

# Maximizing the Diagnostic Yield of Non-Invasive Tests: Simultaneous Optimization of Key Performance Metrics for Fibrotic Liver Disease with a FIB-4 and MASEF (OWLiver) Strategy

I. Martínez-Arranz<sup>1</sup>, R. Mayo<sup>1</sup>, M. Noureddin<sup>2,3</sup>, E. Truong<sup>4</sup>, J. M. Banales<sup>5,6</sup>, M. Arrese<sup>7</sup>, K. Cusi<sup>8</sup>, P. Irizubia<sup>9</sup>, R. Bruha<sup>10</sup>, M. Romero-Gómez<sup>11</sup>, R. Aller<sup>12</sup>, J. Ampuero<sup>13</sup>, J. L. Calleja<sup>14</sup>, L. Ibañez-Samaniego<sup>15</sup>, P. Aspichueta<sup>16,17,18</sup>, A. Martín-Duce<sup>19</sup>, T. Kushner<sup>20</sup>, Javier Crespo<sup>9</sup>, A. J. Sanyal<sup>21</sup>, M. Zulueta<sup>22</sup>.

<sup>1</sup> Rubiò Metabolomics, Derio, Spain. <sup>2</sup> Houston Methodist Hospital, Houston Research Institute Houston, Texas, USA. <sup>3</sup> Houston Research Institute, Houston, Texas, USA. <sup>4</sup> Karsh Division of Gastroenterology and Hepatology, Cedars-Sinai Medical Center, Los Angeles, CA. <sup>5</sup> Biodonostia Research Institute, Donostia University Hospital, University of the Basque Country (UPV-EHU), CIBERehd, IKERBASQUE, Donostia, Spain. <sup>6</sup> Department of Biochemistry and Genetics, School of Sciences, University of Navarra, Pamplona, Spain. <sup>7</sup> Department of Gastroenterology, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile. <sup>8</sup> University of Florida, Gainesville, Florida, USA. <sup>9</sup> Marqués de Valdecilla University Hospital, Cantabria University, IDIVAL, Santander, Spain. <sup>10</sup> General University Hospital and the First Faculty of Medicine, Charles University, Prague, Czech Republic. <sup>11</sup> Valme University Hospital, CIBERehd, Seville, Spain. <sup>12</sup> Clinic University Hospital, University of Valladolid, Valladolid, Spain. <sup>13</sup> Virgen del Rocío University Hospital, Sevilla, Spain. <sup>14</sup> Puerta del Hierro University Hospital, Madrid, Spain. <sup>15</sup> Gregorio Marañón University Hospital, Madrid, Spain. <sup>16</sup> Department of Physiology, Faculty of Medicine and Nursing, University of the Basque Country UPV/EHU, Leioa, Spain. <sup>17</sup> Biocruces Bizkaia Health Research Institute, Barakaldo, Spain. <sup>18</sup> National Institute for the Study of Liver and Gastrointestinal Diseases (CIBERehd, Instituto de Salud Carlos III), Madrid, Spain. <sup>19</sup> Príncipe de Asturias University Hospital, Alcalá University, Madrid, Spain. <sup>20</sup> Division of Liver Diseases, Icahn School of Medicine at Mount Sinai, New York City, New York, USA. <sup>21</sup> Virginia Commonwealth University Medical Center, Richmond, Virginia, USA. <sup>22</sup> CIMA Sciences, USA

## BACKGROUND

Non-invasive tests (NITs) for metabolic dysfunction-associated steatotic liver disease (MASLD) diagnosis and staging face hurdles due to performance limits, especially the sensitivity-specificity trade-off. Robust NITs are needed. The low-cost, simple FIB-4 index is a recommended first-line tool for fibrosis risk. The Metabolomics-Advanced Steatohepatitis Fibrosis Score (MASEF) (part of the OWLiver test) accurately predicts at-risk MASH (MASH with significant fibrosis).

## AIM

To evaluate a novel FIB-4 + MASEF strategy to maximize accuracy, sensitivity, specificity, PPV, and NPV for predicting at-risk MASH, aiming to overcome current limitations of NITs.

