

Non-Invasive Serum and Elastography-Based Markers of Fibrosis Staging in Patients with Metabolic Dysfunction-Associated Steatotic Liver Disease: A Review

Ravi Deshmukh^{1*}, Anneke Visser² and Tomas Novotny³

¹Department of Hepatology, All India Institute of Medical Sciences, New Delhi, India

²Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, Netherlands

³Department of Clinical Biochemistry, Charles University, Prague, Czech Republic

*Correspondence to: Ravi Deshmukh, Department of Hepatology, All India Institute of Medical Sciences, New Delhi, India

Received: 26 January 2024; Accepted: 09 April 2024; Published: 22 July 2024

Volume: 7; Issue: 1; ISSN: 2643-2609

Copyright: © 2024 RD, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Fibrosis stage is what determines outcome in steatotic liver disease, and the disease is far too common for biopsy to be the route to establishing it. Non-invasive tests have therefore become the pathway rather than an adjunct to it, and their thresholds now decide who is reassured, who is referred and who is treated.

Objectives: To review the published evidence on serum scores and elastography for staging fibrosis in metabolic dysfunction-associated steatotic liver disease: how each performs against histology, how performance changes with the threshold chosen, how the modalities compare, and where sequential testing strategies leave the patient.

Material and Methods: Narrative review of diagnostic accuracy studies, individual patient data analyses, meta-analyses, guidelines and care pathway trials indexed in PubMed and the major hepatology journals. Studies were eligible if they reported the accuracy of a serum score or an elastography modality against liver histology in a defined population, compared modalities directly, or evaluated a sequential testing pathway. No new patient data were generated; every accuracy estimate and threshold quoted is one published by the original investigators.

Results: Pooling 207 studies covering more than 80,000 patient investigations, FIB-4 at the conventional 1.3 threshold detected advanced fibrosis with 71 per cent sensitivity, rising to 88 per cent at a threshold below 0.75; at 2.67 its specificity was 96 per cent. Elastography performed better and improved with stage: pooling 63 studies covering 19,199 patients, transient elastography achieved areas under the curve of 0.83 for significant fibrosis, 0.87 for advanced fibrosis and 0.94 for cirrhosis, with optimal cut-offs of 7.1 to 7.9 kilopascals against 3.62 to 3.8 for magnetic resonance elastography. Thresholds achieving roughly 90 per cent sensitivity differed by modality, at 2.32 kilopascals for magnetic resonance elastography and 7.18 for transient elastography. A sequential strategy combining FIB-4 with a serum panel excluded most patients from biopsy while retaining 90.4 per cent specificity.

Conclusion: No single test stages fibrosis adequately across the range. Serum scores are cheap and reasonable for exclusion, though the widely used FIB-4 threshold of 1.3 misses close to three in ten patients with advanced fibrosis. Elastography is more accurate but least discriminating at exactly the early stages where intervention has most to offer. Sequential pathways, rather than any individual test, are what the evidence supports.

Abbreviations: MASLD – Metabolic Dysfunction-Associated Steatotic Liver Disease, MASH – Metabolic Dysfunction-Associated Steatohepatitis, NAFLD – Non-Alcoholic Fatty Liver Disease, NIT – Non-Invasive Test, FIB-4 – Fibrosis-4 Index, NFS – NAFLD Fibrosis Score, ELF – Enhanced Liver Fibrosis, APRI – Aspartate Aminotransferase to Platelet Ratio Index, VCTE – Vibration-Controlled Transient Elastography, LSM – Liver Stiffness Measurement, MRE – Magnetic Resonance Elastography, SWE – Shear Wave Elastography, AUROC – Area Under the Receiver Operating Characteristic Curve, NPV – Negative Predictive Value.

Keywords: MASLD; Liver Fibrosis; Transient Elastography; FIB-4; Non-Invasive Tests

Introduction

In 2023 a multisociety Delphi process replaced the terms non-alcoholic fatty liver disease and non-alcoholic steatohepatitis with metabolic dysfunction-associated steatotic liver disease and steatohepatitis, arguing that the old names relied on exclusionary criteria and stigmatising language [1]. An earlier expert statement

had proposed a related redefinition [2]. The change matters for a review of this kind principally because almost all the diagnostic accuracy literature predates it, and the overlap between the old and new populations has been shown to be high enough that the existing evidence transfers [3,4].

