staining	Yes	No	Yes	Yes
Output measures generated				
Hepatic fibrosis	Yes	Yes	Yes	Yes
Other NAFLD histological lesions *	Yes	Yes	Yes#	Yes
Quantitative, continuous scores	Yes	Yes	Yes (fibrosis)	Yes
Collagen fibrillar structure features	No	Yes	Yes	Yes
Diagnostic and prognostic value				
Identification of NASH-CRN fibrosis stages	Yes	Yes	Yes	No
Fibrosis changes	Fibrosis regression in trials[90–96]	Fibrosis regression in trials[83,87,89] Fibrosis progression[89]	Fibrosis regression in trials[77,97,103]	NA
Prediction of clinical outcomes	Clinical events[42] Portal hypertension§,14,78 Progression to cirrhosis[42]	Clinical events[82,101]	Clinical events[102]	NA

AI, artificial intelligence; AIM-NASH, AI-based histologic measurement non-alcoholic steatohepatitis; DP, digital pathology; n.a., not applicable; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; NASH-CRN, NASH Clinical Research Network.

#Except for hepatocyte ballooning.

*Steatosis, inflammation hepatocyte ballooning.

§Portal hypertension encompasses hepatic venous portal gradient changes, clinically significant portal hypertension or development of varices.

J Hepatol. Author manuscript; available in PMC 2025 February 13.