WED-436

Validation and utility of artificial intelligence-based zonal annotations as an additional assessment tool for the histopathologic review of fibrosis in non-alcoholic steatohepatitis patients

Gwyneth Soon[1], Aileen Wee[2], Wei Qiang Leow[3], Elaine Chng[4], Dean Tai[4], Yayun Ren[4], Feng Liu[5], Lai Wei[6], Arun Sanyal[7]

[1]National University Hospital, Department of Pathology, Singapore, *[2]Yong Loo Lin School of Medicine, National University of Singapore, Department of Pathology, Singapore*, *[3]Singapore General Hospital, Singapore and Duke-NUS Medical School, Department of Anatomical Pathology, Singapore*, *[4]HistoIndex Pte Ltd, Singapore*, *[5]Peking University Hepatology Institute, Peking University People's Hospital, China*, *[6]Hepatopancreatobiliary Center, Beijing Tsinghua Changgung Hospital, China*, *[7]Stravitz-Sanyal Institute of Liver Disease and Metabolic Health, VCU School of Medicine, United States*

Email: gwyneth_st_soon@nuhs.edu.sg

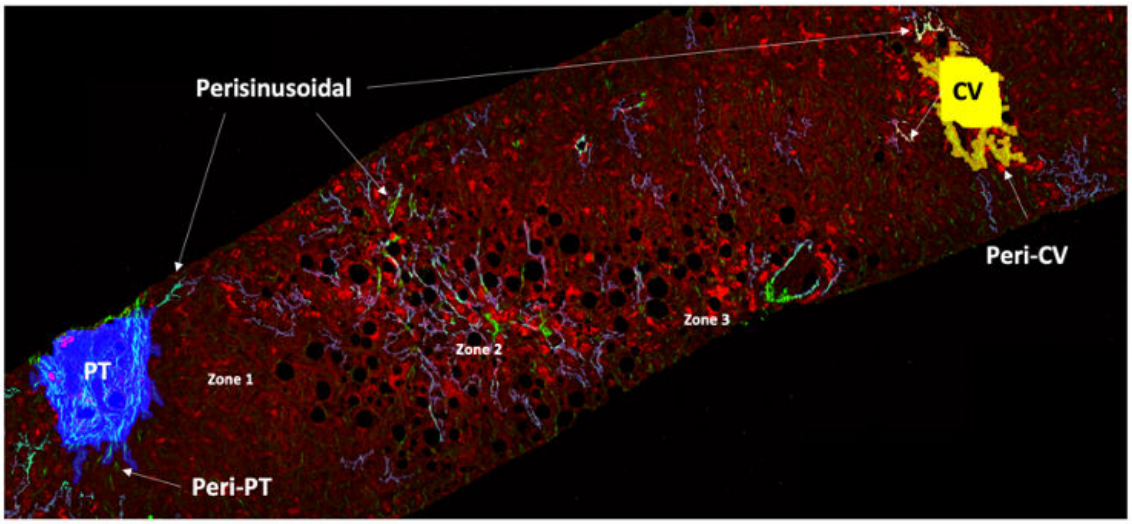
Background and Aims: Inter-observer variability for categorical scores of liver fibrosis among pathologists ranges from fair to moderate weighted kappa. Artificial intelligence (AI) and advances in digitised whole-slide imaging (WSI) have facilitated the use of AI-assistive tools in pathology to improve histopathologic interpretation. Second harmonic generation/two-photon excitation fluorescence (SHG/TPEF) microscopy with qFibrosis staging and continuous values as AI- assistive tools has been shown to help standardize pathologist assessment, contributing to higher overall intra- and inter-rater agreements. With the additional provision of AI-based zonal annotations, we aim to explore whether there will be further improvements on the overall inter-pathologist agreement on fibrosis assessment

Method: Unstained sections of liver biopsies from 50 untreated NASH patients (F0-F4) were evaluated. Fibrosis was quantitated using SHG/TPEF microscopy (qFibrosis). Unassisted reads comprised digitized H&E and Masson trichrome-stained images uploaded to a WSI platform. Level I assisted reads included additional SHG images with qFibrosis outputs. Level II assisted reads further included zonal annotations. The zonal annotations serve to highlight the portal tract (PT), central vein (CV), peri-PT, peri-CV and perisinusoidal regions. To evaluate performance for assisted and unassisted reads, three pathologists with 5 to 40 years' experience interpreted images in 2 sessions: a) Unassisted versus Assisted level I, b) Unassisted versus Assisted level II. Each session consisted of 4 reads, starting with the unassisted read followed by sample randomization before proceeding to the assisted read. This was repeated after a 3–4-week washout period.