The disease is common, and becoming more so, with global prevalence estimates rising steadily and a corresponding increase in

cirrhosis and transplant listing [5,6,7]. What determines whether an individual patient is at risk is not the presence of steatosis but the stage of fibrosis: in a cohort followed for up to 33 years, fibrosis stage was the strongest predictor of disease-specific mortality [8].

That creates an arithmetic problem. Fibrosis stage is established histologically, biopsy carries a measurable complication rate [9], and the number of people with the disease is far too large for biopsy to serve as the route to staging. Non-invasive tests were developed as substitutes and have since become the pathway itself, embedded in guidelines and care algorithms [10,11,12].

Two families of test are in routine use. Serum scores combine widely available laboratory values, with FIB-4 [13], the NAFLD fibrosis score [14] and panels of direct matrix markers [15] the most established. Elastography measures liver stiffness physically, by transient elastography, shear wave techniques or magnetic resonance [16,17]. The aim of this review is to set out what each has been shown to achieve, how sensitive that achievement is to the threshold chosen, and what follows for how they should be sequenced.

Material and Methods

Search strategy: PubMed and the archives of the major hepatology and radiology journals were searched for studies of non-invasive fibrosis assessment in steatotic liver disease. Search terms combined MASLD, NAFLD, fibrosis and cirrhosis with FIB-4, NAFLD fibrosis score, enhanced liver fibrosis, transient elastography, FibroScan, shear wave elastography, magnetic resonance elastography, diagnostic accuracy and care pathway. Reference lists of retrieved guidelines and meta-analyses were screened by hand for further primary sources.

Eligibility: Studies were included if they reported the diagnostic accuracy of a serum score or elastography modality against liver histology in a defined population with steatotic liver disease; compared two or more modalities directly; evaluated a sequential testing strategy; or reported outcomes from a care pathway using these tests. Guidelines and consensus statements were included where they define thresholds in current use. Studies in other liver aetiologies are cited only where they inform methodology.

Data extraction and handling: Design, sample size, reference standard, threshold applied and reported accuracy estimates with their confidence intervals were extracted as published. Every accuracy figure is reported together with the threshold that produced it, because in this field the two are inseparable and quoting one without the other is misleading. No pooling or re-analysis was performed and no patient-level data were obtained. Figures 1 to 4 and Tables 1 and 2 reproduce published values only. This is a narrative synthesis and was not registered as a systematic review.

Results

Serum scores and the cost of the threshold: The most comprehensive comparison pooled 207 studies covering more than 80,000 patient investigations. At the conventional FIB-4 threshold of 1.3, sensitivity for advanced fibrosis was 71 per cent (95% CI 66-75). Lowering the threshold below 0.75 raised sensitivity to 88 per cent (85-91). At the upper threshold of 2.67, specificity was 96 per cent (94-97) (Figure 1) [16]. For significant rather than advanced fibrosis, sensitivity at the 1.3 threshold fell to 65 per cent, rising to 89 per cent at a threshold of 0.48 [16].

The practical consequence deserves stating directly. A threshold of 1.3 is widely used to exclude advanced fibrosis and to decide that no further testing is needed. At that threshold roughly three patients in ten with advanced fibrosis are told they do not need it.

Independent pooling in populations with type 2 diabetes has reached similar conclusions, with low cut-offs giving moderate sensitivity and specificity and high cut-offs trading sensitivity away almost entirely for specificity [18].

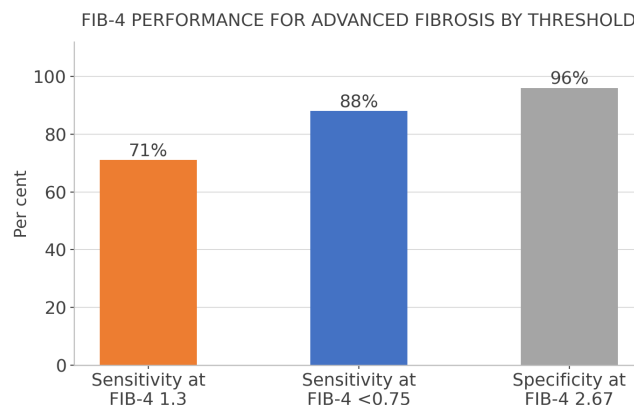


Figure 1: FIB-4 sensitivity and specificity for advanced fibrosis at three thresholds, pooled across 207 studies [16]

Elastography performs better, and better at the top end: Pooling 63 studies covering 19,199 patients, transient elastography achieved areas under the curve of 0.83 for at least F1, 0.83 for at least F2, 0.87 for at least F3 and 0.94 for cirrhosis (Figure 2) [17]. The gradient is the point: the technique is most accurate for the stage that is easiest to suspect clinically and least accurate for the early stages at which intervention would matter most.

