NICE National Institute for Health and Care Excellence

NICE guidance

MRI-based technologies for assessing non-alcoholic fatty liver disease

Diagnostics guidance
Published: 12 January 2023

www.nice.org.uk/guidance/dg50

Your responsibility

This guidance represents the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, healthcare professionals are expected to take this guidance fully into account, and specifically any special arrangements relating to the introduction of new interventional procedures. The guidance does not override the individual responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or guardian or carer.

All problems (adverse events) related to a medicine or medical device used for treatment or in a procedure should be reported to the Medicines and Healthcare products Regulatory Agency using the Yellow Card Scheme.

Commissioners and/or providers have a responsibility to implement the guidance, in their local context, in light of their duties to have due regard to the need to eliminate unlawful discrimination, advance equality of opportunity, and foster good relations. Nothing in this guidance should be interpreted in a way that would be inconsistent with compliance with those duties. Providers should ensure that governance structures are in place to review, authorise and monitor the introduction of new devices and procedures.

Commissioners and providers have a responsibility to promote an environmentally sustainable health and care system and should assess and reduce the environmental impact of implementing NICE recommendations wherever possible.

Contents

This guidance replaces MIB181.

1 Recommendations

1.1 There is not enough evidence to recommend LiverMultiScan or magnetic resonance elastography (MRE) to assess non-alcoholic fatty liver disease (NAFLD) in people:

- with indeterminate or discordant results from previous fibrosis testing
- when transient elastography or acoustic radiation force impulse (ARFI) elastography is unsuitable or has not worked.

1.2 Further research is recommended (see the section on further research) on:

- how the test results affect care decisions
- the test accuracy or prognostic ability of LiverMultiScan
- the test accuracy or prognostic ability of MRE.

Why the committee made these recommendations

Assessing what stage NAFLD is at can help to make decisions about care. Sometimes a biopsy is needed, which is invasive and can cause severe complications like bleeding or death. LiverMultiScan and MRE are non-invasive MRI-based tests that aim to assess the stage of NAFLD to help make decisions about care, and reduce biopsy use.

LiverMultiScan aims to identify a stage of NAFLD called non-alcoholic steatohepatitis (NASH). It is not clear how diagnosing NASH affects care decisions, partly because there are currently no medicines approved for treating NASH. This may change in the future, because there are medicines for NASH in clinical trials. Also, the clinical evidence on test accuracy and how well it can predict clinical outcomes (prognostic ability) is uncertain. There is only 1 low-quality study on whether using LiverMultiScan can reduce biopsy use.

MRE aims to identify how much liver scarring (fibrosis) there is. The company provided thresholds for staging fibrosis. There is no evidence assessing MRE's accuracy using the company's thresholds for advanced fibrosis and cirrhosis. There is also no evidence on

how MRE might affect care decisions for the people who would have it in the NHS.

The cost effectiveness of the tests is likely to depend on how much they can reduce biopsy use. In the economic model, the tests were cost effective when assuming no biopsy was done after them. But, based on available evidence, clinical experts said they would be unlikely to use the tests without a confirmatory biopsy. When assuming confirmatory biopsy would be done after all positive MRI test results, the cost-effectiveness estimates for LiverMultiScan are higher than what NICE normally considers a cost-effective use of NHS resources. MRE may be cost effective but the estimates are uncertain, largely because the cost of using MRE in the NHS is uncertain.

More evidence is needed, particularly on how LiverMultiScan or MRE test results would affect decisions about care. Also, more information is needed about whether MRI-based tests help people make lifestyle choices to help prevent or slow NAFLD progression. So, research is recommended.

2 The diagnostic tests

Clinical need and practice

The condition

2.1 Non-alcoholic fatty liver disease (NAFLD) is the term for a range of conditions caused by a build-up of fat in the liver. NAFLD develops in 4 stages:

- simple fatty liver (steatosis): a largely harmless build-up of fat in the liver cells
- non-alcoholic steatohepatitis (NASH): build-up of fat leads to inflammation
- fibrosis: persistent inflammation causes scar tissue to develop in the liver and nearby blood vessels, but the liver still functions normally
- cirrhosis: severe scarring from chronic inflammation, causing permanent damage, which can lead to liver failure and liver cancer.

