# The Complexity of Ballooned Hepatocyte (BH) Identification: Time to Rethink Trial Endpoints for Nonalcoholic Steatohepatitis?

*Elizabeth M Brunt[1], Andrew Clouston[2], Zachary Goodman[3], Cynthia Guy[4], David E. Kleiner[5], Karolin Lackner[6], Dina G. Tiniakos[7], Aileen Wee[8], Matthew Yeh[9], Wei Qiang Leow[10], Arun Sanyal[11], Brent Tetri[12], Mary Rinella[13], Vincent Wong[14], George Goh Boon Bee[15], Elizabeth Powell[16], Lai Wei[17], Zobair Younossi[18], Michael Charlton[19], Vlad Ratziu[20], Stephen Harrison[21], Elaine Chng[22], Yayung Ren[22], *Dean Tai[22] & *Quentin M Anstee[23]

[1]Dept of Pathology and Immunology, Washington Univ, St. Louis, MO, USA; [2]Envoi Specialist Pathologists, Queensland, Australia; School of Biomedical Sciences; Univ of Queensland, Queensland, Australia; [3]Pathology Dept, and Center for Liver Diseases, Inova Fairfax Hosp, Falls Church, VA, USA; [4]Division of Pathology, Duke Univ Medical Center, Durham, NC, USA; [5]Laboratory of Pathology; Center for Cancer Research, NCI, MD, USA; [6]Institute of Pathology, Medical Univ of Graz, Graz, Austria; [7]Institute of Cellular Medicine, Newcastle Univ, Newcastle Upon Tyne, UK; [8]Dept of Pathology, Yong Loo Lin School of Medicine, National Univ of Singapore, Singapore; [9]Dept of Pathology, Univ of Washington, Seattle, WA, USA; [10]Dept of Anatomical Pathology, Singapore General Hosp, Singapore & Duke-NUS Medical School, Singapore; [11]Dept of Internal Medicine, Virginia Commonwealth Univ, Richmond, VA, USA; [12]School of Medicine, Dept of Internal Medicine-Gastroenterology, Saint Louis Univ, St Louis, MO, USA;[13]Dept of Gastroenterology and Hepatology, Feinberg School of Medicine, Northwestern Univ, Chicago, IL, USA; [14]Dept of Medicine and Therapeutics, Chinese Univ of Hong Kong, Hong Kong;[15]Dept of Gastroenterology and Hepatology, Singapore General Hosp, Singapore;[16]Univ of Queensland & Princess Alexandra Hosp, Brisbane, Australia; [17]Beijing Tsinghau Changgung Hosp, Beijing, China; [18]Betty and Guy Beatty Center for Integrated Research, Inova Health System, Falls Church, VA, USA; [19]Center for Liver Diseases and Transplantation Institute, Univ of Chicago, Chicago, IL, USA; [20]Sorbonne Univ and Pitie-Salpetriere Hosp, Paris, France; [21]Pinnacle Clinical Research, San Antonio, TX, USA; Radcliffe Dept of Medicine, Univ of Oxford, UK; [22]Histoindex Pte Ltd, Singapore; [23]Translational & Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.

AASLD Nov. 12-15, 2021
The Liver Meeting®
DIGITAL EXPERIENCE

## INTRODUCTION

- Hepatocyte ballooning is a key feature discriminating nonalcoholic steatohepatitis (NASH) from steatosis (NAFL).
- High inter/intra-observer variation in ballooning measured has been reported.
- Reliable identification of ballooning is crucial for patient enrollment and drug efficacy evaluation.
- There is a pressing need for reproducible, objective and standardized evaluation of hepatocyte ballooning.
- Artificial intelligence (AI)-based approach may provide a more reliable way to assess the range of injury recorded as “hepatocyte ballooning” as a clinical trial endpoint by expert hepatopathologists.