Optimal cut-offs for advanced fibrosis from that analysis were 7.1 to 7.9 kilopascals for transient elastography and 3.62 to 3.8 for magnetic resonance elastography (Figure 4) [17]. An individual patient data meta-analysis reached comparable conclusions for magnetic resonance elastography while identifying confounders of staging [19], and a direct head-to-head comparison in patients with biopsy-proven disease found magnetic resonance elastography the more accurate of the two [20].

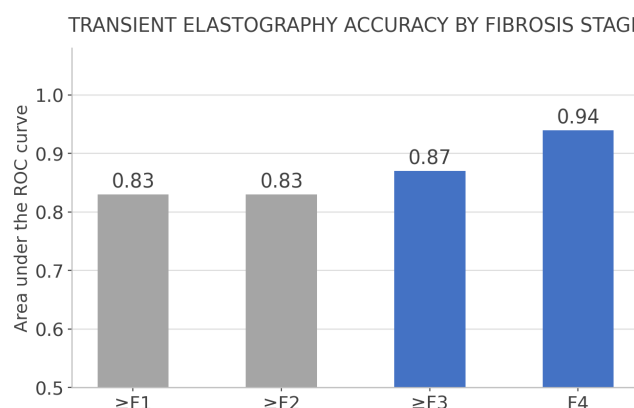


Figure 2: Pooled diagnostic accuracy of transient elastography by fibrosis stage, from 63 studies covering 19,199 patients [17]

Study	n	Design	Principal reported finding
Angulo et al. 2007 [14]	733	Derivation	NAFLD fibrosis score derived and validated
Guha et al. 2008 [15]	192	Validation	European Liver Fibrosis panel validated in NAFLD
Park et al. 2017 [20]	104	Comparison	MRE more accurate than transient elastography
Newsome et al. 2020 [21]	350	Derivation	FAST score combining stiffness, CAP and AST
Chon et al. 2024 [17]	19,199	Pooled	VCTE AUROC 0.87 for advanced fibrosis; cut-off 7.1-7.9 kPa
Wilson et al. 2025 [16]	80,000 +	Pooled	FIB-4 1.3 sensitivity only 71% for advanced fibrosis

Table 1: Principal studies reviewed, with design and the finding as reported by the original investigators

The modalities are not interchangeable: Thresholds achieving approximately 90 per cent sensitivity differ substantially between techniques: 2.32 kilopascals for magnetic resonance elastography, 3.2 for point shear wave elastography, 4.92 for two-dimensional shear wave elastography and 7.18 for transient elastography (Figure 3) [16]. Applying a transient elastography threshold to a shear wave measurement, or the reverse, would misclassify a large proportion of patients.

Attempts to unify cut-offs across modalities are ongoing and have been argued for on exactly these grounds [22,23]. In the meantime the practical requirement is that a stiffness value is reported with the modality that produced it, and that thresholds are not transferred between machines.

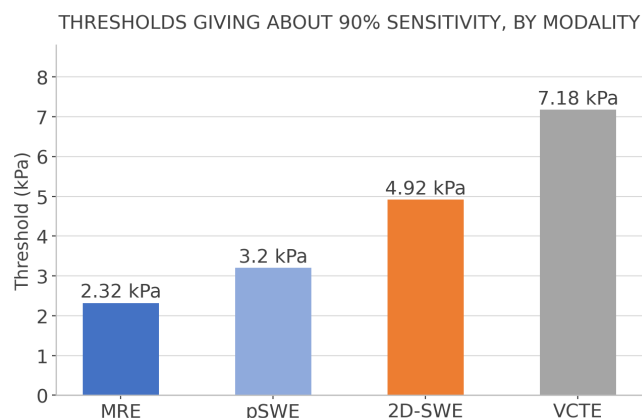


Figure 3: Thresholds required to achieve roughly 90 per cent sensitivity for advanced fibrosis, by elastography modality [16]