Diagnosis

2.2 NAFLD is usually diagnosed using ultrasound. There are several non-invasive tests available to assess the level of fibrosis in NAFLD, including blood-based tests and imaging, such as transient elastography or acoustic radiation force impulse (ARFI) elastography. Specific tests and pathways used vary across the country.

2.3 NICE's guideline on NAFLD recommends considering using the enhanced liver fibrosis (ELF) test to look for advanced liver fibrosis in people with NAFLD. If the result of the ELF test is 10.51 or above, the NICE guideline for cirrhosis in over 16s recommends testing for cirrhosis using transient or ARFI elastography. Routine liver blood tests are not recommended to rule out NAFLD or test for advanced fibrosis.

2.4 The British Society of Gastroenterology (BSG) guideline on NAFLD recommends testing for fibrosis in people with NAFLD using the NAFLD fibrosis score or FIB-4. If these scores indicate an intermediate risk, then transient elastography or the ELF test can be used to further clarify the diagnosis. If the non-invasive tests are not able to exclude advanced fibrosis, BSG recommends that liver biopsy is considered to stage the level of inflammation and fibrosis, and to rule out other concomitant liver disease. Biopsy results are used to decide referral and treatment strategies for people with NAFLD. The NICE guideline for cirrhosis in over 16s recommends that liver biopsy is considered to diagnose cirrhosis when transient elastography is not suitable. NASH is diagnosed using biopsy. However, liver biopsy is an invasive procedure that is associated with well-recognised complications including bleeding and death. NICE's guideline on NAFLD includes a research recommendation to identify which non-invasive tests most accurately identify NASH in people with NAFLD.

2.5 The British Society of Paediatric Gastroenterology Hepatology and Nutrition guideline on fatty liver recommends that, if available, non-invasive markers of fibrosis should be used to assess NAFLD in children. In certain situations, referral for review at a national paediatric liver unit may be appropriate. Liver biopsy may be done in these units to confirm histological diagnosis or assess for NASH, fibrosis or cirrhosis.

Management

2.6 Treatment for NAFLD with no or minimal fibrosis is education on risk factors for advanced fibrosis and advice on weight management. According to NICE's guideline on NAFLD, people with advanced fibrosis may be offered pioglitazone or vitamin E. There are currently no medicines available specifically for NAFLD or NASH, but people with NASH or advanced fibrosis may enter clinical trials for new therapies.

2.7 The NICE guideline on cirrhosis in over 16s recommends that people with cirrhosis have monitoring for end-stage liver disease and liver cancer every 6 months, have testing for varices, are offered treatment for complications of cirrhosis (for example, variceal band ligation), and potentially are offered prophylactic treatment depending on comorbidities.

The interventions

LiverMultiScan

2.8 LiverMultiScan is a standalone software application produced by Perspectum that provides quantitative multiparametric analysis of non-contrast MRI. LiverMultiScan is intended to help clinicians diagnose and stage liver disease by non-invasively imaging the liver.

2.9 LiverMultiScan uses iron-corrected T1 (cT1), proton density fat fraction (PDFF) and T2* MRI protocols for its analyses. cT1 outputs are measured in milliseconds (ms), and correlate with liver fibro-inflammation. MRI PDFF is an MRI estimate of fat content and is expressed as a percentage. T2* is a measure correlated with the iron content of the liver and is used to produce the cT1 scan. The diagnosis indicated by the cT1 output and the clinical recommendations proposed by the company are as follows:

- less than 800 ms: fatty liver
  - no inflammation present
  - reassess with MRI in 3 years
- 800 ms to 875 ms: NASH
  - recommend lifestyle modification
  - manage type 2 diabetes and cardiovascular disease
  - monitor disease status with MRI after 6 months
- more than 875 ms: high-risk NASH
  - reassess with MRI every 6 months
  - consider liver biopsy if cirrhosis is suspected
  - cancer surveillance
  - consider inclusion in NASH therapeutic trials.

Magnetic resonance elastography

2.10 Magnetic resonance elastography (MRE) combines MRI with low-frequency vibrations to create a 2D or 3D elastogram showing the stiffness of tissue. In addition to the usual MRI, an external mechanical driver passes vibrations through a flexible tube to a passive driver placed on a person's abdomen over the liver. The driver is manufactured by Resoundant. MRE is intended to generate acoustic vibrations in the body during an MRI exam to assess tissue elasticity for diagnostic purposes.

