Précision diagnostique des techniques d'élastographie versus biopsie hépatique pour la détection de la fibrose avancée et de la cirrhose compensée précoce (≥F3) : revue systématique et méta-analyse.

Ryan Kara[1], Ling He[1], Delia Munteanu[1], Lyna Bakour[1], David Nassim[2], Dany Nassim[3], Behrad Rezaie[1], Denia Hamidi[1], Ojasvi Sharma[1], Amine Benmassaoud[5], Louis-Martin Boucher[2, 4]

1. Université McGill, Faculté de médecine et des sciences de la santé, Montréal, Québec, Canada
2. Université McGill, Département de radiologie diagnostique, Montréal, Québec, Canada
3. Université de Montréal, Faculté de médecine, Montréal, Québec, Canada
4. Université McGill, Département de radiologie interventionnelle, Montréal, Québec, Canada
5. Université McGill, Département d'hépatologie, Montréal, Québec, Canada

Auteur correspondant : Ryan Kara, ryan.kara@mail.mcgill.ca (ORCID 0009-0009-4669-4128)

PROSPERO CRD420251136886

Comparative Diagnostic Accuracy of MRE, Transient Elastography, Shear-Wave Elastography, and ARFI for Advanced Hepatic Fibrosis: A Systematic Review and Meta-Analysis

Ryan Kara[1], Ling He[1], Delia Munteanu[1], Lyna Bakour[1], David Nassim[2], Dany Nassim[3], Behrad Rezaie[1], Denia Hamidi[1], Ojasvi Sharma[1], Amine Benmassaoud[5], Louis-Martin Boucher[2, 4]

1. McGill University, Faculty of Medicine and Health Sciences, Montreal, Québec, Canada
2. McGill University Department of Diagnostic Radiology, Montreal, Québec, Canada
3. University of Montreal, Faculty of Medicine, Montreal, Québec, Canada
4. McGill University, Department of Interventional Radiology, Montreal, Québec, Canada
5. McGill University Department of Hepatology Montreal, Québec, Canada

Corresponding author: Ryan Kara, ryan.kara@mail.mcgill.ca (ORCID 0009-0009-4669-4128)

PROSPERO CRD420251136886

Déclaration de conflits d’intérêts

- Les auteurs ne déclarent aucun conflit d’intérêts en lien avec cette étude et n’ont reçu aucun financement.

Identifying advanced liver fibrosis (≥F3) is clinically important

Histological fibrosis stage (METAVIR / NASH CRN)[1,2]



≥F3: advanced fibrosis and cirrhosis

Higher risk. Decompensation, portal hypertension, hepatocellular carcinoma and liver-related death.[3,4]

Still reversible. Fibrosis can regress by treating the cause: viral eradication, alcohol abstinence, weight loss.[5,6]

Guides management. Patients approaching cirrhosis need hepatocellular carcinoma surveillance.[7]

1. Rockey DC, et al. Hepatology. 2009;49(3):1017-44.
2. Kleiner DE, et al. Hepatology. 2005;41(6):1313-21.
3. Dulai PS, et al. Hepatology. 2017;65(5):1557-65.
4. Angulo P, et al. Gastroenterology. 2015;149(2):389-97 e10.
5. EASL, EASD, EASO. J Hepatol. 2024;81(3):492-542.
6. D'Ambrosio R, et al. Hepatology. 2012;56(2):532-43.
7. EASL. J Hepatol. 2025;82(2):315-74.

