Simulating clinical confidence intervals for black-box algorithmic predictions of liver steatosis

Brian A. Cohn[1], Mona C. Munteanu[1,2], Victor de Ledinghen[3,4], Rifaat Safadi[5], Chavdar Pavlov[6,7], Teresa Gonzalo[1], Roni Amiel[1], Ronald Quiambao[1]

[1]Liver Research Center, Fibronostics, Orlando, US.
[2]Université Paris-Descartes (Paris V), Paris, France.
[3]Hepatology Unit, Centre Hospitalier Universitaire (CHU), Bordeaux, France.
[4]INSERM U1053, Bordeaux University, Bordeaux, France.
[5]Liver Unit, Institute of Gastroenterology and Liver Diseases, Department of Medicine, Hadassah-Hebrew University Medical Center, Jerusalem, Israel.
[6]'I.M.Sechenov' First Moscow State Medical University, Department of Therapy, Center for Evidence-Based Medicine, Moscow, Russian Federation.

Background and aim

Clinicians have begun using blood-serum biomarkers with artificial intelligence algorithms (AIAs) to assess the degree of liver steatosis without taking a liver biopsy. However, intra-patient and intra-lab variability could affect the inputs, and with >5 biomarkers used by an AIA, noise is compounded. Interpretable measures of an AIA's *confidence* are absent in the clinical workflow. We aim to resolve this gap in interpretability of non-invasive AIA with a stochastic noise injection method and interactive data visualization—allowing clinicians to a) observe steatosis predictions under simulated noise conditions and b) interactively simulate expected regression of steatosis with respect to changes in biomarkers through course of treatment.

Methods

LIVERFASt™ Steatosis, an AIA proprietary technology (Fibronostics, Orlando US), uses surrogate serum biomarkers to infer steatosis (0-1). Some of the serum biomarkers (e.g. cholesterol) are often targets for therapy. With data from a single use-case patient, we exhaustively injected uniform noise (from ± 1 to 64%) into one biomarker at a time (10,000 replicates per biomarker, cond.).

Results

Increasing noise in e.g. apolipoprotein A1 (the major protein component of high-density lipoprotein (HDL) cholesterol in plasma) shows increasing probability of yielding a false positive S1 grade for steatosis (Fig 1a). Steatosis prediction variability increases linearly with noise injection across all biomarkers ($R > 0.99$; $p < 0.001$, Fig 1b); this behavior is visible in our interactive parallel coordinate visualization, which allows a clinician to interactively change the noise levels on all biomarkers to see the effect on the output steatosis scores (e.g. with BMI in Figure 1c).

Conclusion

Noise injection allows for nuanced tailoring and interpretation of results inferred by an AIA. We equip clinicians with an interactive tool to contextualize therapeutic expectations and aid in patient education towards adherence endpoints. In effect, we provide a new, nuanced view of *how* AI algorithms assess liver disease.

Figure 1. a) Expected results of noise injection into apolipoprotein A1, a biomarker input into the AI-based LIVERFASt™ to produce a steatosis score b) linear noise equations for each biomarker, and c) steatosis with a parallel coordinate visualization.