Perspective | Open Access

Metabolism and Target Organ Damage

Bilson *et al. Metab Target Organ Damage.* 2026;6:37 **DOI:**10.20517/mtod.2026.92

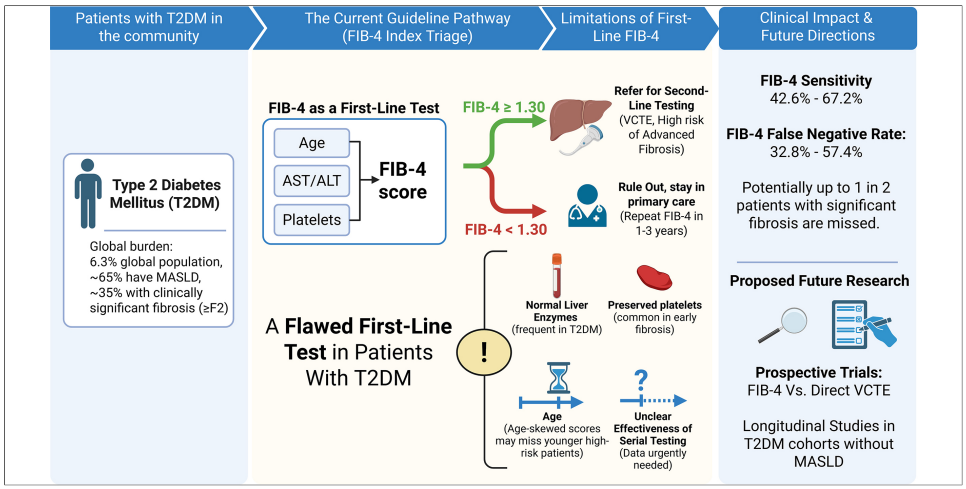
Problems with FIB-4 as an initial screening tool for liver fibrosis in patients with type 2 diabetes

Josh Bilson[1], Ryan M. Buchanan[1,2], Christopher D Byrne[1,3]

Citation: Bilson J, Buchanan RM, Byrne CD. Problems with FIB-4 as an initial screening tool for liver fibrosis in patients with type 2 diabetes. *Metab Target Organ Damage.* 2026;6:37. https://dx.doi.org/10.20517/mtod.2026.92

Received: 20 Apr 2026 **First Decision:** 9 May 2026 **Revised:** 16 Jun 2026 **Accepted:** 18 Jun 2026 **Published:** 8 Jul 2026

Academic Editor: Sonia Michael Najjar **Copy Editor:** Ting-Ting Hu **Production Editor:** Ting-Ting Hu



INTRODUCTION

Type 2 diabetes mellitus (T2DM) affects approximately 6.3% of the global population, rising to 15% among individuals aged 50-69 years[1]. T2DM is an important risk factor for metabolic dysfunction-associated steatotic liver disease (MASLD), with a reported prevalence of 65.4% among individuals with T2DM, more than double that of the general population[2]. T2DM is an important independent risk factor for liver fibrosis, and the presence and severity of liver fibrosis is the most important hepatic predictor of mortality, both for hepatic and extrahepatic disease outcomes[3]. Approximately 35% of people living with MASLD and T2DM are thought to have clinically significant liver fibrosis (F2-F4), highlighting this group as a high-risk population[2]. Given the rising global burden of T2DM and its strong association with MASLD and progressive liver fibrosis, there is increasing emphasis on early identification of at-risk individuals. In response to this growing clinical burden, guideline-recommended sequential screening strategies have been introduced to facilitate the detection of liver fibrosis in patients with T2DM[4-8].

[1]Faculty of Medicine, University of Southampton, Southampton SO17 1BJ, UK.
[2]University Hospital Southampton NHS Foundation Trust, Southampton SO16 6YD, UK.
[3]National Institute for Health and Care Research Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust and University of Southampton, Southampton SO16 6YD, UK.

Correspondence to: Prof. Christopher D Byrne, Faculty of Medicine, University of Southampton, Southampton SO17 1BJ, UK. E-mail: c.d.byrne@soton.ac.uk

Within this context, the Fibrosis-4 (FIB-4) index has become the cornerstone first-line test in sequential screening pathways for liver fibrosis in at-risk populations, where it is used to triage individuals prior to second-line elastography or enhanced liver fibrosis (ELF) test assessment. In this perspective, we propose that the current use of FIB-4 in patients with T2DM may be problematic as a first-line screening tool due to its limited sensitivity, high false-negative rates, and the questionable reliability of serial testing as a safeguard measure against missed clinically significant fibrosis in this high-risk group.