Limitations of liver biopsy

Invasive

Can cause pain, hemorrhage and biliary peritonitis,[8] making it poorly suited for repeated monitoring.[1]

Sampling error

A biopsy only targets a small fraction of the liver, and fibrosis is unevenly distributed.[9]

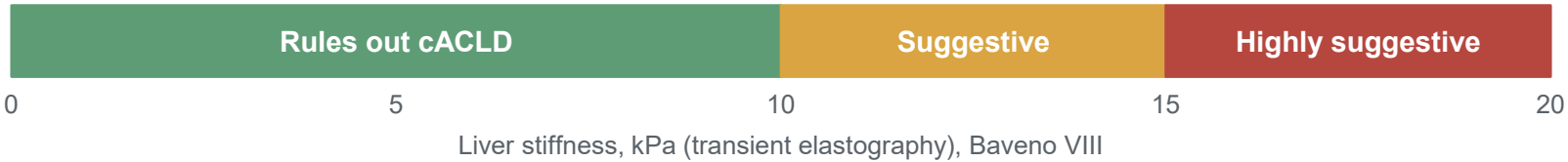
Interpretation

Histological assessment may vary from one pathologist to another, adding further imprecision.[10]

1. Rockey DC, et al. Hepatology. 2009;49(3):1017-44.
8. Piccinino F, et al. J Hepatol. 1986;2(2):165-73.
9. Regev A, et al. Am J Gastroenterol. 2002;97(10):2614-8.
10. Bedossa P. Hepatology. 1994;20(1):15-20.

Compensated advanced chronic liver disease (cACLD) is now identified by liver stiffness

Baveno VIII consensus: cACLD covers the continuum of severe fibrosis and cirrhosis and is defined by liver stiffness measurement.[11]



These values were established with transient elastography, and the consensus lists validation of the other elastography techniques as a research priority.[11]

11. Baveno VIII Faculty. J Hepatol. 2026.

How do MRE, TE, ARFI and SWE compare at ≥F3, and does etiology change the answer?

Prior meta-analyses looked at one modality, one etiology or other thresholds (≥F2, F4), and head-to-head studies are few.[12-14]

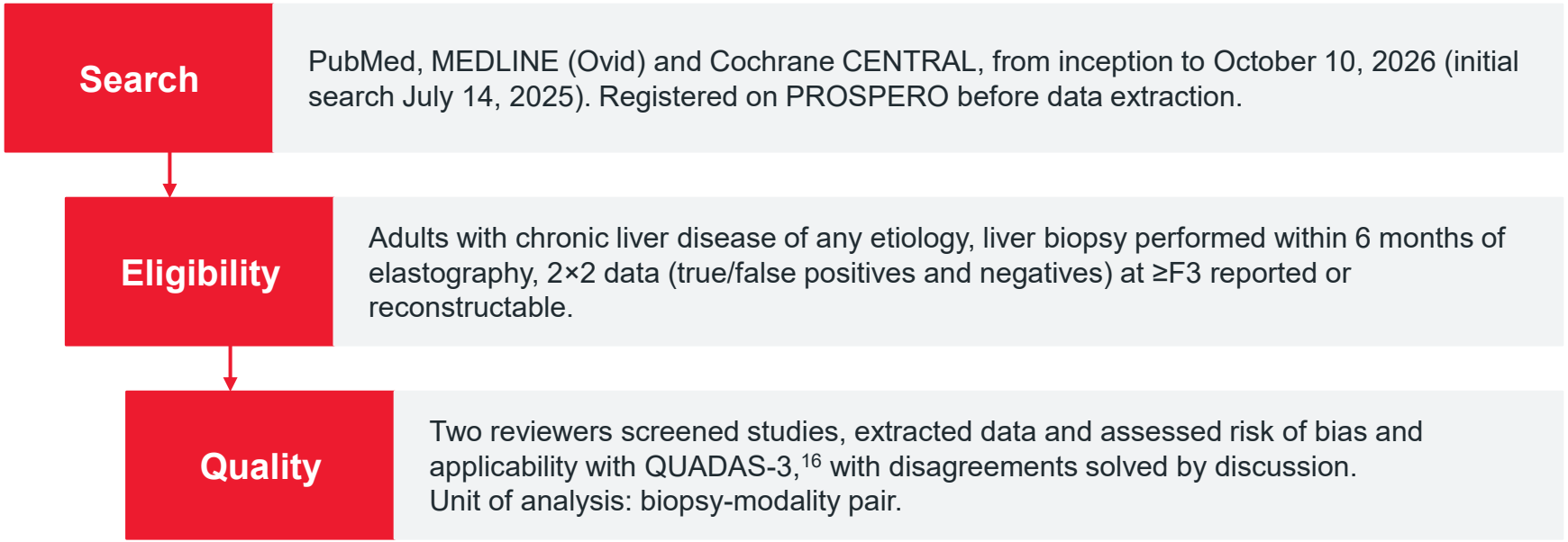
MRE	Magnetic resonance elastography
TE	Transient elastography
ARFI	Acoustic radiation force impulse imaging
SWE	Shear wave elastography