## MATERIALS AND METHODS

- Liver biopsies were obtained from the Seladelpar (NCT03551522) and Resmetirom phase 2 trials (NCT02912260).
- Digitized NAFLD H&E slide images were independently reviewed by 9 expert liver pathologists on two separate occasions
- Each pathologist marked every ballooned hepatocyte (Phase 1) and later provided an overall NAFL/NASH assessment (Phase 2).
- Inter-observer variation was assessed, and a consensus atlas of ballooned hepatocytes was used to train second harmonic generation/two-photon excitation fluorescence (SHG/TPE) imaging-based AI to detect ballooning.

## RESULTS

**Figure 1.** (Left) Heatmap showing number of ballooned cells observed (Dark blue is 0 ballooned cells, through to red) in comparison to ‘non-NASH NAFL’ vs. ‘NASH’ diagnostic call by slide/ pathologist. (Right) Semi-Quantitative Ballooning Score (SQBS) (0-2) by slide and pathologist. (SQBS Ballooning 0: <5 cells circled; 1: 5-75; 2: > 75).

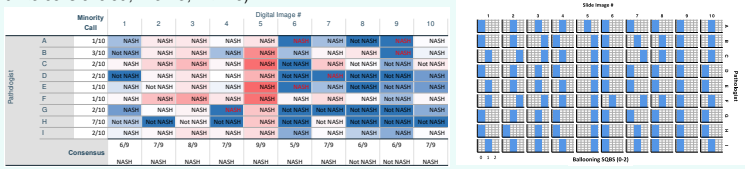


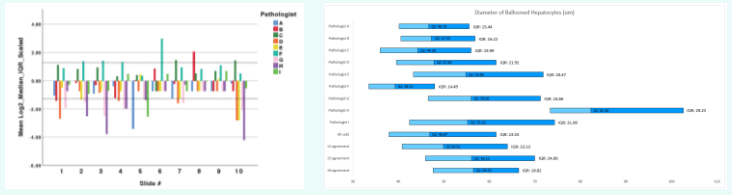
Figure 1 (left):

- Text in red denotes a NASH diagnosis call by a pathologist despite previously reporting no ballooned hepatocytes present at Phase 1.
- No cases for which all pathologists agreed that NASH was absent.
- Kappa value of 0.127 (95%CI 0.024-0.230, P=0.016) for agreement of a NASH diagnosis

Figure 1 (right):

- Using SQBS categories, the level of inter-observer agreement between pathologists remained ‘fair’ (kappa 0.291, 95%CI 0.210-0.371, p<0.0005).

**Figure 2.** (Left) Scaled Count of Cells Circled by Slide and Pathologist. (Right) Ballooned hepatocyte diameter by pathologist.



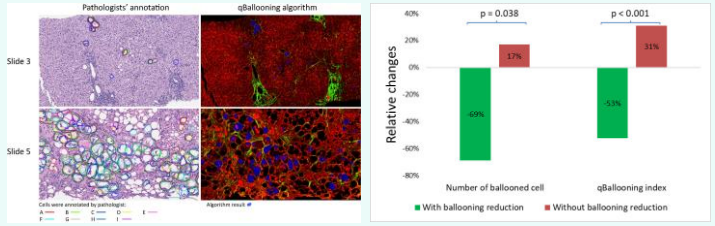
- Individual pathologists consistently tended to identify greater or lesser numbers of ballooned cells.
- Pathologists who considered more cells to be ballooned adopted a more permissive, lower, cell-diameter threshold (pathologists F 39.31 ± 14.49µm and C 33.28 ± 19.99µm) vs. those who identified the least cells to be ballooned (pathologist H 82.30 ± 29.23µm), p<0.001.

**Table 1.** Use of the Histological ‘Ground Truth’ Atlas to tune the qBallooning2 Algorithm. **Relative to majority consensus of ≥5-pathologists. § Based on an estimated mean 8,150 hepatocytes per digital image from nuclear counting and shown for completeness.*