CURRENT GUIDELINE-RECOMMENDED SCREENING STRATEGIES

In primary care settings, current guidelines from European Association for the Study of Liver (EASL)-European Association for the Study of Diabetes (EASD)-European Association for the Study of Obesity (EASO), American Gastroenterology Association (AGA) and the American Association for the Study of Liver Disease (AASLD) recommend calculating the FIB-4 index as a first-line non-invasive test (NIT) in individuals at risk of chronic liver disease[4-6,9]. FIB-4 is based on the combination of age (years), aspartate transaminase (AST; U/L), alanine transaminase (ALT; U/L), and platelet count (10^9/L), calculated as follows: (*Age* * *AST*)/(*Platelet count* * *sqrt* (*ALT*)). This equation yields a numerical score rather than a direct physical or histological measurement. Clinically, this score is interpreted using specific thresholds to stratify risk, making it a simple and clinically accessible tool originally developed in patients with HIV/hepatitis C virus coinfection[10]. Importantly, some studies have shown that FIB-4 has a good ability to detect advanced liver fibrosis (≥ F3) in patients with MASLD [area under the receiver operating characteristic (AUROC) ~0.80], and can exclude the presence of advanced fibrosis with a high negative predictive value (> 90%) when using the lower cutoff of ≥ 1.30 in patients without T2DM and slightly lower in those with T2DM (~85%)[11,12].

Currently, the FIB-4 threshold of ≥ 1.30 is used as a "gatekeeper" or "rule out" test to reassure patients below this threshold that they have a low probability of clinically significant liver fibrosis (≥ F2). Through this approach, guidelines aim to identify patients with earlier stages of liver fibrosis, which mark the onset of increased mortality risk and represent a key window for therapeutic intervention and advanced disease monitoring. Patients exceeding the ≥ 1.30 FIB-4 threshold are referred for further investigation, usually with vibration-controlled transient elastography (VCTE) to measure liver stiffness as a proxy measurement of liver fibrosis. Indeed, a meta-analysis of 37 studies and 5,735 participants indicated that, when used sequentially, FIB-4 (< 1.30) followed by a liver stiffness threshold of < 8.0 kPa to rule out clinically significant liver fibrosis, significantly reduced unnecessary referrals and the need for liver biopsies[13]. Conversely, patients with a FIB-4 < 1.30 are not currently referred for second-line VCTE examination, receive lifestyle modification guidance and are advised to have a repeat FIB-4 test in 1-3 years. This sequential combination of FIB-4 and VCTE has been proposed as a cost-effective approach for identifying patients at increased risk of liver fibrosis with MASLD in primary care[14,15].

This two-tiered approach aims to balance early fibrosis detection with healthcare constraints by minimising unnecessary specialist referrals. By positioning FIB-4 as the initial screening tool, guidelines aim to balance clinical scalability and screening sensitivity with the need to avoid overwhelming secondary care VCTE assessment infrastructures. Inherently, this trade-off accepts a baseline risk of missing cases in the general population in exchange for ensuring secondary health care infrastructure. That said, a high false-negative rate leading to fewer patients being referred on for VCTE testing will inevitably miss opportunities to identify people with F2 and F3 liver fibrosis who may be eligible for newly licensed treatments targeting this level of liver fibrosis such as semaglutide and resmetirom.

The effectiveness of the current strategy assumes that the first-line test reliably identifies low-risk individuals, leaves an acceptably low number of missed cases, and that serial testing over time provides an effective safety net against initial misclassification. As discussed below, we believe that in patients with T2DM, these

assumptions are unlikely to be met, rendering a FIB-4 a problematic first-line screening tool for the identification of patients with clinically significant liver fibrosis (F2) in this high-risk population.

EVIDENCE ON FIB-4 PERFORMANCE AND IMPLICATIONS FOR TWO-TIER SCREENING PATHWAYS

Poor clinical sensitivity and high false-negative rate

As shown in Table 1, several studies have evaluated the performance of FIB-4 thresholds relative to elastography-based thresholds for clinically significant liver fibrosis; typically, 8.0 kPa, equivalent to F2 liver fibrosis by liver histology, as outlined in the EASL-EASD-EASO screening strategy. The sensitivity of FIB-4 thresholds (i.e. ≥ 1.30 or ≥ 2.00 in those ≥ 65 years) used to identify clinically significant liver fibrosis ranged from 31.7% to 72.4% in all studies[16-22]. This corresponds to false negative rates (FNR; 1-sensitivity) of between 27.6% and 68.3%, indicating that these FIB-4 thresholds may lead to a substantial proportion of patients with clinically significant liver fibrosis being misclassified as low risk. Importantly, this pattern is observed across heterogeneous cohorts of patients, including those recruited from primary care referral pathways, diabetes clinic populations, and specialist settings, and importantly persists in populations of patients perceived to be at higher risk of liver fibrosis with T2DM. When considering only those studies specifically involving populations of patients living with T2DM[18-20,22], the sensitivity of FIB-4 thresholds ranges from 42.6% to 67.2% [Table 1]. It is important to highlight that these findings reflect the performance of FIB-4 in patient cohorts with T2DM from a range of settings, including diabetes specialities centres[18,20], outpatient services and liver units[19], and primary care practices[22]. Whilst these findings indicate that FIB-4 is likely to have poor sensitivity and thereby a high false-negative rate leading to people not being referred on for VCTE throughout these patient cohorts, further studies are needed specifically in community-based cohorts of patients living with T2DM who are not known to have liver disease.