Objectives

1. Pool sensitivity and specificity of each technique against biopsy
2. Compare each technique with transient elastography
3. Test whether accuracy differs by etiology
4. Summarize the liver-stiffness thresholds used for ≥F3

12. Hsu C, et al. Clin Gastroenterol Hepatol. 2019;17(4):630-7 e8.
13. Selvaraj EA, et al. J Hepatol. 2021;75(4):770-85.
14. Singh S, et al. Clin Gastroenterol Hepatol. 2015;13(3):440-51 e6.

Search and eligibility



Reported according to PRISMA-DTA.[15]

15. McInnes MDF, et al. JAMA. 2018;319(4):388-96.
16. Whiting PF, et al. Ann Intern Med. 2026;179(4):548-55.

Statistical analysis



Pooled accuracy

Sensitivity and specificity of each technique, pooled with a bivariate random-effects model.[17]

Comparison

Meta-regression of each technique vs transient elastography (Bonferroni alpha = 0.017), and within cohorts that tested both in the same patients.

Subgroups

Analyses repeated by liver disease: MASLD/MASH, viral hepatitis, hepatitis B and C.

Robustness

Analyses repeated after removing overlapping cohorts, non-METAVIR staging, liver stiffness thresholds derived in-sample and unreported biopsy intervals.

MASLD: metabolic dysfunction-associated steatotic liver disease; MASH: metabolic dysfunction-associated steatohepatitis.

