Non-invasive Testing for Fatty Liver Disease for Primary Care Providers

Auguste Byiringiro, Marc Letournieur, Fibronastics Pte Ltd and Hawley Linke, Fibronastics US, Inc.

ABSTRACT

Population data point to increasing prevalence of NAFLD worldwide. Emerging intervention strategies command development of more accessible diagnostic tools to identify individuals at early risk for morbidities associated with untreated NAFLD. While liver biopsies are the diagnostic standard, patients & their clinicians need strong motivators to invoke the risk and uncertainty of biopsy; or even the expense and inconvenience of imperfect imaging technologies. We describe an inviting pathway to reliable diagnoses in a 2-stage process of risk assessment: 1) The initial screening tool, LiverFASST Select, delivers a binary prediction of “Elevated” or “Low” Risk for steatosis; available to virtually any patient using an annual wellness or preventive visit. “Elevated Risk” prediction justifies acquisition of a few strategy biomarkers for non-invasive quantitative risk assessment of steatosis, inflammation, & fibrosis. 2) Using up to 9 biomarkers, the full LiverFASST algorithm predicts the likely degree of liver pathology, providing SAP scores: Steatosis S1-S3 Activity/Inflammation A0 - AA, Fibrosis F0 - F4. The algorithm is demonstrably close to biopsy predictions and can provide the critical diagnoses that motivate patient and physician to develop and implement the best available intervention strategies.

METHOD

- Interview IRB of Austin, TX approved the protocol to assess de-identified medical records containing multiple biomarkers and pathologists-determined SAP scores derived from liver biopsies.
- A database of 2882 unique medical assessments of biomarkers & biopsy reports was created. 1027 assessments were used to train the algorithm. 1835 constituted the validation set.
- KR developed the complex quantitative algorithm utilizing 3 anthropometrics: age, gender BMI. It is up to 9 biomarkers to accurately predict level of steatosis, inflammation activity and fibrosis (comparable to biopsy SAP score).
- Previously, Assistance Publique (AP)-HP compiled 3 sets of markers to create algorithms to assess severity of the 3 stages of NAFLD biopsy-demonstrated pathologies: age & gender plus 4 biomarkers for fibrosis; 1 additional biomarker for inflammation activity; and another 4 biomarkers for steatosis. For the creation of LiverFASST, three neural networks (1 each for S, A, and F) were developed and aligned against the AP-HP determinations for accuracy relative to biopsy.
- Subsequently a LiverFASST Select algorithm was trained from 1670 medical records to make a binary decision for the (ELEVATED/LOW) probability of NAFLD/ NASP based on age, gender, BMI & 1 or more biomarkers out of 6 frequent tests. Patients predicted at ELEVATED risk require follow-up quantitative diagnostic tests.

The ML algorithm created new SAP scoring using the markers below. LiverFASST uses one less biomarker than the AP-HP algorithm.

Fibrosis: Age, Gender, α 2 Macroglubulin, Apolipoprotein A1, Bilirubin, GGT, Hepatocellular
Inflammation: Age, Gender, α 2 Macroglubulin, Apolipoprotein A1, Bilirubin, GGT, Hepatocellular
Steatosis: Age, Gender, α 2 Macroglubulin, Apolipoprotein A1, Bilirubin, GGT, Hepatocellular, ALT, Total Cholesterol, Fasting Glucose, Triglycerides

BACKGROUND

Liver characterization: 1) STEATOSIS - A liver biopsy to define the degree of steatosis based on the percentage of hepatocytes containing fat, or macrovesicular steatosis; 2) INFLAMMATION - A liver biopsy to define the degree of inflammation. Pathologists assign a histological grade of inflammation to each patient; 3) FIBROSIS – A liver biopsy to define the degree of fibrosis. Pathologists score fibrosis usually by assigning a stage of fibrosis - from S0 to S3.

TRAINING SET: The training dataset (n=1027) included 60% male, was slightly older (mean age =51 yrs) and with a lower BMI. The validation dataset (n=1835) was 17% male, mean age = 45 yrs, and higher BMI.

Following algorithmic assignment of disease diagnoses, 235 medical assessments yielded simultaneously negative diagnostic predictions for steatosis, inflammation activity, and fibrosis. These “SAFOP” patients were therefore considered to have a functionally “healthy liver.” The training dataset and the validation dataset included 192 and 43 healthy liver assessments respectively.