Although FIB-4 is frequently described as having a high negative predictive value (NPV), this assumption does not appear to be supported in populations with T2DM. As shown in Table 1, NPVs vary considerably across studies, particularly in cohorts of patients with T2DM, where NPVs range from 62.4% to 87.6%. This variance is likely to have profound implications for clinical decision-making. In a general population where the NPV exceeds 90%-95%, a low FIB-4 score (< 1.30) provides clinicians with adequate reassurance to "rule out" the presence of clinically significant liver disease. However, when the NPV falls to the lower ranges observed in T2DM-specific cohorts - reaching as low as 62.4% in primary care settings - the test loses its utility as an effective gatekeeper. Clinically, an NPV of 62.4% indicates that nearly 38% of patients with T2DM who are triaged as "low risk" have clinically significant liver fibrosis. Hence, we argue that instead of functioning as a reliable clinical filter, FIB-4 likely introduces an unacceptable risk of diagnostic delay to a substantial proportion of an already high-risk population of patients. Given that a test's NPV is inherently dependent on disease prevalence, its utility as a "rule-out" marker is limited in higher-risk populations, including those with T2DM where the prevalence of liver fibrosis is higher; where even modest reductions in sensitivity translate into substantial numbers of missed cases [Figure 1].

Reasons for low FIB-4 performance in patients with T2DM

It is important to acknowledge the biological and methodological factors which detrimentally impact the performance of FIB-4 in patients with T2DM. While the FIB-4 score heavily weighs elevations in transaminase (AST and ALT) as markers of hepatic dysfunction, patients with T2DM and MASLD frequently present with normal liver enzyme concentrations, often even in the presence of liver fibrosis[23]. Similarly, peripheral platelet counts are frequently preserved in early-to-intermediate fibrosis stages in patient populations with metabolic derangements, as portal hypertension typically develops at more advanced disease stages.

Table 1. Performance of FIB-4 as a first-tier screening test for significant liver fibrosis identified by liver stiffness measurement thresholds

