Validation of the HistoIndex AI digital pathology platform as an aiding tool to increase pathologist concordance on fibrosis staging in NASH

Desiree Abdurrachim[1], Serene Lek[2], Yongqi Zhou[1], Chun Kit Wong[1], Charlene Ong[1], Ashmita Saigal[3], Radha Krishnan[4], Geoffrey Walford[5], Elaine Chng[6], Dean Tai[6], Gideon Ho[6], Annaswamy Raji[7], Thomas Forest[8], Richard Baumgartner[7], Saswata Talukdar[3], Chih-Liang Chin[3], Samuel S Engel[7], Asad Abu Bakar Ali[1], David E Kleiner[9], Arun J Sanyal[10]

[1] MSD, Singapore, [2] Clinnovate Health UK Ltd, UK, [3] Merck & Co., Inc., South San Francisco, CA, USA, [4] MSD (UK) Limited, London, UK, [5] Merck & Co., Inc., Boston, MA, USA, [6] HistoIndex Pte Ltd, Singapore, [7] Merck & Co., Inc., Rahway, NJ, USA, [8] Merck & Co., Inc., West Point, PA, USA, [9] National Cancer Institute, NIH, USA, [10] Department of Internal Medicine, Virginia Commonwealth University, USA

AASLD The Liver Meeting® Nov. 10-14, 2023

INTRODUCTION

- Intra- and inter-observer variability in histological staging of fibrosis in NASH clinical trials lead to suboptimal selection of patients and confound assessment of fibrosis response.
- Efforts in reducing pathologist variability as well as improving scoring accuracy are therefore critical for a reliable NASH clinical trial.
- Aim: To prospectively evaluate the utility of the HistoIndex artificial intelligence (AI) tool to improve the reliability of fibrosis staging in NASH

METHODS

- Histology slides from two trials (NCT #03517540, #03912532) including 80 baseline/screening biopsies and 40 paired baseline and end-of-treatment biopsies were read by 4 pathologists in a cross-over modality design (Fig. 1), either without AI or with AI (Fig. 2).
- Fibrosis stage distribution (based on pathologist median without AI) is F0: 6, F1: 12, F2: 48, F3: 27, F4: 25.

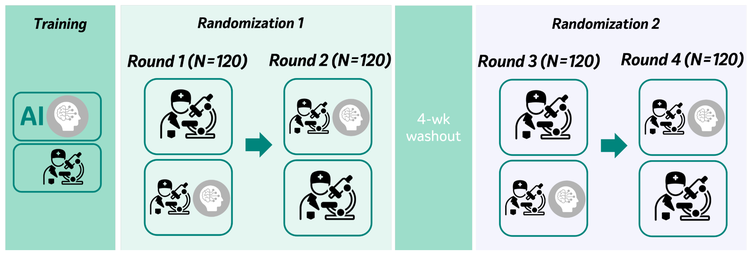

Fig 1. Study design. 4 expert hepato-pathologists, masked to each other, read a total of 120 biopsy sections twice each, masked to study source, with and without the AI aiding tool respectively, in random order reading 30 biopsies each week. The process was repeated after a 4-week washout.

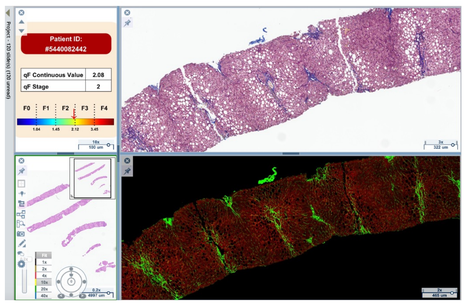

Fig 2. Evaluation platform. During the session with AI, pathologists were provided with an unstained second harmonic generation/two photon excitation fluorescence (SHG/TPEF) image (D), along with the AI quantitative fibrosis (qF) continuous values and the corresponding qF stage (A), in addition to the conventional H&E (B) and Masson's Trichrome image (C).

RESULTS

Key result #1: AI-assisted reads improved inter-observer kappa for fibrosis staging (Fig. 3A and Fig. 4), with the greatest impact shown for F0-F2 population (Fig. 3B). In clinical trials, this kappa improvement would have reduced the number of cases requiring adjudication by a third reader by 30%.