## METHOD

This multicenter study analyzed a cohort of 1122 patients at elevated risk for MASLD due to various metabolic comorbidities. FIB-4 and MASEF were determined and compared against the histopathological reference standard. To identify optimal diagnostic thresholds for FIB-4, a grid-based evaluation of all possible FIB-4 cutoff pairs was performed. For each pair, the MASEF score (part of the OWLiver test) was applied exclusively to the indeterminate zone (patients with FIB-4 values between the two cutoffs). This approach allowed for the systematic assessment of the combined strategy's performance across a multidimensional space. A rigorous statistical analysis, including ROC curve modeling and joint maximization of accuracy, sensitivity, specificity, PPV, and NPV, was conducted focusing on identifying a method capable of the simultaneous optimization of all five metrics.

## RESULTS

Fibrosis Test	N	Cut-off	Accuracy	Sensitivity	Specificity	NPV	PPV
MASEF Score	1355	(0.33 - 0.33)	0.71 (0.69-0.74)	0.73 (0.69-0.77)	0.71 (0.68-0.73)	0.86 (0.84-0.89)	0.51 (0.47-0.55)
FIB-4 → NFS	752	(1.3 - 2.67) / (-1.455 - -1.455)	0.66 (0.63-0.69)	0.60 (0.54-0.67)	0.68 (0.65-0.72)	0.82 (0.78-0.85)	0.42 (0.37-0.48)
NFS → FIB-4	752	(-1.455 - -1.455) / (1.3 - 1.3)	0.45 (0.42-0.48)	0.82 (0.77-0.87)	0.31 (0.28-0.35)	0.82 (0.77-0.87)	0.31 (0.28-0.35)
FIB-4 → NFS	752	(1.3 - 1.3) / (-1.455 - -1.455)	0.66 (0.63-0.69)	0.67 (0.60-0.73)	0.66 (0.61-0.70)	0.84 (0.80-0.87)	0.42 (0.37-0.48)
FIB-4 → MASEF Score	1122	(1.3 - 2.67) / (0.33 - 0.33)	0.75 (0.73-0.78)	0.56 (0.50-0.61)	0.83 (0.81-0.86)	0.82 (0.80-0.85)	0.58 (0.52-0.63)
FIB-4 → MASEF Score	1122	(1.3 - 1.3) / (0.33 - 0.33)	0.65 (0.62-0.68)	0.63 (0.58-0.69)	0.66 (0.62-0.69)	0.81 (0.78-0.84)	0.43 (0.39-0.47)
NFS → MASEF Score	754	(-1.455 - -0.676) / (0.33 - 0.33)	0.52 (0.48-0.55)	0.76 (0.71-0.82)	0.42 (0.38-0.47)	0.82 (0.78-0.87)	0.33 (0.29-0.38)
NFS → MASEF Score	754	(-1.455 - -1.455) / (0.33 - 0.33)	0.45 (0.42-0.49)	0.82 (0.76-0.87)	0.31 (0.27-0.35)	0.82 (0.76-0.87)	0.31 (0.27-0.35)
MASEF Score → VCTE	315	(0.33 - 0.33) / (7.6 - 7.6)	0.64 (0.58-0.69)	0.80 (0.73-0.87)	0.52 (0.45-0.59)	0.79 (0.71-0.85)	0.55 (0.48-0.62)
FIB-4 → VCTE	310	(1.3 - 2.67) / (7.6 - 7.6)	0.65 (0.60-0.70)	0.66 (0.57-0.74)	0.64 (0.57-0.71)	0.72 (0.66-0.79)	0.58 (0.50-0.65)
FIB-4 → VCTE	310	(1.3 - 1.3) / (7.6 - 7.6)	0.63 (0.57-0.68)	0.73 (0.65-0.80)	0.55 (0.48-0.63)	0.73 (0.66-0.80)	0.54 (0.48-0.62)
VCTE → MASEF Score	315	(7.6 - 7.6) / (0.33 - 0.33)	0.54 (0.49-0.60)	0.91 (0.86-0.95)	0.27 (0.21-0.34)	0.81 (0.70-0.89)	0.47 (0.41-0.54)
MASEF Score → FIB-4	1122	(0.33 - 0.33) / (1.3 - 1.3)	0.73 (0.70-0.75)	0.75 (0.71-0.80)	0.71 (0.68-0.74)	0.88 (0.85-0.90)	0.52 (0.47-0.56)