2.11 MRE is used for detecting and evaluating different stages of fibrosis and is usually added to a conventional abdominal MRI protocol. MRE outputs are provided in kilopascals (kPa). The company stated that MRE liver stiffness outputs can be used to stage liver fibrosis as follows:

- more than 2.9 kPa: any fibrosis
- more than 3.3 kPa: significant fibrosis
- more than 3.9 kPa: advanced fibrosis
- more than 4.8 kPa: cirrhosis.

Before the second committee meeting, the company stated that these thresholds were suggestions and that other thresholds could be used depending on intended use.

The comparator

No further testing before a decision to do a biopsy or any other care decision

2.12 After testing as described in sections 2.2 to 2.5, in the absence of MRI-based testing, no other tests would be done before a decision to do a biopsy or any other care decision.

3 Committee discussion

The diagnostics advisory committee considered evidence on MRI-based technologies for assessing non-alcoholic fatty liver disease (NAFLD) from several sources, including a diagnostics assessment report and an overview of that report. Full details of all the evidence are in the project documents for this guidance on the NICE website.

Reducing liver biopsies would substantially benefit patients and carers

3.1 A patient expert explained that liver biopsy has many drawbacks for people with liver disease and their carers. These include the risk of complications, needing to take time off work or education to attend the procedure and for recovery, and the time it takes for biopsy results to be determined and communicated to people and their primary care team. MRI is much less invasive and has fewer associated risks, although some people cannot tolerate MRI scans. MRI may also not be suitable for people with a very high body mass index (BMI) because of the size of the scanner bore. The committee considered evidence from a survey (McKay et al. 2021) that reported that some people found biopsy very uncomfortable and caused psychological stress, but most found MRI to be harmless and tolerable. A patient expert described the impact of having to watch their child have biopsies, and their preference for non-invasive tests. The committee noted that liver biopsy can also have issues such as sampling error (that is, a biopsy can only sample a small part of the liver, which may miss affected areas). Perspectum, the company that manufactures LiverMultiScan, highlighted literature that states that biopsies sample 1/50,000th of the liver. MRI-based testing can image the whole liver. Liver biopsy results also depend on the experience of the pathologist who reports the biopsy and show intra- and inter-operator variability. The committee also noted that the risk of complications from liver biopsy is higher for people with a very high BMI, who are at higher risk of having NAFLD. Clinical experts noted that biopsy use is generally decreasing, but it is still an important option in some situations. At consultation, stakeholders highlighted the increasing prevalence of NAFLD in children, and that any technologies that could improve the non-invasive diagnosis of liver conditions for children would be beneficial. They noted that the risk associated with liver biopsy is greater in children

because they take place under general anaesthetic. The committee concluded that technologies that could reduce the need for liver biopsy would be likely to substantially benefit people and carers, in terms of health and impact on their lives.

The impact of a diagnosis of NASH on clinical management is very uncertain

3.2 Perspectum stated that LiverMultiScan should be used to distinguish non-alcoholic steatohepatitis (NASH) from simple fatty liver (see section 2.9). The committee recognised that the risk of disease progression and adverse outcomes is increased with stage of NAFLD. Clinical experts highlighted that clinical management of NASH (if fibrosis is not detected) is generally the same as for simple fatty liver. They explained that there is currently no difference in the extent of lifestyle-based interventions offered based on stage of liver disease. They emphasised that the level of fibrosis or presence of cirrhosis are the main drivers of decisions about care. A clinical expert commented that if a specialist in secondary care identified a person with NASH but no fibrosis, they would discharge them back to primary care. The company stated that cT1 results are correlated with adverse outcomes in liver disease. It further commented that the potential benefits of NASH detection could be to inform biopsy use, monitoring frequency, treatment options for comorbidities, and the extent of lifestyle advice offered. The external assessment group (EAG) noted that its assessment had modelled the use of MRI to inform biopsy use. Clinical experts commented that the progression of NAFLD can be slow and the impact of more frequent monitoring on earlier detection of disease progression is uncertain. They also highlighted that quality standard recommendations on managing NAFLD from the British Association for the Study of the Liver and the British Society of Gastroenterology NAFLD Special Interest Group state that people with low risk of significant fibrosis should be reassessed in the community every 3 years. The NICE guideline on NAFLD does not include any recommendations for interventions for people with a diagnosis of NASH, although it does include a research recommendation for non-invasive tests for NASH (see section 2.4). The committee acknowledged that new medicines for NASH are currently being developed and used in clinical trials, but these are not routinely used in the NHS.