Study	Setting (Country)	Cohort(s)	FIB-4 thresholds	Reference measure for ≥ F2 fibrosis	Performance	Comments
Davyduke *et al.* (2019)[16]	Primary care MASLD referral pathway (Canada)	*N* = 565 (17.7% with T2DM)	< 1.30 *vs.* ≥ 1.30	VCTE > 8 kPa *vs.* ≥ 8 kPa	**Se: 31.7%** **FNR: 68.3%** Sp: 89.9% NPV: 91.4%	Performance metrics were calculated from data presented in Figure 6 within the cited reference
Viganò *et al.* (2022)[17]	Patients referred for MASLD diagnosis (Italy)	*N* = 1,338 (32% with DM)	< 1.30 *vs.* ≥ 1.30	VCTE > 8 kPa *vs.* ≥ 8 kPa	**Se: 72.4%** **FNR: 27.6%** Sp: 63.8% NPV: 83.1%	Performance metrics were calculated from data presented within the cited reference
Bhuvanesswar *et al.* (2025)[18]	Diabetes Specialities Centre (Asia)	*N* = 1,070 with T2DM	< 1.30 *vs.* ≥ 1.30	VCTE > 8.1 kPa *vs.* ≥ 8.1 kPa*	**Se: 67.2%** **FNR: 32.8%** Sp: 47.1% NPV: 77.2%	Performance metrics were calculated from data presented in Table 3 within the cited reference
Mantovani *et al.* (2026)[19]	Multicentre outpatient services and liver units (Italy)	*N* = 1,203 with T2DM	< 65 years; < 1.30 *vs.* ≥ 1.30 ≥ 65 years; < 2.00 *vs.* ≥ 2.00	VCTE > 8 kPa *vs.* ≥ 8 kPa	**Se: 50.4%** **FNR: 49.6%** Sp: 66.3% NPV: 83.3%	Age-adjusted thresholds applied
Caviglia *et al.* (2026)[20]	Patients at their first referral to diabetes clinics (Italy)	*N* = 800 with T2DM	< 65 years; < 1.30 *vs.* ≥ 1.30 ≥ 65 years; < 2.00 *vs.* ≥ 2.00	VCTE > 8 kPa *vs.* ≥ 8 kPa	**Se: 43.0%** **FNR: 57.0%** Sp: 82.0% NPV: 87.6%	Age-adjusted thresholds applied
Shaheen *et al.* (2026)[21]	Primary care MASLD referral pathway (Canada)	Calgary: *N* = 8,126 *n* = 2,772 with DM Edmonton: *N* = 985 *n* = 218 with DM	< 1.30 *vs.* ≥ 1.30	Calgary: 2D-SWE > 8 kPa and ≥ 8 kPa Edmonton: VCTE > 8 kPa *vs.* ≥ 8 kPa	**Calgary whole cohort** **Se: 58.6%** **FNR: 41.4%** Sp: 72.7% NPV: 94.5% **Calgary with DM** **Se: 63.1%** **FNR: 36.9%** Sp: 67.2% NPV: 90.8% **Edmonton whole cohort** **Se: 54.5%** **FNR: 45.5%** Sp: 81.4% NPV: 91.9% **Edmonton with DM** **Se: 64.7%** **FNR: 35.3%** Sp: 74.0% NPV: 82.2%	Performance metrics were calculated from data presented in table 2 within the cited reference
Heyens *et al.* (2026)[22]	Fourteen primary care practices (Belgium and Netherlands)	Discovery cohort (complete) *N* = 482 Validation cohort without T2DM *N* = 493 With T2DM *N* = 229	< 1.30 *vs.* ≥ 1.30	VCTE > 8 kPa *vs.* ≥ 8 kPa	**Discovery cohort** **Se: 40.6%** **FNR: 59.4%** Sp: 80.1% NPV: 91.8% **Validation cohort without T2DM** **Se: 33.3%** **FNR: 66.7%** Sp: 88.4% NPV: 96.1% **Validation cohort with T2DM** **Se: 42.6%** **FNR: 57.4%** Sp: 72.5% NPV: 62.4%	Performance metrics were calculated from data presented in Tables 3 and 5 within the cited reference

*Minor variations in liver stiffness cutoffs (8.0 kPa *vs.* 8.1 kPa) reflect investigator-specific or manufacturer-calibrated thresholds utilised across cohorts to define ≥ F2 fibrosis. Bold text highlights the sensitivity and false-negative rate of FIB-4 threshold. 2D-SWE: Two-dimensional shear-wave elastography; DM: diabetes mellitus; F2: stage F2 fibrosis; FIB-4: Fibrosis-4 Index; FNR: false negative rate; kPa: kilopascals; MASLD: metabolic dysfunction-associated steatotic liver disease; NPV: negative predictive value; Se: sensitivity; Sp: specificity; T2DM: type 2 diabetes mellitus; VCTE: vibration-controlled transient elastography.