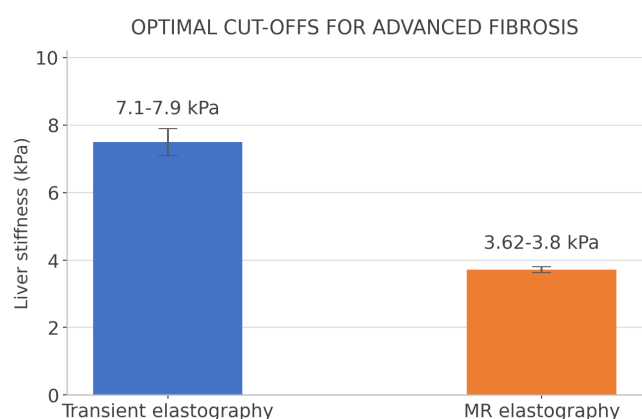


Figure 4: Optimal reported cut-off ranges for advanced fibrosis by transient and magnetic resonance elastography [17]

Technical failure is not evenly distributed: Transient elastography fails more often in the patients who most need it. The XL probe was developed specifically because measurement in overweight and obese patients was frequently unreliable with the standard probe [24], and subsequent work established how M and XL probe readings should be interpreted against one another in steatotic liver disease [25]. Since obesity is close to universal in this population, the proportion of patients in whom a reliable measurement cannot be obtained is not a minor footnote.

Serum scores have a different but equally systematic problem. FIB-4 incorporates age directly, so its performance drifts with the age structure of the population tested, and pooled subgroup analyses have found lower sensitivity in older cohorts [18]. The direct matrix marker panels avoid that dependency but cost more and are less widely available [15,26].

Sequential testing is what the evidence supports: Because no single test performs adequately across the range, guidelines have converged on two-step pathways in which a cheap, sensitive first test excludes the majority and a second, more specific test is applied to the remainder [10,11,26]. In a real-world evaluation of one such strategy, applying a serum panel threshold to patients with an indeterminate FIB-4 gave 67.9 per cent sensitivity and 90.4 per cent specificity for advanced fibrosis, with an area under the curve of 0.791, and avoided biopsy in 71.8 per cent of the patients to whom it was applied [27].

Pathways of this kind have also been tested at the level of health system implementation. A randomised trial in patients with type 2 diabetes evaluated automated fibrosis score calculation with electronic reminder messages as a route to detecting advanced liver disease [12], and a prospective primary care referral pathway showed that structured use of non-invasive tests changed which patients reached hepatology [28]. Scores combining stiffness with laboratory values, such as the FibroScan-AST score [21] and the Agile scores [29], represent a further step in the same direction.

Test	Best use	Reported performance	Principal limitation
FIB-4	First-line exclusion	Sensitivity 71% at 1.3; 88% below 0.75 [16]	Age-dependent; misses 29% at the usual threshold [16,18]
Serum matrix panels	Second-line after FIB-4	Specificity 90.4% in a two-step strategy [27]	Cost and availability [15,26]
Transient elastography	Second-line confirmation	AUROC 0.87 for advanced fibrosis [17]	Unreliable in obesity without the XL probe [24,25]
MR elastography	Problem cases	More accurate than VCTE head to head [20]	Cost and access [19,23]

Table 2: Where each test fits, with reported performance and the limitation described for it

Discussion

Two things follow from this evidence, and the first is uncomfortable. The FIB-4 threshold of 1.3, which is embedded in most published pathways as the point at which a patient can be reassured, has a pooled sensitivity of 71 per cent for advanced fibrosis [16]. Applied at population scale to a disease this prevalent, that is a large absolute number of people with advanced fibrosis who are told their liver has been checked. Lowering the threshold to 0.75 recovers most of them, at the cost of referring considerably more patients onward.

The second is that elastography inverts the accuracy profile one would want. It performs best for cirrhosis, which is often clinically apparent and where the management question is portal hypertension rather than staging [30], and worst for the transition from mild to significant fibrosis, which is where identifying a patient early enough to intervene actually matters [17]. That is a limitation of the physical measurement rather than of any particular device.

Choosing thresholds honestly

The recurring theme across this literature is that a test does not have a sensitivity; a test at a threshold has a sensitivity. FIB-4 moves from 71 to 88 per cent sensitivity across a threshold change most clinicians would regard as minor [16]. Elastography thresholds giving equivalent sensitivity differ threefold between magnetic resonance and transient techniques [16]. Reporting a result without the threshold and the modality attached conveys less information than it appears to.

This also bears on how the nomenclature change should be handled. Because the diagnostic literature was generated under the older definitions, thresholds carried forward into MASLD pathways were validated in populations defined differently, even if the overlap is substantial [1,3,4]. That is an argument for revalidation rather than for discarding the evidence.