Routine availability of these treatments would likely increase the clinical impact of a NASH diagnosis. The committee concluded that, based on current practice, the impact of a diagnosis of NASH on clinical management is very uncertain.

Introducing MRI for people with NAFLD could have a large impact on radiology capacity

3.3 The current pathway for NAFLD does not include testing with MRI. The committee noted that introducing routine MRI testing into the care pathway for NAFLD would significantly increase demand on MRI services. Perspectum suggested that the number of people with NAFLD referred for MRI would be low compared with the overall demand for MRI. Radiologist experts highlighted that wait times for MRI scans in the NHS are already long, with services working at full capacity (see section 3.12). Introducing MRI for NAFLD would either increase waiting times for MRI (for all indications, including cancer detection) or need further MRI capacity to be added to the NHS, including more scanners and trained staff, at considerable cost. If recurring scans are needed for monitoring purposes, then this would further add to the impact on MRI services. The committee concluded that greater adoption of MRI testing for NAFLD would have a large impact on the NHS. It further concluded that the benefit of introducing MRI into the NAFLD pathway would have to be very clear to justify the impact on MRI services.

Clinical effectiveness

No test accuracy evidence was found for MRI when transient or ARFI elastography is unsuitable or has not worked

3.4 Ultrasound-based tests such as transient elastography and acoustic radiation force impulse (ARFI) elastography are typically done before liver biopsy is considered (see sections 2.2 to 2.5). However, these tests are not suitable for people with a high BMI because they have a high chance of not working, particularly with central obesity. Clinical experts considered that MRI-based tests

could have particular benefit for the NHS when transient or ARFI elastography has not worked or is unsuitable. This population was identified as one of interest for this evaluation during scoping. Some studies identified included this population, but diagnostic accuracy data was not reported separately from the general study population.

There is very little evidence on the impact of MRI-based tests on decisions about care

3.5 Only 1 study was identified showing the impact of an MRI-based test (LiverMultiScan) on decisions about care, specifically the level of biopsy use (see section 3.6). There was no data on the impact of magnetic resonance elastography (MRE) on decisions about care. Clinical advice to NICE during scoping was that assessment of liver health by MRI-based technologies could help motivate people with NAFLD to engage with lifestyle changes. This could help slow or even reverse progression of liver disease. The committee acknowledged that McKay et al. (2021) provided evidence that LiverMultiScan improved some people's understanding of NAFLD. However, no data was available to determine whether LiverMultiScan or MRE affected people's adherence to lifestyle advice or interventions. Perspectum highlighted that there was evidence from other clinical areas, such as cardiology, that imaging can influence people's lifestyle choices. People with NAFLD are likely to have overweight or obesity and it was unclear to the committee the extent to which information provided by the tests would further incentivise lifestyle changes (for example, losing weight, which could have benefits beyond slowing NAFLD progression). Clinical experts highlighted that how test results affect behaviour is complex, and that they already use some quantitative data, such as weight and BMI, to incentivise lifestyle changes. The EAG highlighted the possibility that negative test results could disincentivise lifestyle changes. A further suggested benefit of MRI-based tests was on decisions about monitoring frequency. The committee recalled that the NAFLD progression can be slow. The impact of more frequent monitoring on earlier detection of disease progression is uncertain (see section 3.2). No data was identified on the impact of the tests on decisions about monitoring frequency or impact of test use on earlier detection of more advanced liver disease. Clinical experts also said that MRI could be used to help with targeting a subsequent biopsy, but no data was found on this use. People with a

South Asian family background may have a more centralised distribution of body fat than the wider population, which may increase risk of NAFLD. People with a learning disability or people taking antipsychotics may be more likely to have metabolic disorders leading to NAFLD. The committee noted that the technologies may be particularly beneficial for these groups, but concluded that there was not enough data to support this.