**Table 1. Diagnostic performance of sequential non-invasive tests.**

Each row represents a diagnostic strategy applying one or more non-invasive tests consecutively (e.g., FIB-4 → NFS). N indicates the number of subjects evaluated for each combination. The cut-off column shows the decision thresholds used for each test in the sequence. Accuracy, Sensitivity, Specificity, NPV (Negative Predictive Value) and PPV (Positive Predictive Value) for the detection of at-risk MASH are shown for each combination of NITs.

## CONCLUSIONS

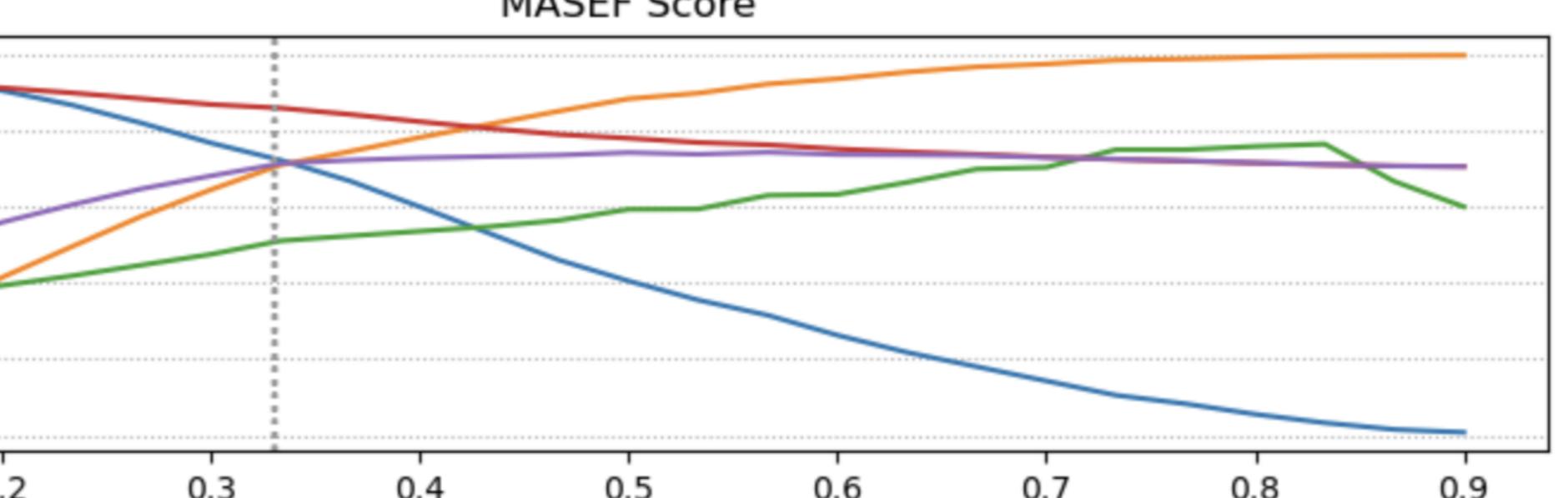
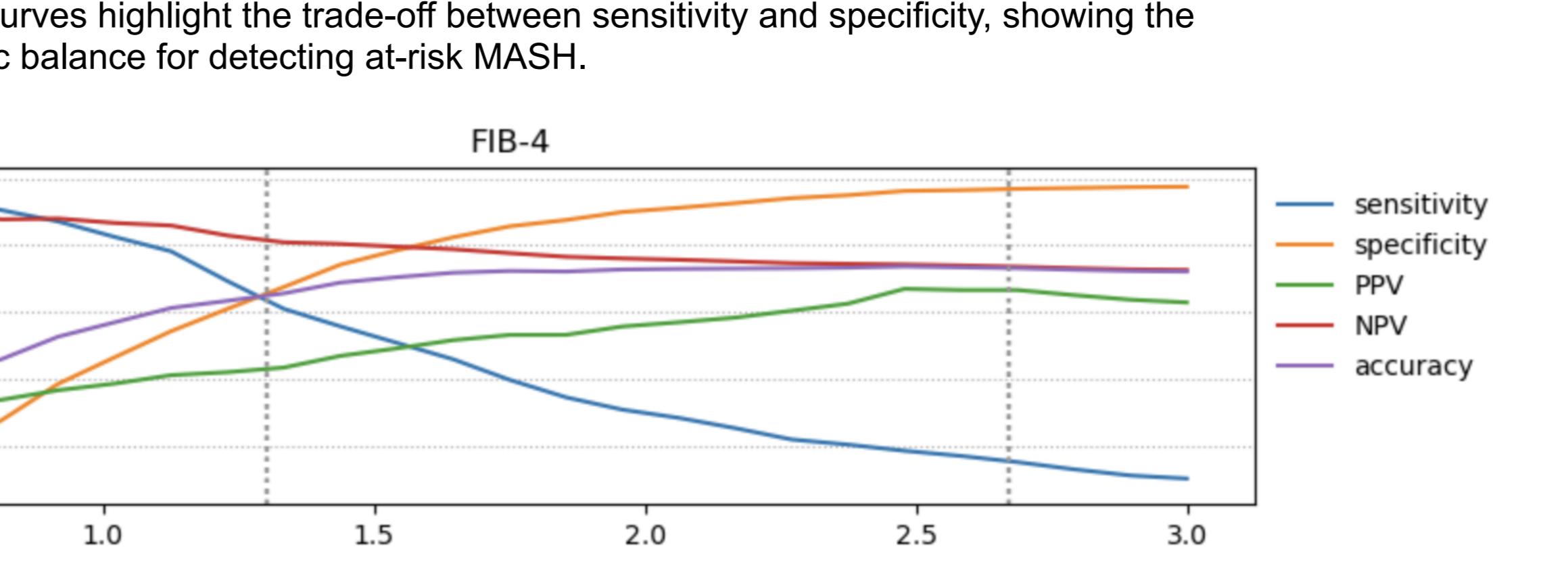
- Analysis of the 1,122-patient cohort led to the identification of optimized FIB-4 thresholds (1.08 and 2.61) that enabled a balanced trade-off across all five key diagnostic metrics for detecting at-risk MASH. Using this strategy, the combined FIB-4 and MASEF (cutoff 0.33) approach achieved an accuracy of 78%, sensitivity of 71%, specificity of 82%, PPV of 64%, and NPV of 88% for the detection of at-risk MASH.
- This study demonstrates that the simultaneous maximization of all five key diagnostic metrics is an achievable benchmark for non-invasive testing with the combination of FIB-4 index and MASEF score (OWLiver test). This approach can enhance clinical confidence, streamline diagnostic pathways, and ultimately improve the management of patients with MASLD by providing a more complete and reliable assessment of disease severity.

## REFERENCES

Noureddin M., Truong E., Mayo R., et al. Serum identification of at-risk MASH: The metabolomics-advanced steatohepatitis fibrosis score (MASEF). *Hepatology*. 2024 Jan 1;79(1):135-148. doi: 10.1097/HEP.0000000000000542. Epub 2023 Jul 24.

## CONTACT INFORMATION

Presenting author: Mirella Zulueta, MD, PhD.  
mzulueta@cimasciences.com



**Figure 2: Accuracy of the sequential diagnostic strategy combining FIB-4 and MASEF Score.**  
The figure shows the diagnostic accuracy of the MASEF Score for detecting at-risk MASH when applied after an initial FIB-4 screening step. The x-axis represents the low FIB-4 cutoff, and the y-axis represents the high FIB-4 cutoff. Subjects with FIB-4 values below the low cutoff are classified as negative (not at-risk MASH), whereas those above the high cutoff are classified as positive (at-risk MASH). Samples falling between both cutoffs are further classified using the MASEF Score. The color scale indicates the overall accuracy of this sequential diagnostic approach, highlighting the regions where the combination yields optimal performance.