17. Reitsma JB, et al. J Clin Epidemiol. 2005;58(10):982-90.

133 studies, 168 biopsy-modality pairs, 26,588 patients

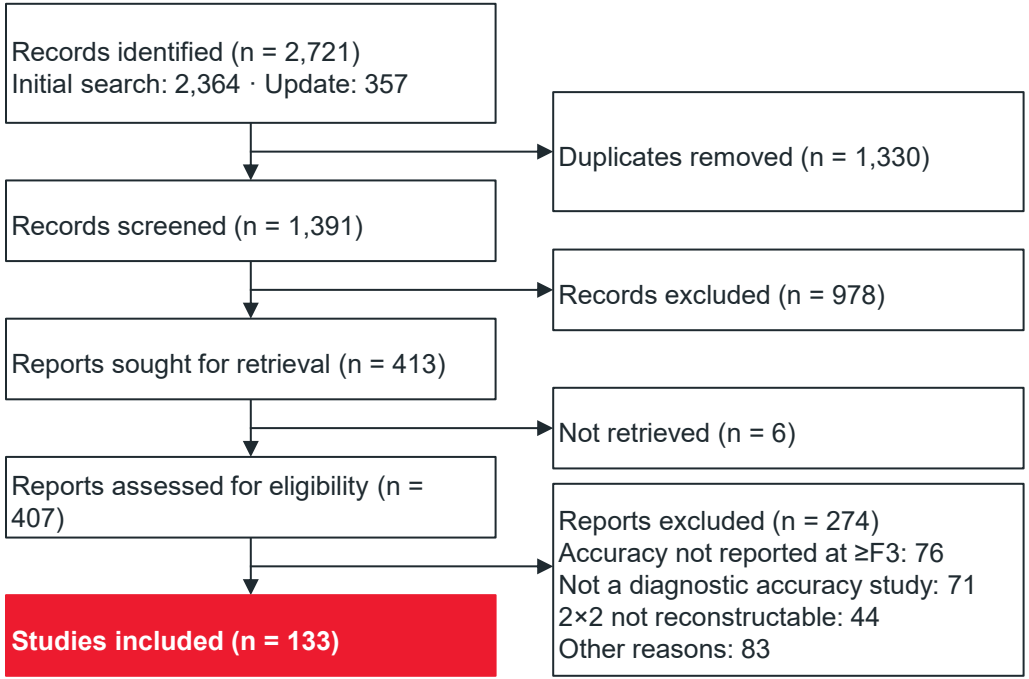


Figure 1. PRISMA 2020 flow diagram (simplified).
Initial search July 14, 2025. Updated October 10, 2026.

Modality	Biopsy-modality pairs	Observations
TE	87	20,737
ARFI	37	5,076
SWE	26	3,503
MRE	18	1,747
Total	**168**	**31,063**

Cohort: independent patient sample; 133 studies gave 136 cohorts.

Biopsy-modality pair: one imaging modality compared with biopsy in one cohort.

Observation: one patient assessed by one technique.

Patients: 26,588 unique individuals, each counted once.

Comparable diagnostic accuracy across the four techniques

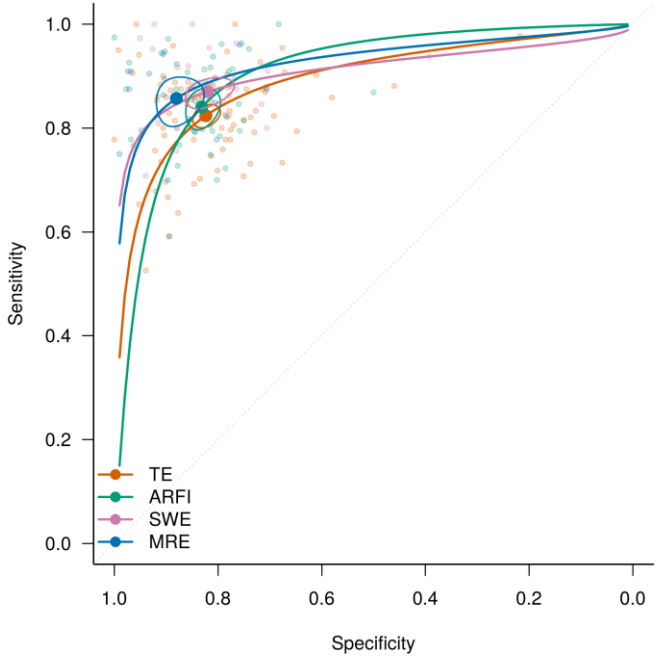


Figure 2. Summary receiver operating characteristic (ROC) curves for advanced fibrosis (≥F3) by modality.
Filled circles: pooled operating points with 95% confidence regions.
Faint points: individual biopsy-modality pairs.

Modality	n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	sAUC
MRE	18	85.7 (81.5-89.1)	**88.0 (83.9-91.2)**	**0.92**
TE	87	82.4 (80.5-84.1)	82.4 (80.1-84.4)	0.89
ARFI	37	84.1 (80.6-87.0)	83.1 (80.3-85.7)	0.90
SWE	26	**87.0 (84.4-89.2)**	82.0 (78.0-85.5)	0.90

Table 1. Pooled diagnostic accuracy for advanced fibrosis (≥F3), by modality.
Bivariate random-effects model; confidence regions overlap for all four techniques. Highest value per column in red.
n: biopsy-modality pairs; CI: confidence interval; sAUC: summary area under the ROC curve.