Study limitations

The limitations belong to the underlying literature. Almost all accuracy estimates derive from biopsy-referred populations, in which the prevalence of advanced fibrosis is far higher than in the primary care populations where these tests are increasingly applied, so predictive values do not transfer [16,18]. Liver biopsy itself is an imperfect reference standard, subject to sampling variability and observer disagreement [31]. Pooled estimates carry substantial heterogeneity, with performance varying by region, age and reference standard, and the choice between biopsy and elastography as the comparator materially changes the result [16,17,18]. Several of the newer composite scores were derived and validated in overlapping cohorts [21,29].

Two gaps deserve naming. No study reviewed here reports outcomes for patients managed by a non-invasive pathway against those managed by biopsy, which is the comparison that would establish whether the substitution is safe as well as accurate. And thresholds have been derived almost entirely for cross-sectional staging rather than for detecting change over time, which is what monitoring a treated patient requires [22,32].

Conclusion

Fibrosis stage determines outcome in steatotic liver disease, and non-invasive tests are now the route by which it is established for almost everyone. The evidence supports a sequential approach rather than reliance on any single test. It does not support the widely used FIB-4 threshold of 1.3 as a safe exclusion point, since pooled sensitivity at that threshold is 71 per cent, and a lower threshold recovers most of the missed cases. Elastography is the more accurate family of tests but is least discriminating at the early stages where intervention has most to offer, and its thresholds are modality-specific and not transferable. Reporting the threshold and the modality alongside every result, and revalidating both in the populations where the tests are actually used, are the immediate requirements.

Acknowledgements

Everything discussed below was measured by others. We are grateful to the groups that ran the diagnostic studies, pooled the results and tested the pathways described here.

Funding

Nobody funded the preparation of this article, in the public sector, in industry or among charities.

Conflict of Interest

There are no competing interests to record for any author.

Ethics Statement

The article works from findings that were published previously. It gathered no new patient data and enrolled nobody, and so fell outside the requirement for ethical approval.

References

1. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, et al. "A multisociety Delphi consensus statement on new fatty liver disease nomenclature". *J Hepatol* 79(2023):1542-1556.
2. Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, et al. "A new definition for metabolic dysfunction-associated fatty liver disease: an international expert consensus statement". *J Hepatol* 73(2020):202-209.
3. Schneider KM, Schneider CV. "A new era for steatotic liver disease: evaluating the novel nomenclature in the UK Biobank". *J Hepatol* 80(2024):e58-e60.
4. Song SJ, Lai JC, Wong GL, Wong VW, Yip TC. "Can we use old NAFLD data under the new MASLD definition?". *J Hepatol* (2024). doi: 10.1016/j.jhep.2023.07.021.
5. Teng ML, Ng CH, Huang DQ, Chan KE, Tan DJ, Lim WH, et al. "Global incidence and prevalence of nonalcoholic fatty liver disease". *Clin Mol Hepatol* 29(2023):S32-S42.
6. Wong VW, Ekstedt M, Wong GL, Hagström H. "Changing epidemiology, global trends and implications for outcomes of NAFLD". *J Hepatol* 79(2023):842-852.
7. Younossi Z, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, et al. "Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention". *Nat Rev Gastroenterol Hepatol* 15(2018):11-20.
8. Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, et al. "Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up". *Hepatology* 61(2015):1547-1554.
9. Thomaidis-Brears HB, Alkhoury N, Allende D, et al. "Incidence of complications from percutaneous biopsy in chronic liver disease: a systematic review and meta-analysis". *Dig Dis Sci* 67(2022):3366-3394.
10. European Association for the Study of the Liver, Asociación Latinoamericana para el Estudio del Hígado. "EASL-ALEH clinical practice guidelines: non-invasive tests for evaluation of liver disease severity and prognosis". *J Hepatol* 63(2015):237-264.
11. Castera L, Friedrich-Rust M, Loomba R. "Noninvasive assessment of liver disease in patients with nonalcoholic fatty liver disease". *Gastroenterology* 156(2019):1264-1281.
12. Zhang X, Yip TC, Wong GL, Leow WX, Liang LY, Lim LL, et al. "Clinical care pathway to detect advanced liver disease in patients with type 2 diabetes through automated fibrosis score calculation and electronic reminder messages: a randomised controlled trial". *Gut* 72(2023):2364-2371.
13. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. "Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection". *Hepatology* 43(2006):1317-1325.

14. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, Farrell GC, et al. "The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD". *Hepatology* 45(2007):846-854.
15. Guha IN, Parkes J, Roderick P, et al. "Noninvasive markers of fibrosis in nonalcoholic fatty liver disease: validating the European Liver Fibrosis Panel and exploring simple markers". *Hepatology* 47(2008):455-460.
16. Wilson MP, Singh R, Mehta S, Murad MH, Fung C, Low G. "Comparing FIB-4, VCTE, pSWE, 2D-SWE, and MRE thresholds and diagnostic accuracies for detecting hepatic fibrosis in patients with MASLD: a systematic review and meta-analysis". *Diagnostics* 15(2025):1598.
17. Chon YE, Jin YJ, An J, Kim HY, Choi M, Jun DW, et al. "Optimal cut-offs of vibration-controlled transient elastography and magnetic resonance elastography in diagnosing advanced liver fibrosis in patients with non-alcoholic fatty liver disease: a systematic review and meta-analysis". *Clin Mol Hepatol* 30(2024):S117-S133.
18. Yoo JJ, et al. "Diagnostic accuracy of the Fibrosis-4 index for advanced liver fibrosis in nonalcoholic fatty liver disease with type 2 diabetes: a systematic review and meta-analysis". *Clin Mol Hepatol* (2024). doi: 10.3350/cmh.2024.0212.
19. Liang JX, Ampuero J, Niu H, Imajo K, Noureddin M, Behari J, et al. "An individual patient data meta-analysis to determine cut-offs for and confounders of NAFLD-fibrosis staging with magnetic resonance elastography". *J Hepatol* 79(2023):592-604.
20. Park CC, Nguyen P, Hernandez C, Bettencourt R, Ramirez K, Fortney L, et al. "Magnetic resonance elastography vs transient elastography in detection of fibrosis and noninvasive measurement of steatosis in patients with biopsy-proven nonalcoholic fatty liver disease". *Gastroenterology* 152(2017):598-607.
21. Newsome PN, Sasso M, Deeks JJ, et al. "FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study". *Lancet Gastroenterol Hepatol* 5(2020):362-373.
22. Zhang Y, Wong VW. "Towards unification of liver stiffness measurement cutoffs: editorial". *Clin Mol Hepatol* 31(2025):264-267.
23. Sohn W. "Essential tools for assessing advanced fibrosis in metabolic dysfunction-associated steatotic liver disease: editorial". *Clin Mol Hepatol* 31(2025):277-280.
24. Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D, et al. "Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients". *Hepatology* 55(2012):199-208.
25. Wong VW, Irls M, Wong GL, Shili S, Chan AW, Merrouche W, et al. "Unified interpretation of liver stiffness measurement by M and XL probes in non-alcoholic fatty liver disease". *Gut* 68(2019):2057-2064.
26. Reinson T, Buchanan RM, Byrne CD. "Noninvasive serum biomarkers for liver fibrosis in NAFLD: current and future". *Clin Mol Hepatol* 29(2023):S157-S170.
27. Kim SU, et al. "Sequential diagnostic approach using FIB-4 and ELF for predicting advanced fibrosis in metabolic dysfunction-associated steatotic liver disease". *Diagnostics* 14(2024):2517.
28. Srivastava A, Gailer R, Tanwar S, et al. "Prospective evaluation of a primary care referral pathway for patients with non-alcoholic fatty liver disease". *J Hepatol* 71(2019):371-378.
29. Sanyal AJ, Foucquier J, Younossi ZM, Harrison SA, Newsome PN, Chan WK, et al. "Enhanced diagnosis of advanced fibrosis and cirrhosis in individuals with NAFLD using FibroScan-based Agile scores". *J Hepatol* (2023). doi: 10.1016/j.jhep.2022.10.034.
30. de Franchis R, Bosch J, Garcia-Tsao G, Reiberger T, Ripoll C. "Baveno VII: renewing consensus in portal hypertension". *J Hepatol* 76(2022):959-974.
31. Sumida Y. "Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis". *World J Gastroenterol* 20(2014):475-485.
32. Haldar D, Kern B, Hodson J, Armstrong MJ, Adam R, Berlakovich G, et al. "Outcomes of liver transplantation for non-alcoholic steatohepatitis: a European Liver Transplant Registry study". *J Hepatol* 71(2019):313-322.

Citation: Deshmukh R, Visser A, Novotny T. "Non-Invasive Serum and Elastography-Based Markers of Fibrosis Staging in Patients with Metabolic Dysfunction-Associated Steatotic Liver Disease: A Review." *Ann Gastroenterol Dig Disord* (2024) Volume 7, Issue 1 doi: 10.5281/zenodo.22179529