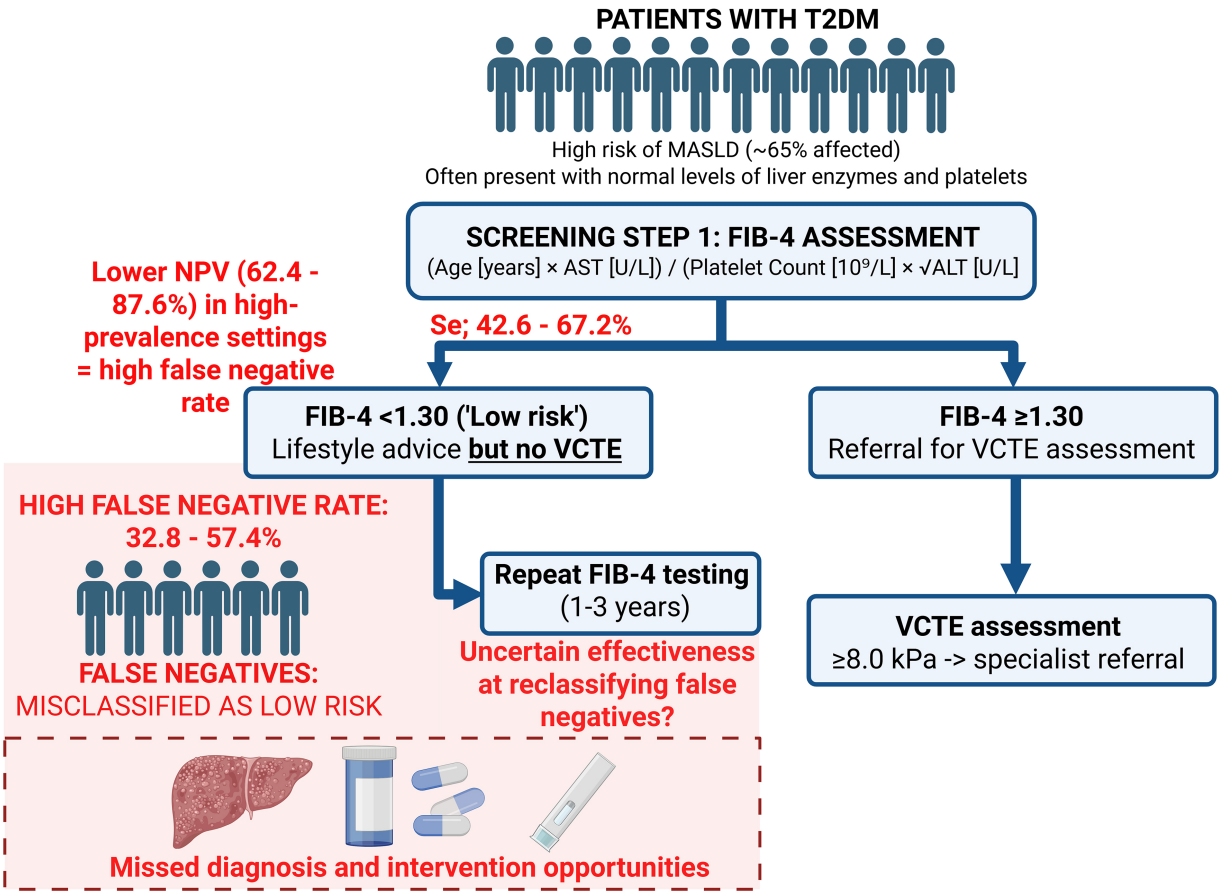


Figure 1. The performance gap of FIB-4 as a first-line tool in guidelines-recommended screening approach in patients with T2DM in a community setting. Schematic representation of the current sequential screening pathway for liver fibrosis in at risk populations. Due to the high prevalence of liver fibrosis in patients with T2DM, the NPV and Se are low, compared to those observed in patient populations with a lower prevalence of liver fibrosis. Consequently, a substantial proportion of patients with clinically significant liver fibrosis (VCTE ≥ 8.0; aligning with ≥ F2 fibrosis) fall below the 1.30 threshold, leading to false reassurance and missed opportunities for diagnosis and intervention. Created in BioRender. Bilson, J. (2026) https://BioRender.com/rfa80yb. ALT: Alanine transaminase; AST: aspartate transaminase; FIB-4: Fibrosis-4 index; MASLD: metabolic dysfunction-associated steatotic liver disease; NPV: negative predictive value; Se: sensitivity; T2DM: type 2 diabetes mellitus; VCTE: vibration-controlled transient elastography.

The FIB-4 index also suffers from methodological limitations regarding age and disease prevalence. Since age is a direct multiplier in the numerator, older patients (≥ 65 years) naturally default to higher FIB-4 scores, leading to high-false positive rates that necessitate age-adjusted cutoffs (e.g., ≥ 2.00). Conversely, in younger or middle-aged patients with T2DM, their younger age suppresses the final score, resulting in a higher false-negative classification. Additionally, the high pre-test probability and prevalence of clinically significant liver fibrosis within patient populations with T2DM degrade the negative predictive value of screening tools that possess sub-optimal clinical sensitivity. Collectively, these biological and methodological factors restrict the ability of FIB-4 to perform as an effective first-tier liver fibrosis screening tool in patients with T2DM.

Uncertainty surrounding serial FIB-4 testing for reclassifying "low-risk" patients

To mitigate false negatives, guidelines recommend repeating FIB-4 testing every 1-3 years (depending on the presence of metabolic risk factors) in patients with a FIB-4 of < 1.30[4-8]. When evaluating these guidelines, it is crucial to distinguish between data exploring the prognostic value and evidence of screening reclassification utility. A substantial body of longitudinal evidence robustly establishes that the prognostic value of tracking FIB-4 scores over time; changes are positively associated with fibrosis progression, increased liver stiffness, and adverse clinical outcomes in patients with MASLD[24-28]. However,

demonstrating that a rising score correlates with disease progression in an established cohort is fundamentally different from proving that serial testing functions as an effective clinical safety net within a screening pathway [Figure 1].