Comparison with transient elastography

Comparison	Overall (n = 168)	MASLD/MASH (n = 44)	Viral hepatitis (n = 68)
MRE vs TE	**2.10 (1.32-3.34), p = 0.002**	**2.83 (1.45-5.50), p = 0.002**	NP
SWE vs TE	1.45 (1.01-2.07), p = 0.042	1.45 (0.74-2.82), p = 0.28	1.56 (0.89-2.73), p = 0.12
ARFI vs TE	1.18 (0.86-1.61), p = 0.30	1.04 (0.51-2.10), p = 0.92	1.24 (0.80-1.93), p = 0.33
Overall modality effect	p = 0.012	p = 0.042	p = 0.133

Table 2. Relative diagnostic odds ratios versus transient elastography, overall and by etiology.
Values are relative diagnostic odds ratio (DOR) (95% confidence interval), p.
n: biopsy-modality pairs; NP: not pooled; MASLD/MASH: metabolic dysfunction-associated steatotic liver disease / steatohepatitis.

Indirect comparisons (Table 2). Only the MRE vs TE relative DOR remained significant after Bonferroni correction (alpha = 0.017).

Direct comparisons. Within cohorts that tested both techniques in the same patients, no difference was significant:

Comparison	Relative DOR (95% CI)	Cohorts
MRE vs TE	1.59 (0.75-3.38)	5
SWE vs TE	1.35 (0.87-2.10)	10
ARFI vs TE	0.87 (0.57-1.34)	10

Accuracy by liver disease

Modality	Etiology	n	Sensitivity, % (95% CI)	Specificity, % (95% CI)	DOR
TE	Viral	39	81.9 (78.9-84.5)	85.5 (83.2-87.5)	26.7
TE	MASLD/MASH	20	81.5 (77.1-85.2)	**77.1 (70.2-82.9)**	14.9
SWE	Viral	10	87.9 (84.4-90.7)	84.7 (80.0-88.5)	40.3
SWE	MASLD/MASH	8	86.8 (79.1-91.9)	**78.0 (65.4-87.0)**	23.2
ARFI	Viral	19	86.7 (82.5-90.0)	83.1 (80.5-85.5)	32.0
ARFI	MASLD/MASH	7	82.3 (66.3-91.7)	82.7 (74.5-88.7)	22.3
MRE	Viral	2	NP	NP	NP
MRE	MASLD/MASH	9	85.8 (80.5-89.8)	87.5 (82.6-91.2)	42.2

Table 3. Pooled diagnostic accuracy by modality and etiology.
Pooled where at least five biopsy-modality pairs were available.
n: biopsy-modality pairs; NP: not pooled; DOR: diagnostic odds ratio; MASLD/MASH: metabolic dysfunction-associated steatotic liver disease / steatohepatitis.

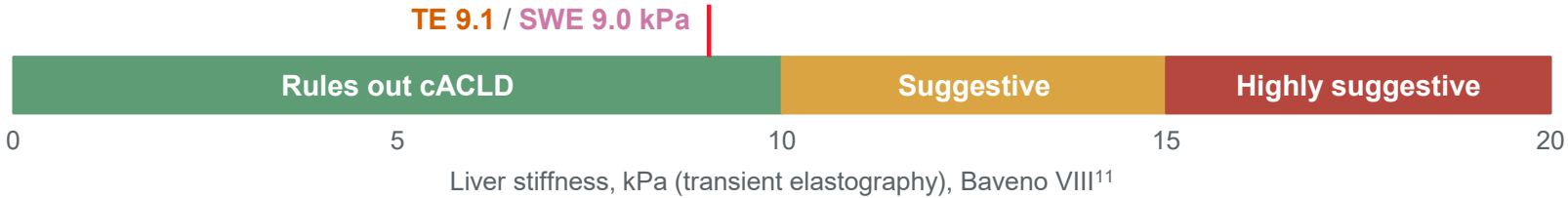
- In MASLD/MASH, specificity was lower: **77.1%** for TE and **78.0%** for SWE, versus 85.5% and 84.7% in viral hepatitis.
- ARFI specificity was similar across etiologies.
- Within MASLD/MASH, the effect of modality was significant (p = 0.042), with a higher DOR for MRE than TE (2.83, 1.45-5.50).
- Hepatitis B vs C (all modalities): specificity 83.3% vs 86.3%, DOR 25.2 vs 35.3.