Direct evidence on the effect of LiverMultiScan on biopsy use is low quality

3.6 Only RADIcAL1 (a phase 4 open-label randomised controlled trial comparing LiverMultiScan [n=403] with local standard care [n=399] in people with suspected NAFLD) gave direct evidence on the impact of MRI-based tests on decisions about care. This trial assessed the number of liver biopsies avoided by using LiverMultiScan. The study was not published, but data was available from a clinical study report provided by the company. The EAG highlighted that the population in RADIcAL1 was broader than the population for this assessment. The committee noted that a relatively small number of people had a liver biopsy (55 out of 802). It further noted that the authors of the study report commented that the low number of people having biopsies was likely because there are no current treatment options for NASH. Therefore, unless the clinician suspects advanced fibrosis, the clinical management will be the same for simple fatty liver or NASH. A lower proportion of people had 'unnecessary' biopsies (defined by the study authors as biopsy with a negative NASH result) in the LiverMultiScan trial arm (9 out of 22; 41%) compared with the standard care arm (16 out of 31; 52%), although this was not statistically significant (EAG calculated odds ratio 0.65, 95% confidence interval 0.22 to 1.96). The committee and EAG noted concerns with the quality of the study, including the low number of people who had biopsy, and the lack of information about previous testing or rationale for deciding to do a biopsy. The committee also noted that there was no information on the number of necessary biopsies that had been missed because of the result of LiverMultiScan (see section 3.10).

No data for MRE was identified in the scope population, or at the company's specified cut-off values for advanced fibrosis or

cirrhosis

3.7 The EAG's diagnostic test accuracy review included data for people with NAFLD who had not had a diagnosis of advanced fibrosis or cirrhosis. No studies were identified that assessed test accuracy in the exact populations defined during scoping: people who have indeterminate or discordant results from fibrosis testing, or when transient or ARFI elastography has not worked or is unsuitable to assess fibrosis. Clinical experts reiterated that the MRE test was likely to be most useful in populations when non-invasive tests for fibrosis (such as transient elastography) could not be used, for example, because of high BMI. No diagnostic accuracy data for MRE was identified that used MRE to test for advanced fibrosis or cirrhosis using the thresholds defined by the company (3.9 kPa for advanced fibrosis, and 4.8 kPa for cirrhosis). Before the second committee meeting, the company stated that these thresholds were examples, and that other thresholds could be used depending on intended use. The committee concluded that more evidence was needed for MRE in the population of interest, at thresholds defined by the company for significant and advanced fibrosis and cirrhosis (see the section on further research).

Cost effectiveness

The disutility from missed diagnosis of liver disease is highly uncertain

3.8 In the EAG's model, liver disease that was missed by using MRI-based tests was assumed to be correctly identified 6 months later. The EAG used a value of 0.03 quality adjusted life years (QALYs) per year for the disutility associated with the liver disease that was initially missed by the tests. This disutility over the 6-month time horizon of the model had a large effect on the incremental cost-effectiveness ratio (ICER). This was because the QALY losses from false-negative results from MRI testing were often larger than the QALY gains from avoiding liver biopsy. This meant the MRI tests caused a loss of QALYs. The source of the disutility value was taken from NICE's guideline on NAFLD, and was based on the difference in QALYs expected between treated and untreated NASH. The committee and Perspectum questioned the validity of this value. The EAG agreed

and explained that it was unable to identify any alternative data to inform this parameter. Clinical experts highlighted that the progression of NAFLD can be slow and is often asymptomatic, and that the disutility accumulated over the short time horizon of the model is likely low. The EAG stated the disutility could be interpreted as a loss of QALYs from delayed diagnosis, which could happen some time after a diagnosis is made. Before the second committee meeting, Perspectum suggested that the QALY loss associated with false-negative results should be 0 because there was no strong evidence to determine the size of the disutility. The EAG stated if missing liver disease has no impact on health-related quality of life, then there would be no value to doing the test. However, it provided scenario analyses in which no disutility associated with missed liver disease was included in the base-case model, and cautioned that this should be considered exploratory analysis. It also highlighted that all ICERs exceeded £100,000 per QALY gained in this analysis. The committee concluded that the disutility associated with a missed diagnosis of liver disease is highly uncertain, but this should not be modelled as 0.

The model potentially underestimates the costs and impact of more MRI use

3.9 The EAG's model included costs of doing MRI, but not any costs for changes to NHS infrastructure that may be needed for more MRI use. The EAG commented that the implications for NHS service provision would be significant. This is because of increased staffing levels and changes in infrastructure needed to accommodate the high demand for MRI scans for people with NAFLD. The committee recalled its conclusion that more MRI testing in NAFLD would have a significant impact on demand for MRI. This could mean purchasing more MRI scanners or increased waiting times for MRI scans (see section 3.3). The committee concluded that the true cost of introducing MRI to the care pathway for NAFLD would likely be higher than estimated in the model.