Most longitudinal studies to date include heterogeneous populations and have not specifically studied patients with T2DM from primary care settings who may have normal liver enzyme concentrations and are most at risk of false-negative classification because they have a low FIB-4 score. This uncertainty is reinforced by a large longitudinal study, involving ~40,000 participants with repeated FIB-4 testing (within 5 years) found that whilst repeated FIB-4 testing improved the identification of individuals at an increased risk of severe liver disease, almost half of all events occurred in those consistently in the low-risk group (i.e. in those with a FIB-4 of < 1.30)[28]. This paradox underscores the fundamental disconnect between a test's longitudinal prognostic tracking value and its utility as a clinical screening "safety net". That said, it should be noted that further research is urgently needed to directly explore the reclassification utility of serial FIB-4 testing specifically within T2DM screening pathways.

In addition, interpretation of longitudinal increases in FIB-4 is complicated by the influence of age, which is both a component of FIB-4 and a strong independent predictor of adverse clinical outcomes. While statistical adjustment may partially account for the impact of age, residual confounding is difficult to exclude when evaluating change over time. Taken together, although serial FIB-4 measurements clearly provide prognostic information, their role as an effective safety net for false-negative cases, especially in patients with T2DM, in screening pathways remains unclear and warrants further attention.

Advantages and implications of potential alternative approaches

The limitations highlighted above raise important concerns regarding the use of FIB-4 as a first-line test in sequential screening pathways, especially in patients with T2DM. While lowering FIB-4 thresholds may lead to improvements in sensitivity and subsequently, a reduction in false negatives, this would occur at the expense of a significant increase in the proportion of individuals referred for second-line testing. Conversely (as mentioned above), maintaining current thresholds seemingly results in high false-negative rates, which are likely exacerbated in patients with T2DM [Table 1], resulting in a large proportion of patients with clinically significant liver fibrosis being missed and not referred for further assessment [Figure 1]. Alternative strategies, such as referring all patients with T2DM directly to VCTE, would maximise diagnostic sensitivity but would introduce substantial logistical trade-offs. While cost-effective in some long-term economic models due to early intervention benefits[15,29], a direct-to-VCTE pathway would require substantial up-front investment, extensive staff training, and would likely create immediate bottlenecks in primary and secondary care settings.

To address the performance gap of FIB-4 in patients with T2DM, several non-invasive biomarker alternatives warrant further evaluation alongside the VCTE imaging pathway discussed above. Scoring systems specifically optimised or modified for patients with T2DM or metabolic dysfunction, such as metabolic dysfunction-associated fibrosis 5 (MAF-5) score and the modified FIB-4[30,31], appear to improve diagnostic sensitivity in patients with hepatic steatosis and T2DM. Alternatively, evidence indicates that implementing the ELF test as a gatekeeper test offers superior diagnostic accuracy for early-stage liver fibrosis (≥ F2) compared to FIB-4[32]. Importantly, given that the ELF test relies on direct serum markers of extracellular matrix remodelling, it retains high diagnostic performance even when transaminase concentrations are normal - a common observation in patients living with T2DM. However, the widespread adoption of first-line ELF testing is largely restricted by high recurring laboratory costs and lack of universal availability in routine primary care[14,29].

FUTURE DIRECTIONS AND CONCLUSION

In sequential testing strategies, the sensitivity of the first-line test defines the upper bound of pathway sensitivity for those ruled out at step one. We have concerns, as outlined above, regarding the use of the FIB-4 index as the first-line test for clinically significant liver fibrosis, especially in patient groups where the prevalence of liver fibrosis is not low. Specifically, we propose that in patients with T2DM, the current use of FIB-4 as a first-line test has an unacceptably high false-negative rate, poor clinical sensitivity for detecting patients with clinically significant liver fibrosis and that there is no clear evidence that serial testing functions as an effective clinical safety net in this high-risk population.

We reason that further research is needed to confirm or refute our concerns, particularly in community-based cohorts living with T2DM who are not known to have liver disease. Specifically, large-scale, prospective diagnostic accuracy studies are needed within primary care T2DM populations. These studies should look to evaluate the head-to-head performance of first-line FIB-4 triage against universal second-line VCTE or ELF panel testing to definitively establish the true rate of missed clinically significant liver fibrosis in a real-world setting. Studies, such as the REFLEX clinical trial[33], are also required to ascertain the cost-effectiveness of liver fibrosis screening strategies in patients with T2DM. Finally, longitudinal registry studies should track patients with T2DM who test as initial false negatives over a 3-5-year period to provide direct evidence on whether routine serial FIB-4 testing reliably triggers threshold crossing (≥ 1.30) in a timely manner, or if it simply delays essential specialist care, treatment opportunities with the newly licensed drugs for F2 and F3 fibrosis and advanced disease monitoring.

DECLARATIONS

Acknowledgments

The Graphical Abstract was created in BioRender. Bilson, J. (2026) https://BioRender.com/wrrea8i.