Liver-stiffness diagnostic thresholds for advanced fibrosis (≥F3)

Modality	Overall	MASLD/MASH	Viral hepatitis	Mixed
TE (kPa)	**9.1 (8.3-10.2), n = 87**	9.2 (8.3-9.9), n = 20	8.9 (8.1-10.1), n = 39	9.6 (8.6-11.9), n = 22
SWE (kPa)	**9.0 (8.1-9.7), n = 22**	8.2 (7.4-8.8), n = 6	9.1 (8.6-10.1), n = 8	9.0 (8.0-9.5), n = 8
SWE (m/s)	1.59, n = 4	1.43, n = 2	1.77, n = 2	-
MRE (kPa)	3.6 (3.3-3.8), n = 18	3.6 (3.2-3.6), n = 9	3.7, n = 2	3.7 (3.4-4.1), n = 7
ARFI (m/s)	1.49 (1.36-1.68), n = 26	1.34 (1.32-1.47), n = 5	1.61 (1.44-1.72), n = 15	1.47, n = 4
ARFI (kPa)	6.8 (5.9-7.5), n = 11	8.2, n = 2	6.4, n = 4	6.7 (4.9-7.7), n = 5

Table 4. Median reported liver-stiffness thresholds for ≥F3 by modality and etiology.
Median (interquartile range) of the thresholds reported by the included studies; not derived optimal values.
n: biopsy-modality pairs; MASLD/MASH: metabolic dysfunction-associated steatotic liver disease / steatohepatitis.

The median ≥F3 thresholds for TE (9.1 kPa) and SWE (9.0 kPa) both fall below the 10 kPa Baveno VIII cut-off that rules out cACLD.[11]



11. Baveno VIII Faculty. J Hepatol. 2026.

Risk of bias and sensitivity analyses

QUADAS-3 domain	Low, n (%)	High, n (%)
D1 Participants	156 (93)	12 (7)
D2 Index test	37 (22)	**131 (78)**
D3 Target condition	162 (96)	6 (4)
D4 Analysis	154 (92)	14 (8)
Overall risk of bias	26 (15)	**142 (85)**
Applicability	140 (83)	28 (17)

Table 5. Risk of bias and applicability of the included biopsy-modality pairs (n = 168).
High index-test risk mostly reflects thresholds chosen in the same cohort.
n: biopsy-modality pairs; QUADAS-3[16]: Quality Assessment of Diagnostic Accuracy Studies version 3.

Sensitivity analyses

Pre-specified or externally derived thresholds (37 pairs). Largest change: sensitivity 2.3 points (MRE), specificity 2.0 points (TE); ARFI not pooled (2 pairs).

Overlapping cohorts removed. Largest change: sensitivity 0.4 points (SWE), specificity 0.2 points (SWE).

Unreported biopsy interval removed (9 pairs). Largest change: sensitivity 0.9 points (MRE), specificity 1.1 points (MRE).

METAVIR-only pairs (98/168). Largest change: sensitivity 1.8 points (MRE), specificity 1.9 points (ARFI).

Deeks' test. possible small-study effects for ARFI only (p = 0.006).

Most biopsy-modality pairs (142/168) were at high risk of bias, mainly because the liver stiffness threshold was derived in the same cohort. Restricting to the 37 pairs with pre-specified or externally derived diagnostic thresholds changed pooled sensitivity and specificity by 1.3 and 1.6 points (83.6 → 82.3%, 83.1 → 81.5%).

16. Whiting PF, et al. Ann Intern Med. 2026;179(4):548-55.

Limitations

In-sample thresholds. Most studies chose their diagnostic liver stiffness cut-off in the same patients, which may inflate accuracy.

Heterogeneity. Substantial, especially for specificity, and only partly explained by covariates.

Few head-to-head studies. Comparisons relied mostly on indirect evidence. MRE had the fewest included studies.

Databases. Embase not searched, although MEDLINE alone identifies nearly all studies included in diagnostic accuracy reviews.[18]

18. van Enst WA, et al. J Clin Epidemiol. 2014;67(11):1192-9.