Based on current data, a biopsy is likely to be done after a positive MRI test result

3.10 The EAG's base-case model assumed that all people with a positive result from

MRI testing would then be referred for a confirmatory biopsy. In comments submitted on the diagnostics assessment report, the companies suggested that a positive test result was sufficient for a diagnosis for some people, and that a confirmatory biopsy was not always necessary. The EAG noted that there was no data to inform any assumption about the effect of a positive MRI test result on the clinical decision to do a biopsy. It explained that, based on clinical advice, it considered it appropriate to assume a confirmatory biopsy would be needed. However, before the second committee meeting, the EAG did a scenario analysis in which no confirmatory biopsy was done after a positive MRI result. In this analysis, LiverMultiScan and MRE dominated (were less expensive and more effective than) the biopsy-only strategy. The EAG cautioned that the costs and QALYs associated with false positives from MRI tests (which would be assumed to be detected if confirmatory biopsy was done) were unknown and not modelled. Also, this assumption did not take into account that biopsies may be done to obtain information other than staging NAFLD. Clinical experts highlighted that biopsy is important for differential diagnosis. For LiverMultiScan, a positive result was stated by the company to distinguish NASH from simple fatty liver (see section 3.2), and not to stage fibrosis. If a test for fibrosis was needed, a biopsy may still be done. There was also limited data on the impact of tests on biopsy use, and the only study identified (RADIcAL1, see section 3.6) found about a 30% decrease in biopsy use, rather than 100% as in the EAG's scenario. Perspectum stated that LiverMultiScan would not entirely replace biopsy but could help identify people who could benefit most from it. Clinical experts commented that current data on test performance was not sufficient to be confident in a diagnosis without a biopsy. The committee concluded that if further data provides reassurance on test performance, a follow-up biopsy may not always be needed, but it is inappropriate to assume that the tests can replace biopsy entirely.

How much the tests affect decisions about care for people who do not want a biopsy is unclear

3.11 Perspectum stated that between 5% (South Warwickshire Foundation Trust) and 50% (feedback from clinicians to company) of people offered liver biopsy decline the procedure. It suggested that MRI could be an appropriate method of staging liver disease in these people. The EAG modelled a scenario in which no biopsy was done in either the intervention or comparator arms (to represent people who

would not have a biopsy in current practice). It estimated the number of additional QALYs LiverMultiScan would need to generate to achieve an ICER of £30,000 per QALY gained. The committee noted that it was still unclear for this population what changes to care would be made based on the MRI test results and therefore how the additional QALYs could be generated.

Time to test is not accounted for in the model

3.12 Perspectum stated that people can wait up to 18 months for a liver biopsy. It questioned whether this was incorporated in the model, noting that reducing time to testing could be an uncaptured benefit for the MRI-based tests. The EAG confirmed that wait times had not been included in the model. Clinical and patient experts stated that their experience of wait times for liver biopsy were much lower than suggested by the company, between 2 days and 6 months. The committee considered data submitted by Perspectum before the second committee meeting on the average waiting time to a first outpatient appointment, and to treatment, in gastroenterology departments across 7 hospitals, representing the longest waits in the 7 main regions of England. These varied between 19 and 38 weeks. The committee noted that there are currently significant wait times for MRI, although this can vary across the country, and that introducing MRI to the NAFLD care pathway could further increase the wait (see section 3.3). Therefore, it was not appropriate to assume that an MRI test would be done as soon as needed, and any quicker time to test compared with biopsy was very uncertain. A patient expert commented that if confirmatory biopsy was needed after a positive MRI test result (see section 3.10), introducing MRI could also increase the time to diagnosis compared with a pathway in which liver biopsy is done without a preceding MRI test. The EAG commented that it had looked for data on the impact of the MRI tests on time to diagnosis but was unable to find any. The committee concluded that there was no evidence to demonstrate that adding MRI to the NAFLD pathway would reduce time to diagnosis, and while there was the potential for this to be a benefit of testing, the size of any benefit was highly uncertain. The extent of any benefit may depend on what actions could be taken based on MRI tests alone. The extent of impact on QALYs of a quicker diagnosis from testing is also uncertain (see section 3.8).

