Machine learning-enabled continuous scoring of histologic features facilitates prediction of clinical disease progression in patients with non-alcoholic steatohepatitis

Abstract #3788

STUDY BACKGROUND

Non-alcoholic steatohepatitis (NASH) disease severity is typically assessed via ordinal scoring of tissue biopsies[1]. However, ordinal systems lack the sensitivity required to detect the heterogeneity that exists in liver tissue or capture subtle changes that may indicate improvement or worsening of disease[2]. Additionally, pathologist scoring of NASH histologic features (steatosis, ballooning, lobular inflammation, and fibrosis) is difficult, error-prone, and subject to intra- and inter-reader variability[3].

Here, we report machine learning-facilitated continuous scoring of histologic features by PathAI's AI-based Measurement of NASH Histology (AIM-NASH) algorithms in a retrospective analysis of a datasets from two completed NASH clinical trials. The prognostic utility of continuous scores was evaluated and compared against ordinal scoring for predicting patient outcomes.

MAPPING OF CONTINUOUS SCORES

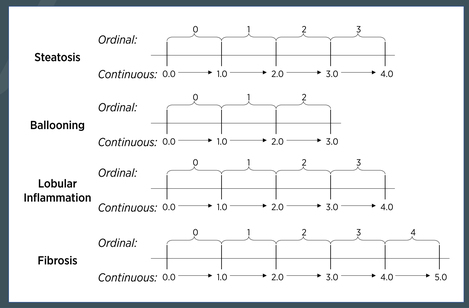

Figure 2. Continuous NASH scores were produced by mapping GNN-derived ordinal grades/stages to bins, such that ordinal scores are spread over a continuous range (unit distance of 1). Inter-bin cutoffs were learned during model training. Outer bins have long-tailed distributions that are not penalized during training. To ensure balanced linear mapping of outer bins, logit values in the first and last bins were restricted during post-processing to minimum and maximum values, respectively. These values were defined by outer-edge cutoffs chosen to maximize the uniformity of logit value distributions across training data. All GNN continuous feature training and ordinal mapping was performed for each NASH CRN histologic component separately.

METHODS

- Liver biopsies and corresponding central pathologist (CP) Ishak and Clinical Research Network (CRN) disease severity scores were collected from subjects enrolled in the Ph3, placebo-controlled STELLAR (ST) trials of selonsertib (NCT03053050 and NCT03053063)[4]. All subjects had CRN fibrosis stage 3 or 4 at baseline (BL) and were assessed for progression to either cirrhosis or adjudicated liver-related events (LREs), respectively. Treatment arms were combined for the present analysis due to lack of measured drug efficacy.

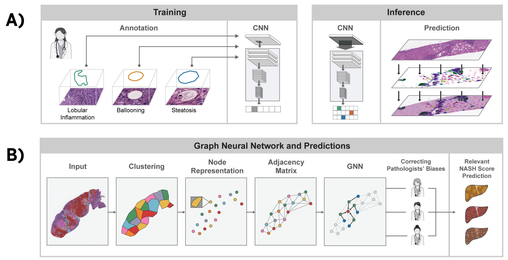

Figure 1. (A) Schematic illustrating convolutional neural network (CNN) model development approach for pixel-level feature predictions. (B) Schematic illustrating graph neural network (GNN) model development approach for slide-level model-derived CRN grading and staging.

- The AIM-NASH algorithms were developed using roughly 16,000 H&E and trichrome whole slide images (WSIs) from 6 NASH clinical trials and from patients with CRN stage F0-F4 fibrosis.
- Pathologist NASH CRN grades and stages and feature annotations were used to train machine learning models to generate pixel-level feature predictions in WSIs in addition to slide-level score predictions (**Fig. 1a**; manuscript in preparation). Continuous NAS and CRN fibrosis scores were produced by mapping GNN-derived ordinal grades/stages to bins, such that ordinal scores are spread over a continuous range (unit distance of 1) (**Fig. 1b, Fig. 2**).
- Algorithms were deployed on 1519 WSIs of baseline biopsies (N=725 pts from STELLAR 3, N=794 pts from STELLAR 4) to generate ML-derived predictions for ordinal and continuous CRN scores for steatosis, hepatocellular ballooning, lobular inflammation, and fibrosis.

RESULTS

- GNN continuous score interpretability was evaluated by correlating continuous scores against mean scores across three pathologists in a held-out validation dataset (completed Ph2 NASH clinical trial that enrolled patients with F1-3 fibrosis). When ordinal bins (**Fig. 2**) were determined by AIM-NASH (as opposed to median pathologist consensus), continuous scores were significantly correlated with mean pathologist scores, confirming alignment between ML-derived continuous scores and directional bias of panel-based pathologist scoring (**Fig.3**).

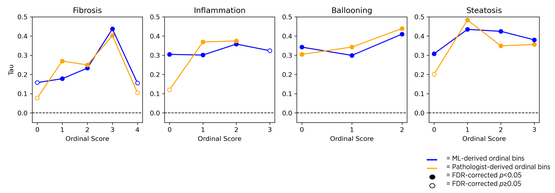

Figure 3. ML GNN NASH CRN continuous scores are correlated with mean pathologist scores per feature per disease severity grade. Kendall rank correlation coefficients are shown as a function of NASH CRN ordinal bin determined by AIM-NASH (blue) and median pathologist consensus (orange).

