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

Louis PETITJEAN1, Li CHEN1, Arras N. MATTIS2, Mojgan HOSSEINIR, Michael GOEDKEN4, Cynthia BEHLING5, Arun J. SANYAL6, Mathieu PETITJEAN1

1 Pharmasen, Princeton, N.J. USA  2 Department of Pathology, University of California San Francisco, CA, USA  3 Department of Pathology, University of California San Diego, CA, USA  4 Rutgers Research Pathology Services, New Brunswick, NJ, USA  5 Department of Pediatrics University of California, San Diego, CA  6 Virginia Commonwealth University, Richmond, VA, USA

1 Introduction

Manual histological evaluation of liver biopsy is the gold standard for fibrosis and steatosis staging in Non-Alcoholic Steatohepatitis (NASH), but it is limited by its inter and intra-reader variability. Quantitative Digital Pathology image analysis and AI (FibroNest™) methods have the potential to overcome the limitation of these standards.

2 Aim

We have previously shown that the Phenotypic Fibrosis Composite Score calculated by the FibroNest methods correlate with the NASH-CRN histological fibrosis stages and steatosis grades established from collagen-stained histology slides. Here, we report the performance of the FibroNest method to quantify Ballooning, Inflammation and Steatosis from digital images of H&E human liver biopsy sections.

3 Method

1. Image Dataset:
   • Retrospective cohort of 85 patient with NASH diagnosed by histologic assessment of liver biopsy according to NASH-CRN.
   • 20X digital images of H&E stained FFPE sections of liver biopsies.

2. Machine Learning Training:
   • Cohort of 21 selected images
   • Annotations by four expert pathologists:
     - Positive features
     - False or "Look-Alike" features
     - Healthy Tissue landmarks
     - "Noise" annotation / point / feature

3. Predictive ML Model:
   • Deep-learning based model
   • Composed of anatomical feature maps for lobular, portal inflammation and hepatocellular ballooning
   • Macro-vascular steatosis feature confidence

4. Composite scores calculation:
   • The selection of a model probity cut off defines real-estate "objects" that can be quantified with count, density, morphology at the tissue level or for 200X equivalent FOV (total 64 components). Principal components are combined into composite scores.

5. Validation Cohort:
   • Training cohort is re-incorporated in the Validation Cohort.
   • Validation cohort is process and the agreement of the composite score with NASH-CRN grades is evaluated.
   • Iterations (steps 1 to 5) are performed adjust and refine the predictive model and the performance of the scores (not reported here).

6. Contact Information

Louis.Petitjean@pharmasen.com

4 Steatosis Results

- Effective exclusion of microvessel steatosis and glycogenotic hepatocytes.
- Model translates to Pre-clinical tissues
- Grade 1 vs 2 confusion is the combination of ML accuracy and Pathologist's annotation.
- This approach does not meet FDA outcome definitions ("either percent of steatosis hepatocytes or non-fibrotic tissue fat area ratio")
- New FibroNest release resolves these issues and results and results higher performance (see EASL2020 poster # FR0038)

3 Nuclear Inflammation Results

- ML Model accuracy (F1) for Lobular foci is 55%.
- Further improvements are possible if the histological definition of a "Lobular foci" is improved.
- The semi-automatic and systematic quantification method is attractive as an investigational endpoint in NASH studies.
- The performance of the method is reduced for significantly hematoxylin overlaid tissues.
- Quantitative Image Analysis ("single Nuclei" analysis, FibroNest V3.1) resolves these issues.

2 Hepatocellular Ballooning Results

- Significant disagreement between pathologists' annotations as reported elsewhere due to a poor definition of "ballooned hepatocytes" of "definite" vs "diagnostically borderline" and (b) accounting methods "none / few / many / ballooned" hepatocytes or non-pathological ballooning.
- The ML Model accuracy (F1) for Ballooned Hepatocytes is 37%.
- Thresholding the "ballooned Hepatocyte" topographic probability maps at high probability (>65%) enables the detection of a Density score that correlates poorly with Pathology grades.
- Further improvements are possible if the histological definition of "ballooning" is improved.

Digital Pathology AI methods based on supervised ML and annotations provided by pathologists yield to moderate performance results.