Results: When assisted by the level I AI tool, the concordance rate between pathologists improved to near-perfect agreement, with 0.82 linear weighted kappa, as compared to 0.72 for the unassisted review. Mean overall percentage agreement (PA) between pathologists improved from 89.38% to 92.93% (p=0.032). Mean linear weighted kappa for intra-observer agreement was also higher, achieving 0.91 kappa compared to 0.79 for unassisted reads. When compared with the assistive level I tool, the concordance between pathologists with assistive level II tools showed marginal improvement from 0.82 to 0.84. The overall PA increased slightly from 92.9% to 93.8% (p=0.473).

Conclusion: qFibrosis as an AI-assistive tool can improve inter-pathologist weighted kappa to near-perfect (close to 93%) agreement. Additional AI-based zonal annotations provide negligible improvement in inter-pathologist weighted kappa.

Figure:
Zonal annotations and the impact of AI-assistive tools on intra- and inter-rater agreements.



		Unassisted	Assisted Level I	Assisted Level II
Inter-observer	Mean percentage agreement	89.4%	92.9%	93.8%
	Mean weighted kappa (Linear)	0.72	0.82	0.84
Intra-observer	Mean percentage agreement	92.1%	96.5%	95.63%
	Mean weighted kappa (Linear)	0.79	0.91	0.88

Impact of artificial intelligence-based zonal annotations as an additional assistive tool for the histopathologic review of fibrosis in non-alcoholic steatohepatitis patients

Gwyneth Soon[1], Aileen Wee[2], Wei-Qiang Leow[3], Elaine L. K. Chng[4], Yayun Ren[4], Dean Tai[4], Feng Liu[5], Lai Wei[6], Arun J. Sanyal[7]

[1]Department of Pathology, National University Hospital, Singapore, [2]Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, [3]Department of Anatomical Pathology, Singapore General Hospital, Singapore and Duke-NUS Medical School, Singapore, [4]HistoIndex Pte. Ltd., Singapore, [5]Peking University People's Hospital, Peking University Hepatology Institute, Beijing Key Laboratory of Hepatitis C and Immunotherapy for Liver Diseases, Beijing, China, [6]Hepatopancreatobiliary Center, Beijing Tsinghua Changgung Hospital, Tsinghua University, Beijing, China, [7]Stravitz-Sanyal Institute of Liver Disease and Metabolic Health, Virginia Commonwealth University School of Medicine, Richmond, United States

INTRODUCTION

- Inter-observer variability among pathologists for categorical scores of liver fibrosis in non-alcoholic steatohepatitis (NASH) ranges from fair to moderate weighted kappa.[1,2]
- Artificial intelligence (AI) and advances in digitised whole-slide imaging (WSI) have facilitated the use of AI-assistive tools in pathology to improve histopathologic interpretation.
- We previously demonstrated that second harmonic generation/two-photon excitation fluorescence (SHG/TPEF) microscopy with qFibrosis staging and continuous values as AI- assistive tools can help standardize pathologist assessment, contributing to higher overall intra- and inter-rater agreements.[3]

AIM

- The aim of this exploratory study was to test whether the provision of additional AI-based zonal annotations can further improve the overall inter-pathologist agreement with regards to fibrosis assessment in patients with NASH.

METHODS

- Liver biopsies from 40 untreated NASH patients (F0-F4) were evaluated.
- Stained Hematoxylin and eosin (H&E) and Masson trichrome (MT) sections were digitized and images uploaded to a digital viewing platform (Aperio eSlide Manager).
- SHG/TPEF microscopy was performed on unstained sections:
  - Fibrosis was quantitated to provide a qFibrosis stage and continuous value[4]
  - AI-generated zonal annotations were generated to highlight the portal tract (PT), central vein (CV), peri-PT, peri-CV and perisinusoidal regions
- To evaluate performance for assisted and unassisted reads, three pathologists with 5 to 40 years' experience interpreted images in 3 modalities over 2 sessions (Figure 1).
- Each session consisted of 4 reads, starting with the unassisted read followed by sample randomization before proceeding to the assisted read. This was repeated after a 3–4-week washout period.