- During a median follow-up of 16.5 or 15.8 months, 14.9% (N = 108/725) of F3 subjects progressed to cirrhosis, and 2.5% (N = 20/794) of F4 subjects had liver-related events (LREs), respectively[4]. Algorithm-derived continuous and ordinal CRN scores, plus pathologist scores of biopsies collected at BL were evaluated for association with progression to cirrhosis or LRE (**Table 1**).

Baseline Feature	HR (95% CI)	p-Value
Patients with bridging fibrosis (F3)		
ML NASH CRN continuous fibrosis stage	3.91 (2.28, 6.70)	<0.0001
ML NASH CRN ordinal fibrosis stage	2.38 (1.59, 3.57)	<0.0001
Ishak fibrosis stage (pathologist)	1.72 (1.18, 2.52)	0.005
ML NASH CRN continuous steatosis grade	0.77 (0.60, 0.99)	0.038
CRN steatosis grade (pathologist)	1.40 (0.76, 2.58)	0.27
Patients with cirrhosis (F4)		
ML NASH CRN continuous steatosis grade	0.46 (0.22, 0.99)	0.049
CRN steatosis grade (pathologist)	1.15 (0.15, 8.52)	0.89
Ishak fibrosis stage (pathologist)	6.81 (1.59, 29.09)	0.010

Table 1. Association between algorithm- and pathologist-derived scores and progression to cirrhosis or LRE. HR, hazard ratio.

- Associations between ML-derived continuous scores at baseline and clinical disease progression through end of follow-up were determined using Kaplan-Meier and Cox proportional hazards regression analysis. Rounded cutoffs were chosen to maximize hazards.
- BL continuous fibrosis stage cutoffs of 3.6 and 4.6 stratified subjects into slow vs. rapid progressors to cirrhosis or LREs, respectively (Kaplan-Meier analysis, **Fig. 4**).
- Subjects with BL F3 and continuous scores above the cutoff of 3.6 will progress more rapidly to cirrhosis. Subjects with BL F4 and continuous scores above the cutoff of 4.6 will progress more rapidly to LRE.

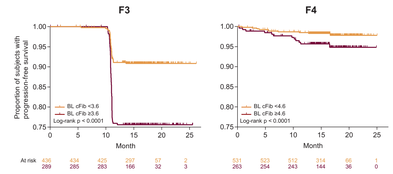

Figure 4. Continuous score cut-offs stratify subjects with BL F3 or F4 fibrosis into rapid and slow progressors.

- Algorithm-derived continuous scoring showed higher discriminatory accuracy to predict progression to cirrhosis and to LREs than algorithm-derived ordinal scoring (receiver operating characteristic analysis, **Fig. 5**).

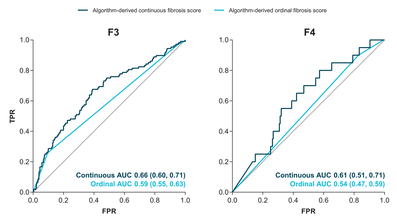

Figure 5. Comparing algorithm-derived continuous and ordinal scores for predicting disease progression to cirrhosis (left), and LRE (right). In both cases, the continuous AUC is significantly greater at one-sided p<0.025.

CONCLUSIONS

Application of the AIM-NASH-algorithm to WSIs from two completed Ph3 NASH clinical trials demonstrates the utility of ML-derived continuous histologic scoring methods relative to ML or pathologist ordinal scoring for predicting disease progression in patients with NASH and F3 or F4 fibrosis.

These results support further investigation into the value of ML-based continuous histologic scoring methods for detecting sub-ordinal yet clinically meaningful therapeutic effect in NASH clinical trials.

AUTHORS

Janani S. Iyer[1], Charles Biddle-Snead[1], Quang Le[1], Pratik Mistry[1], Isaac Finberg[1], Victoria Mountain[1], Chuhan Chung[2], Robert P. Myers[3], Andrew Billin[2], Tim Watkins[2], Ilan Wapinski[1], Michael Montalto[1], Andrew Beck[1], Murray Resnick[1], Katy Wack[1]

[1]PathAI, Inc., Boston, Massachusetts
[2]Gilead Sciences, Inc., Foster City, California
[3]The Liver Company, Inc., Palo Alto, California

contact email: katy.wack@pathai.com

REFERENCES

1. Kleiner, DE, et al. Hepatol. 2005; 41:1313-1321.
2. Taylor-Weiner, A, et al. Hepatol. 2021; 74:133-147.
3. Davison, BA et al. J Hepatol. 2020;73:1322-1332.
4. Harrison, SA, et al. J. Hepatol. 2020;73:26-39.

ACKNOWLEDGMENTS

We thank Bioscience Communications for assistance with figure design, and M.Lin (PathAI) for poster preparation.
This poster template was developed by SciStories LLC.
https://scistories.com/

PathAI

The International Liver Congress, 22-26 June 2022, London, UK