Conclusions

1

Use the technique you have

MRE, TE, ARFI and SWE perform comparably at ≥F3, with a modest advantage for MRE that was not confirmed in direct comparisons.

2

9.0 kPa for SWE

Baveno VIII values are based on TE[11], and SWE had a similar median diagnostic threshold (9.0 vs 9.1 kPa). Therefore, we propose 9.0 kPa to identify cACLD with SWE, pending prospective validation.

3

Account for etiology

Lower specificity in MASLD/MASH suggests liver stiffness diagnostic thresholds should consider the underlying liver disease.

Next: prospective head-to-head studies with pre-specified, ultrasound system-specific diagnostic liver stiffness thresholds.

11. Baveno VIII Faculty. J Hepatol. 2026.

References

1. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD, American Association for the Study of Liver Diseases. Liver biopsy. Hepatology. 2009;49(3):1017-44.

2. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005;41(6):1313-21.

3. Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. Hepatology. 2017;65(5):1557-65.

4. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2015;149(2):389-97 e10.

5. European Association for the Study of the Liver, European Association for the Study of Diabetes, European Association for the Study of Obesity. EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). J Hepatol. 2024;81(3):492-542.

6. D'Ambrosio R, Aghemo A, Rumi MG, Ronchi G, Donato MF, Paradis V, et al. A morphometric and immunohistochemical study to assess the benefit of a sustained virological response in hepatitis C virus patients with cirrhosis. Hepatology. 2012;56(2):532-43.

7. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on the management of hepatocellular carcinoma. J Hepatol. 2025;82(2):315-74.

8. Piccinino F, Sagnelli E, Pasquale G, Giusti G. Complications following percutaneous liver biopsy. A multicentre retrospective study on 68,276 biopsies. J Hepatol. 1986;2(2):165-73.

9. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol. 2002;97(10):2614-8.

10. Bedossa P. Intraobserver and Interobserver Variations in Liver Biopsy Interpretation in Patients With Chronic Hepatitis C. Hepatology. 1994;20(1):15-20.

11. Baveno VIII Faculty. Baveno VIII - Advancing consensus in portal hypertension. J Hepatol. 2026.

12. Hsu C, Caussy C, Imajo K, Chen J, Singh S, Kaulback K, et al. Magnetic Resonance vs Transient Elastography Analysis of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Pooled Analysis of Individual Participants. Clin Gastroenterol Hepatol. 2019;17(4):630-7 e8.

13. Selvaraj EA, Mozes FE, Jayaswal ANA, Zafarmand MH, Vali Y, Lee JA, et al. Diagnostic accuracy of elastography and magnetic resonance imaging in patients with NAFLD: A systematic review and meta-analysis. J Hepatol. 2021;75(4):770-85.

14. Singh S, Venkatesh SK, Wang Z, Miller FH, Motosugi U, Low RN, et al. Diagnostic performance of magnetic resonance elastography in staging liver fibrosis: a systematic review and meta-analysis of individual participant data. Clin Gastroenterol Hepatol. 2015;13(3):440-51 e6.

15. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Prisma-DTA Group, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018;319(4):388-96.

16. Whiting PF, Tomlinson E, Rutjes AWS, Davenport CF, Yang B, Westwood ME, et al. QUADAS-3: A Revised Tool for the Quality Assessment of Diagnostic Test Accuracy Studies. Ann Intern Med. 2026;179(4):548-55.

17. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58(10):982-90.

18. van Enst WA, Scholten RJ, Whiting P, Zwinderman AH, Hooft L. Meta-epidemiologic analysis indicates that MEDLINE searches are sufficient for diagnostic test accuracy systematic reviews. J Clin Epidemiol. 2014;67(11):1192-9.

Acknowledgments: Paule Kelly-Rhéaume (McGill University Health Centre), search strategy; Kiwon Lee, PhD (McGill University), statistical review. Analyses in R 4.5.3 (mada, metafor).