Authors' contributions

Led the writing and generation of figures and tables: Bilson J
Contributed to writing and provided expert advice and guidance: Buchanan RM, Byrne CD

Availability of data and materials

Not applicable.

AI and AI-assisted tools statement

Not applicable.

Financial support and sponsorship

Byrne CD is supported in part by the Southampton National Institute for Health and Care Research, Biomedical Research Centre (NIHR203319).

Conflicts of interest

Byrne CD is an Honorary Editor-in-Chief and Bilson J is a Youth Editorial Board Member of the journal *Metabolism and Target Organ Damage*. Byrne CD and Bilson J were not involved in any steps of editorial processing, notably including reviewers' selection, manuscript handling and decision making. Byrne CD and Buchanan RM have received research funding from Echosens, outside of this work.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes - global burden of disease and forecasted trends. *J Epidemiol Glob Health.* 2020;10:107-11. DOI PubMed PMC
2. En Li Cho E, Ang CZ, Quek J, et al. Global prevalence of non-alcoholic fatty liver disease in type 2 diabetes mellitus: an updated systematic review and meta-analysis. *Gut.* 2023;72:2138-48. DOI PubMed
3. Cusi K, Abdelmalek MF, Apovian CM, et al. Metabolic dysfunction-associated steatotic liver disease (MASLD) in people with diabetes: the need for screening and early intervention. a consensus report of the american diabetes association. *Diabetes Care.* 2025;48:1057-82. DOI PubMed
4. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology.* 2023;77:1797-835. DOI PubMed PMC
5. European Association for the Study of the Liver. EASL Clinical Practice Guidelines on non-invasive tests for evaluation of liver disease severity and prognosis - 2021 update. *J Hepatol.* 2021;75:659-89. DOI PubMed
6. European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol.* 2016;64:1388-402. DOI PubMed
7. Kanwal F, Shubrook JH, Adams LA, et al. Clinical care pathway for the risk stratification and management of patients with nonalcoholic fatty liver disease. *Gastroenterology.* 2021;161:1657-69. DOI PubMed PMC
8. ElSayed NA, Aleppo G, Aroda VR, et al. 4. comprehensive medical evaluation and assessment of comorbidities: standards of care in diabetes-2023. *Diabetes Care.* 2023;46:S49-67. DOI PubMed PMC
9. Wattacheril JJ, Abdelmalek MF, Lim JK, Sanyal AJ. AGA Clinical Practice update on the role of noninvasive biomarkers in the evaluation and management of nonalcoholic fatty liver disease: expert review. *Gastroenterology.* 2023;165:1080-8. DOI PubMed
10. Sterling RK, Lissen E, Clumeck N, et al. ; APRICOT Clinical Investigators. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology.* 2006;43:1317-25. DOI PubMed
11. Xiao G, Zhu S, Xiao X, Yan L, Yang J, Wu G. Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: a meta-analysis. *Hepatology.* 2017;66:1486-501. DOI PubMed
12. Alkayyali T, Qutranji L, Kaya E, Bakir A, Yilmaz Y. Clinical utility of noninvasive scores in assessing advanced hepatic fibrosis in patients with type 2 diabetes mellitus: a study in biopsy-proven non-alcoholic fatty liver disease. *Acta Diabetol.* 2020;57:613-8. DOI PubMed
13. Mózes FE, Lee JA, Selvaraj EA, et al. ; LITMUS Investigators. Diagnostic accuracy of non-invasive tests for advanced fibrosis in patients with NAFLD: an individual patient data meta-analysis. *Gut.* 2022;71:1006-19. DOI PubMed PMC
14. Younossi ZM, Paik JM, Henry L, Pollock RF, Stepanova M, Nader F. Economic evaluation of non-invasive test pathways for high-risk metabolic dysfunction-associated steatotic liver disease (MASLD) in the United Kingdom (UK). *Ann Hepatol.* 2025;30:101789. DOI PubMed
15. Vilar-Gomez E, Lou Z, Kong N, Vuppalanchi R, Imperiale TF, Chalasani N. Cost effectiveness of different strategies for detecting cirrhosis in patients with nonalcoholic fatty liver disease based on United States Health Care System. *Clin Gastroenterol Hepatol.* 2020;18:2305-14.e12. DOI PubMed
16. Davyduke T, Tandon P, Al-Karaghouli M, Abraldes JG, Ma MM. Impact of implementing a “FIB-4 First” strategy on a pathway for patients with NAFLD referred from primary care. *Hepatol Commun.* 2019;3:1322-33. DOI PubMed PMC
17. Viganò M, Pugliese N, Cerini F, et al. Accuracy of FIB-4 to detect elevated liver stiffness measurements in patients with non-alcoholic fatty liver disease: a cross-sectional study in referral centers. *Int J Mol Sci.* 2022;23:12489. DOI PubMed PMC
18. Bhuvanesswar KC, Srivastava BK, Amutha A, et al. Prevalence of hepatic steatosis and fibrosis in Asian Indian individuals with type 2 diabetes. *Diabetes Ther.* 2025;16:1797-811. DOI PubMed PMC
19. Mantovani A, Lombardi R, Dalbeni A, et al. Two-tier screening approach for liver fibrosis stratification in outpatients with type 2 diabetes mellitus: a multicenter cross-sectional study. *Diabetes Obes Metab.* 2026;28:644-53. DOI PubMed
20. Caviglia GP, Ferro A, D’Ambrosio R, et al. Effectiveness of a model of care based on fibrosis-4 and liver stiffness measurement for the screening of patients with type 2 diabetes mellitus at risk of advanced liver disease: results from an Italian prospective multicenter study. *Am J Gastroenterol.* 2026;121:375-82. DOI PubMed
21. Shaheen AA, Baguley E, Swain MG, et al. Diabetes and obesity reduce FIB-4 accuracy in MASLD referral pathways. *JHEP Rep.* 2026;8:101735. DOI PubMed PMC
22. Heyens LJM, van Malde DPA, Dogay Us G, et al. ; MASLD research group. Is FIB-4 the right tool for screening for liver fibrosis? *Ann Hepatol* 2025;31:102176. DOI PubMed
23. Makker J, Tariq H, Kumar K, et al. Prevalence of advanced liver fibrosis and steatosis in type-2 diabetics with normal transaminases: a prospective cohort study. *World J Gastroenterol.* 2021;27:523-33. DOI PubMed PMC