Even when less than the optimal number of blood biomarkers are available for the LiverFASST Select algorithm (E.g.: only FBS and trig) the algorithm provides useful information to inspire physician and/or patients to investigate without causing unnecessary apprehension. Adding the readily available biomarkers ALT and GGT provide good sensitivity and specificity.

TABLE 1 RESULTS OF STAGE 1 FOR BINARY (ELEVATED/LOW) RISK ASSESSMENT

Statistical values related to LiverFASST Select Risk Assessment based on increasing application of specific, selected biomarkers. (See Methods)

<table>
<thead>
<tr>
<th></th>
<th>SAP</th>
<th>A</th>
<th>STD</th>
<th>MAE</th>
<th>RFA</th>
<th>AUC</th>
<th>PR</th>
<th>EM</th>
<th>TNR</th>
<th>FPR</th>
<th>LP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td>87.6</td>
<td>1.2</td>
<td>0.89</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td>87.6</td>
<td>1.2</td>
<td>0.89</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FBS</td>
<td></td>
<td>87.6</td>
<td>1.2</td>
<td>0.89</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trig</td>
<td></td>
<td>88.7</td>
<td>1.1</td>
<td>0.83</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2 CLINICAL CHARACTERISTICS OF INDIVIDUAL ASSESSMENTS IN THE DATASETS FOR THE QUANTITATIVE ALGORITHM.

The training dataset (n=1027) included 60% male, was slightly older (mean age =51 yrs) and with a lower BMI. The validation dataset (n=1835) was 17% male, mean age = 45 yrs, and higher BMI.

Following algorithmic assignment of disease diagnoses, 235 medical assessments yielded simultaneously negative diagnostic predictions for steatosis, inflammation activity, and fibrosis. These “SAFOP” patients were therefore considered to have a functionally “healthy liver.” The training dataset and the validation dataset included 192 and 43 healthy liver assessments respectively.

Even when less than the optimal number of blood biomarkers are available for the LiverFASST Select algorithm (E.g.: only FBS and trig) the algorithm provides useful information to inspire physician and/or patients to investigate without causing unnecessary apprehension. Adding the readily available biomarkers ALT and GGT provide good sensitivity and specificity.

TABLE 3 COMPARISON TO THE AP-HP TEST ACCURACY

Statistical metrics used to describe the accuracy of regression models were computed to assess the performance of the new models. MACE, MaxAE and R2 (Coefficient of Determination) for the prediction of interest: ActTest and SteatoTest biopsy-validated scores with the new SAP prediction models. On average, the new models make predictions that are very close to the AP-HP derived scores.

Table 3: For the three new models, the order of magnitude of the MAE is 1E-3. This is satisfactory since AP-HP-derived scores are given with a precision of 1E-2.

TABLE 4 CORRELATION BETWEEN THE ESTIMATED SAP SCORE AND THE PREDICTED NAFLD/MASH DIAGNOSIS

The algorithm retrieves 3 separate scores to create the composite LiverFASST SAP score - SAkyta – to determine probable outcome in FUP biopsy scoring. (X, Y, and Z are integers from [0-3], [0-4] and [0-4] respectively).

Diagnoses are unambiguous except for SA0-A2 Fx. When the steatosis stage is at least 1 and the activity stage is 2, the patient is more likely to be labelled NAFLD (probability of 0.99999) or NASH (probability of 0.99997). In this specific case the LiverFASST algorithm is only able to provide an open diagnosis: NAFLD or NASH.

DISCLOSURE

Fibronastics provided the financial support for this study.

CONTACT: ronald.quiambao@fibronotics.com

FIGURE 1 DISTRIBUTION OF THE INDIVIDUAL ASSESSMENT SCORES TEND TO REPRESENT EARLIER STAGES OF FIBROSIS AND INFLAMMATION ACTIVITY

Unlike some other non-biopsy diagnostic tools, the majority of the assessments were in early stages of fibrosis and inflammation activity: steatosis assessments, however, were more balanced across the range of scores.

HISTOGRAMS OF: (left to right) Steatosis, Inflammation Activity, B Fibrosis scores (Training and validation datasets are similar; only the training set shown)