RESULTS

Figure 1. Case example demonstrating level I and level II assistive tools to aid pathologists in fibrosis evaluation, alongside with conventional stained digital images

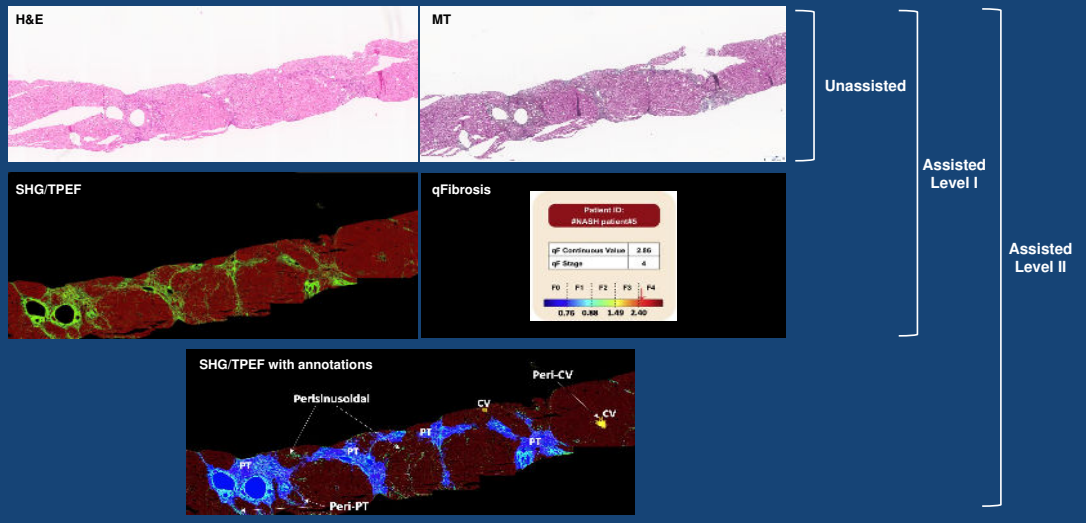


Figure 1: Study pathologists assessed fibrosis in 3 modalities: a) Unassisted read which consisted of only digitized H&E and MT-stained images; b) Assisted level I read which consisted of additional tools of SHG/TPEF image with corresponding qFibrosis readout alongside digitized stained images; c) Assisted level II read which consisted of additional tools of SHG/TPEF image, SHG/TPEF image with zonal annotations and corresponding qFibrosis readout alongside digitized stained images.

Table 1. Inter- and intra-observer variability as measured by mean percentage agreement and weighted kappa (linear) for assisted level I and II versus unassisted reads.

		Unassisted	Assisted Level I	Assisted Level II
Inter-observer	Mean percentage agreement	89.4%	92.9%	93.8%
	Mean weighted kappa (Linear)	0.72	0.82	0.84
Intra-observer	Mean percentage agreement	92.1%	96.5%	95.6%
	Mean weighted kappa (Linear)	0.79	0.91	0.88

- When assisted by the level I AI tool, the concordance rate between pathologists (inter-observer) improved to near-perfect agreement, with 0.82 linear weighted kappa, as compared to 0.72 for the unassisted review.
- Mean overall percentage agreement (PA) between pathologists improved from 89.38% to 92.93% (p=0.032).
- Mean linear weighted kappa for intra-observer agreement was also higher, achieving 0.91 kappa compared to 0.79 for unassisted reads.
- When compared with the assistive level I tool, the concordance between pathologists with assistive level II tools showed marginal improvement from 0.82 to 0.84. The overall PA increased only slightly from 92.9% to 93.8% (p=0.473).

CONCLUSIONS

- qFibrosis as an AI-assistive tool can improve inter-pathologist weighted kappa to near-perfect (close to 93%) agreement, as previously reported.
- Additional AI-based zonal annotations provide negligible improvement in inter-pathologist weighted kappa.
- With AI-assistive tools, such as qFibrosis, pathologic assessment can be further standardized and will remain an important element in determining subject eligibility and assessing treatment effects in NASH clinical trials.
- Pathologists' assessment of fibrosis remains an important element and its usability will only be further enhanced by AI-assistance.
- A larger validation study is planned with inclusion of an adjudication panel to establish "ground truth" for fibrosis to help determine the robustness of qFibrosis as an assistive tool in assessing fibrosis in NASH.

REFERENCES

1. Davison BA, et al. J Hepatol 2020;73:1322–1332.
2. Kleiner DE, et al. Hepatology 2005;41:1313–1321.
3. Soon GST, et al. Clin Gastroenterol Hepatol.2022;10:S1542-3565(22)00555-9.
4. Liu F, et al. Hepatology 2020;71:1953–1966

CONTACT INFORMATION

gwyneth_st_soon@nuhs.edu.sg

EASL CONGRESS Vienna, Austria 21–24 June 2023
DOI: 10.3252/poe.eu.EASLCONGRESS.2023
Metabolism, alcohol and toxicity
Gwyneth Soon
WED–436
EASL2023