24. Yeo YH, Ho HJ, Huang TW, et al. Dynamic FIB-4 score changes and HCC risk in patients with MASLD and elevated liver enzymes: a nationwide cohort study. *Liver Cancer.* 2026;Epub ahead of print. DOI PubMed PMC
25. Zhou XD, Li YT, Kim SU, et al. ; VCTE-Prognosis Study Group, Paired Liver Biopsy Group. Longitudinal changes in fibrosis markers: monitoring stiffness/fibrosis progression and prognostic outcomes in MASLD. *Clin Gastroenterol Hepatol.* 2026;24:394-406. DOI PubMed
26. Anstee QM, Berentzen TL, Nitze LM, et al. Prognostic utility of fibrosis-4 index for risk of subsequent liver and cardiovascular events, and all-cause mortality in individuals with obesity and/or type 2 diabetes: a longitudinal cohort study. *Lancet Reg Health Eur.* 2024;36:100780. DOI PubMed PMC
27. Schreiner AD, Zhang J, Moran WP, et al. FIB-4 as a time-varying covariate and its association with severe liver disease in primary care: a time-dependent cox regression analysis. *J Clin Gastroenterol.* 2024;58:917-22. DOI PubMed PMC
28. Hagström H, Talbäck M, Andreasson A, Walldius G, Hammar N. Repeated FIB-4 measurements can help identify individuals at risk of severe liver disease. *J Hepatol.* 2020;73:1023-9. DOI PubMed
29. Noureddin M, Jones C, Alkhouri N, Gomez EV, Dieterich DT, Rinella ME; NASHNET. Screening for nonalcoholic fatty liver disease in persons with type 2 diabetes in the United States is cost-effective: a comprehensive cost-utility analysis. *Gastroenterology.* 2020;159:1985-7.e4. DOI PubMed
30. van Kleef LA, Francque SM, Prieto-Ortiz JE, et al. Metabolic dysfunction-associated fibrosis 5 (MAF-5) score predicts liver fibrosis risk and outcome in the general population with metabolic dysfunction. *Gastroenterology.* 2024;167:357-67.e9. DOI PubMed
31. Kim J, Ito T, Arai T, et al. Modified FIB-4 index in type 2 diabetes mellitus with steatosis: a non-linear predictive model for advanced hepatic fibrosis. *Diagnostics.* 2024;14:2500. DOI PubMed PMC
32. Kjaergaard M, Lindvig KP, Thorhauge KH, et al. Using the ELF test, FIB-4 and NAFLD fibrosis score to screen the population for liver disease. *J Hepatol.* 2023;79:277-86. DOI PubMed
33. Buchanan R, Reinson T, Bilson J, et al. OS-030 A multicentre randomised controlled trial of screening for MASLD in patients with type 2 diabetes not known to have liver disease in primary care. *J Hepatol.* 2026;84:S25-6. DOI