Digital Pathology quantitative image analysis and AI method detects the treatment effect of pegbefelmin in Cirrhosis patients with a performance that benchmarks manual histological assessment

Li Chen1, Anne Minnich2, Edgar D. Charles3, Zachary Goodman4, Mathieu Petitjean5, Arun J. Sanyal6
1 Pharmasset Inc, Princeton, NJ, USA, 2 Bristol Myers Squibb, Princeton, NJ, USA, 3Inova Health Systems, Falls Church, VA, USA, 4Virginia Commonwealth University, Richmond, VA, USA

Background and aims

Manual histological evaluation of liver biopsy is the gold standard for fibrosis staging in Non-Alcoholic Steatohepatitis (NASH), but it is limited by its inter- and intra-reader variability. Digital Pathology image analysis (FibroNest™) has the potential to overcome the current limitation of such standards. This exploratory post-hoc analysis compared FibroNest content scores with NASH-CRN categorical stages in patients with NASH from the phase 2b FALCON study (NCT03486912).

Method

Study design and treatment arms

Eligible adults were 18-75 years of age (N=145) with NASH diagnosed by histologic assessment of liver biopsy according to NAS-H CRN criteria and stage 4 fibrosis, defined as Cirrhosis. During the 48-week double-blind comparison period, patients received 10mg, 20mg, or 40mg pegbefelmin subcutaneous or placebo once weekly.

Liver tissue histology

- Liver biopsies were obtained six months before or during screening and at week 48.
- Formalin-fixed, paraffin embedded sections (~4 microns) of adequate liver biopsies were stained with Masson Trichrome for Collagen.

Whole slide imaging

- The same slides prepared for and reviewed by the pathologist were digitized at 40x (0.25 microns/pixel) on an Aperio AT2 WSI system. A total of 276 images were acquired.
- Each Digital Image was evaluated for adequacy, using 20 criteria specific to Biopsy tissue (quality, size, histology Tissue processing (embedding, sectioning, mounting), Trichrome staining (uniformity, significant degrees of intensity) and scanning (white balance correction, stripes, dust), all resulting into a Digital Biopsy adequate Score (DBA) ranging from 0 to 10. Digital images that are not acceptable has DBA<5, and minimally acceptable is 5-50.

Two assessments of the antifibrotic effect were studied on the same slides:

NASH-CRN and ISHAK staging for fibrosis severity

- Biopsy slides are read by an independent pathologist trained on the NASH-CRN and ISHAK staging system.
- NAS-H CRN F3 (severe fibrosis) and F4 (cirrhosis) and ISHAK (Stages 4, 5, or 6) are used to stage patients at baseline and end of treatment.
- Responders are identified with at least a full categorical stage change.

Digital pathology and artificial intelligence

- The digital images were read using FibroNest™, a single-fiber, high-content, quantitative Digital Pathology image analysis and AI method. This provides a continuous phenotypic Fibrosis Composite Severity Score (Ph-FCS, FibroNest Digital Pathology biomarker for fibrosis) that ranges from 1 to 10 for the full spectrum of fibrosis severity observed in the liver. This allows identifying fibrosis improvements that may be missed by stage definitions and other statistical quantification of change from baseline. An exploratory biomarker PT-Ph-FCS was also normalized and calculated based on the Parenchymal (e.g. non-destructive) Tissue topology.
- A 0.3 absolute reduction in Ph-FCS (4-fold higher than the analytical variability) identified any reduction in fibrosis [5].
- A 25% relative reduction of Ph-FCS (corresponding to an absolute change of 0.75 to 2 for 3xPh-FCS<3) is used to identify significant reduction in responders [4].

Single-fiber, high-content, quantitative image analysis

Using Quantitative Image Analysis (FibroNest™) the fibrosis phenotype is described for its collagen content and structure (12 traits), the morphometric traits of the collagen fibers (13 traits), and fibrosis architecture traits. In each image, each morphometric and texture trait is represented by a histogram distribution.

- The histogram for each trait is described by seven quantitative fibrosis parameters (αPFs, 315 in total) to account for mean, variance, distortion and progression.
- Principal αPFs (1,23) are automatically detected to account for disease severity if their group mean value difference is statistically (P<0.05, T-Test) greater than 20%.
- Principal αPFs are combined into a normalized Phenotypic Composite Fibrosis Severity score (Ph-FCS), a continuous quantifier of the fibrosis severity phenotype.

Digital pathology method

Augmented digital pathology images

- Ph-FCS correlates with NASH CRN F3 and F4 stages (Fig. A). The performance of the FCS as a diagnostic biomarker to classify F3 from F4 is moderate (AUROC=0.753) but is on-par with histological methods and Non-Invasive Tests.

Augmented digital pathology images

- F3-F4 Diagnostic performance of the Ph-FCS

References


Conclusion

Quantitative digital pathology image analysis and AI generates continuous scores for fibrosis that enhance conventional histological staging and resolve the continuum of cirrhosis. The definition of meaningful change criteria using this continuous scoring remains to be improved.

Contact and presenting author: mathieu.petitjean@pharmasset.com

Results

- Groups sizes with paired biopsies were 22, 24, 20, 27 for the placebo, 10mg, 20mg, 40mg groups following removal of images considered non-evaluable for FibroNest algorithms (i.e. DBA < 5).
- Responders were identified with a 1-unit reduction for the histological stage (Fig. D).
- Using an absolute reduction of 0.3 (4-fold higher than the analytical variability), the Ph-FCS resolved 15% to 20% (resp. 0% to 10%) more responders than NASH CRN (resp. Ishak) categorical stages which is consistent with an increased detection threshold (Fig. D).
- A 25% relative reduction of Ph-FCS (corresponding to an absolute change of 0.75 to 2 for 3xPh-FCS<3) detected fewer responders than when using NASH-CRN or Ishak (Fig. B).
- While the Ph-FCS absolute reduction cut off of 0.3 and relative reduction cut off at 25% are consistent with the one defined in the context of pre-cirrhotic patient studies [5], it might be appropriate to change them in the context of a study in cirrhotic patient to adjust the amount of change to the severity of the disease. The definition of a “meaningful Ph-FCS change” is still under development and will probably be settled with the use of retrospective cohorts with long term follow up and liver related events.
- There was no difference between the Ph-FCS and the PT-Ph-FCS which is attributed to the lack of antifibrotic effect in the treatment study, as reported elsewhere.