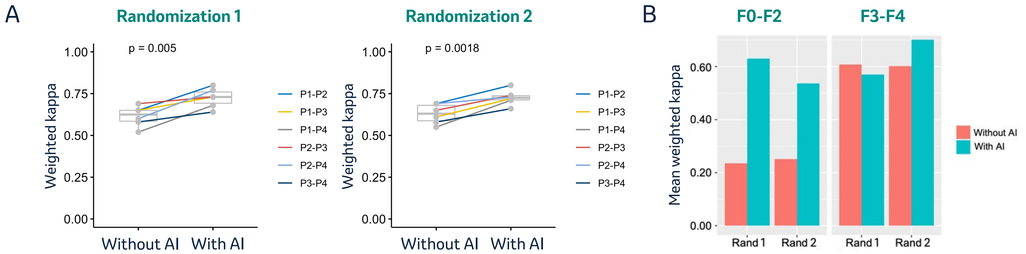

Fig 3. (A) Inter-reader weighted kappa for fibrosis staging without vs. with AI aiding, (B) mean inter-reader weighted kappa when evaluated separately for early-stage fibrosis (F0-F2 population) and late-stage fibrosis (F3-F4 population)

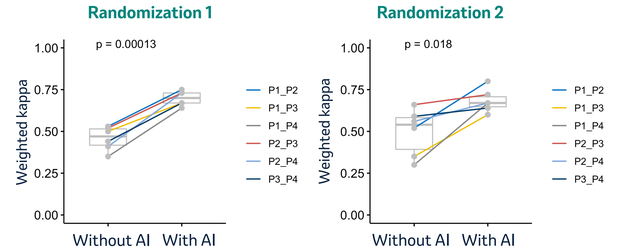

Fig 4. Inter-reader weighted kappa for evaluation of inclusion/exclusion to clinical trial (F2-F3 vs. F0, F1, F4) .

Key result #2: AI-assisted reads improved the rates of concordance between 4 pathologists for inclusion of NASH with F2-F3, exclusion of NASH (F0, F1, F4), and assessment of fibrosis response (Fig. 5). This increase was associated with decreased variance around the median reads.

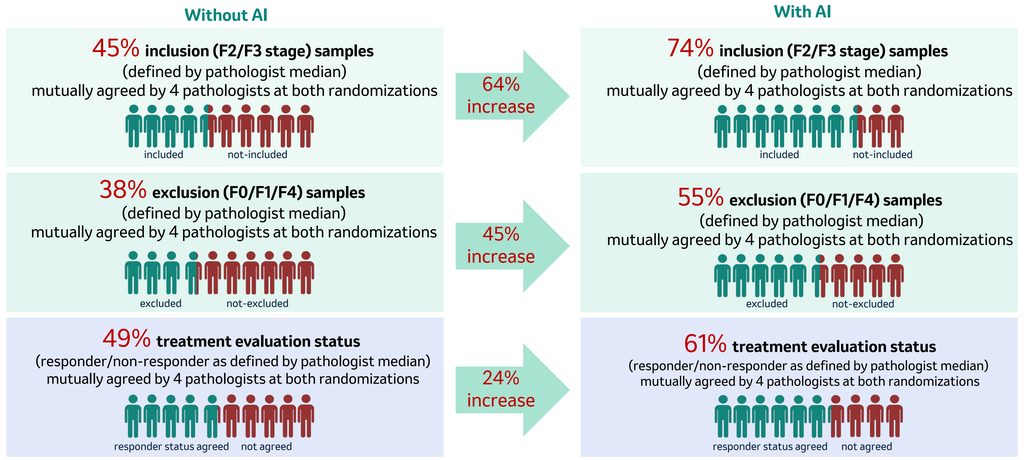

Fig 5. Rates of concordance between 4 pathologists without AI vs. with AI (green man: proportion agreed by 4 pathologists)

Key result #3: Overall, at least 3 out of 4 pathologists considered SHG/TPEF image useful in 83% cases (Fig. 6A) and qF values useful in 55% cases (Fig. 6C); this was greatest for F1-F2 (Fig. 6B and Fig. 6D).

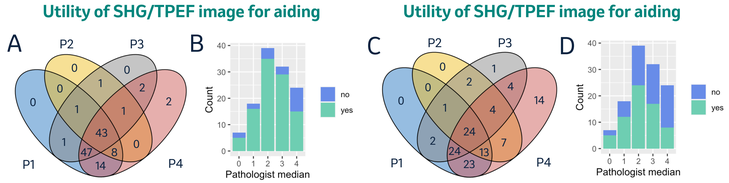

Fig 6. Venn diagram for the number of samples pathologists found (A) SHG useful or (C) qFibrosis useful, with the corresponding distribution for the different F-stage in which at least 3 pathologist found (B) SHG useful or (D) qFibrosis useful.

CONCLUSION

- HistoIndex AI tool enhances pathologist confidence and inter-rater reliability for assessment of fibrosis stage in NASH.
- They validate the utility of SHG/AI as an aid for pathologist assessment of fibrosis.
- These data support the use of SHG/AI to enhance the efficiency of clinical trials and reliability of fibrosis readouts of response from trials.

Contact: desiree.abdurrachim@msd.com | serene.lek@histoindex.com | arun.sanyal@vcuhealth.org