| qBallooning2 training-set cell-selection criteria | Number of ballooned cells identified by Pathologists | Number of ballooned cells identified by qBallooning2 | Overlap between qBallooning2 and majority consensus of ≥5-Pathologists | Positive Predictive Value Proportion of ballooned cells called by qBallooning2 are 'True Positive' * | False Discovery Rate Proportion of ballooned cells called by qBallooning2 are 'False Positive' * | True Positive Rate (Sensitivity) Proportion of ballooned cells identified by qBallooning2 * | False Negative Rate Proportion of ballooned cells missed by qBallooning2 * | Estimated True Negative Rate (Specificity) *§ |
|---|---|---|---|---|---|---|---|---|
| Agreement of any 1 pathologist | 1188 | 346 | 54 | 54/346 (16%) | 292/346 (84%) | 54/133 (41%) | 79/133 (59%) | >99% § |
| Agreement of any 2 pathologists | 481 | 250 | 51 | 51/250 (20%) | 199/250 (79.6%) | 51/133 (38%) | 82/133 (62%) | >99% § |
| Agreement of any 3 pathologists | 284 | 170 | 37 | 37/170 (22%) | 133/170 (78.2%) | 37/133 (28%) | 96/133 (72%) | >99% § |
| Agreement of any 4 pathologists | 188 | 114 | 25 | 25/114 (22%) | 89/114 (78%) | 25/133 (19%) | 108/133 (81%) | >99% § |
| Agreement of any 5 pathologists | 133 | 88 | 22 | 22/88 (25%) | 66/88 (75%) | 22/133 (17%) | 111/133 (83%) | >99% § |
| Agreement of any 6 pathologists | 86 | 59 | 16 | 16/59 (27%) | 43/59 (73%) | 16/133 (12%) | 117/133 (88%) | >99% § |
| Agreement of any 7 pathologists | 59 | 40 | 15 | 15/40 (38%) | 25/40 (62.5%) | 15/133 (11%) | 118/133 (89%) | >99% § |
| Agreement of any 8+ pathologists | 26 | 24 | 5 | 5/24 (21%) | 19/24 (79%) | 5/133 (4%) | 128/133 (96%) | >99% § |

- The SHG/TPE “qBallooning2” algorithm was optimized to detect the ballooned cells for each level of inter-observer concordance
- Comparing the qBallooning2 index to cells identified by expert pathologist consensus as a reference standard, hepatocyte ballooning could be detected with 17-73% specificity and 5-57% sensitivity
- Based on the performance analysis shown in Table 1, the “qBallooning2” algorithm that had been optimized based on concordance of ≥5-pathologists was selected for further study.

**Figure 3.** (Left) Two examples showing the BH identification by pathologists and qBallooning index. (Right) The AI algorithm reading from a separate drug trial, showing agreement with the study pathologist’s interpretation




- Pilot data demonstrates that qBallooning index has the capacity to detect change in ballooning
- qBallooning performance may be improved by further refinement and validation will be required before implementation.

## CONCLUSIONS

- Substantial divergence in BH identification amongst expert liver pathologists suggests ballooning is a spectrum, too subjective for its presence or absence to be unequivocally determined.
- By digitized slide evaluation, 9 expert liver pathologists had poor agreement on the exact numbers of BH per slide.
- By separate, blinded evaluation, diagnostic categorization of NAFL vs NASH likewise did not always correlate with presence of BH in the slides.
- An AI algorithm, based on consensus cells of ≥ 5 pathologists was created and successfully quantified change in BH in a separate clinical trial.
- BH result from complex alterations to hepatocytes; and variable appreciation by pathologists, including cell size.
- A consensus atlas of cells may be used to train AI for assessment of efficacy of therapy as well as understanding of etiology.
- A second phase of this work is planned for consensus amongst pathologists.

## ACKNOWLEDGEMENTS

This study has been supported by Histoindex Pte Ltd ; the Newcastle NIHR Biomedical Research Centre; the Intramural Research Program of the NIH, National Cancer Institute; and the LITMUS (Liver Investigation: Testing Marker Utility in Steatohepatitis) consortium funded by the Innovative Medicines Initiative (IMI2) Program of the European Union under Grant Agreement 777377; this Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and EFPIA.

## CONTACT INFORMATION

**Joint senior and corresponding authors:**
Ebrunt@wustl.edu
Quentin.Anstee@newcastle.ac.uk
Dean.Tai@histoindex.com