The impact of complications from biopsy is uncertain

3.13 Values used in the EAG's model for the costs and disutility associated with complications from liver biopsy were questioned by stakeholders. Perspectum stated that the cost used for biopsy complications (£8.54 per liver biopsy) was too low and was based on assumptions. It suggested that costs from a recent economic evaluation of LiverMultiScan for people with autoimmune hepatitis (Bajre et al. 2022) would be more appropriate (£168.67 per liver biopsy). The EAG did a threshold analysis that found that the cost of complications would have to increase by more than 10,000% for the ICERs from the base-case model to reach a threshold of £30,000 per QALY gained, if using LiverMultiScan to test for advanced NASH. Perspectum also queried the disutility values assigned to complications from liver biopsy, including those from death. It stated that the QALY loss from biopsy complications used in the model was based on a calculation error from a previous economic analysis (Stevenson et al. 2012), so was underestimated by a factor of 10. The EAG did a scenario analysis using the corrected disutility from biopsy complications, and found that the ICERs remained above £100,000 per QALY gained for all strategies. It also did a threshold analysis that found that the QALY decrement from complications in its base case would have to increase by almost 20,000% to get an ICER of £30,000 per QALY gained if using LiverMultiScan to test for advanced NASH. The committee agreed that the impact of complications from biopsy was uncertain. It concluded that the costs and QALYs would have to change by a large amount from the values used in the EAG's base case for the interventions to be cost effective.

In the EAG's base case, LiverMultiScan is dominated or has much higher ICERs than are usually considered acceptable

3.14 Perspectum stated that LiverMultiScan was intended to distinguish NASH from simple fatty liver. So, the committee focused its considerations for LiverMultiScan on the cost-effectiveness estimates provided by the EAG for NASH, advanced NASH, and high risk of progressive disease (defined as NASH or at least F2 fibrosis). In the base case, LiverMultiScan was dominated by the biopsy-only pathway. This was because of the QALY losses incurred by false-negative results, which the committee recalled were very uncertain (see section 3.8). This QALY loss was removed in a scenario analysis, which improved cost effectiveness (the

test no longer reduced QALYs). However, the ICERs were above £118,000 per QALY gained. The EAG commented that this scenario is not plausible as it would imply there was no impact of a correct or incorrect diagnosis, and consequently no point to testing. The EAG commented that the cost effectiveness of LiverMultiScan would be above £30,000 per QALY gained even if the test was assumed to have 100% accuracy. Threshold analyses indicated that the population prevalence of the condition being tested for would have to be much lower than found in the study by Eddowes et al. (which was used in the EAG's base case) for LiverMultiScan to be cost effective. The committee noted that the extent of decrease in biopsy use caused by LiverMultiScan use estimated by the model (up to about 30% reduction) was similar to that seen in RADIcAL1.

MRE could be cost effective, but this is highly dependent on the cost per test

3.15 Resoundant, the company that manufactures the technology used in MRE, stated that no additional cost per scan would be necessary for MRE if the hardware was already available. The EAG used 2 cost per scan estimates for MRE. The first estimate assumed that MRE was already installed, so the cost of a scan was the only cost of doing an MRI (that is, no additional cost for using MRE). Based on accuracy estimates to detect significant fibrosis (at least F2), MRE dominated the biopsy-only pathway. The second estimate assumed that MRE would need to be installed and added an additional £59.50 per scan on top of the cost of doing an MRI. The EAG noted that this additional cost was uncertain, being based on several assumptions, many of which there was no evidence for. Using this cost, the ICER for detecting significant fibrosis was over £225,000 per QALY gained. Radiology experts in the committee stated that MRE is not widely available in the NHS, so the second cost-effectiveness estimate is more realistic. The EAG did not model MRE to detect advanced fibrosis (at least F3) or cirrhosis because no data was identified using the thresholds the company specified to NICE (see section 3.7). The committee also had concerns that using different costs based on whether or not a test is already available could result in inequality based on geographical location. The committee concluded that the cost of MRE was a significant factor in whether or not the test could be cost effective, and that the true cost was highly uncertain. Further clarification of the cost per person of MRE testing in the NHS would benefit future decision making.

Test accuracy data for MRE is needed in the scope population

3.16 The EAG provided cost-effectiveness estimates for MRE using data from Imajo et al. (2021). It highlighted that this population was broader than the scope population, and that a subgroup analysis based on the scope population was not possible. Clinical experts commented that MRE could have a role in the NHS if used when previous tests such as transient or ARFI elastography either could not be done, had not worked, or gave discordant results, in line with the scope population. The committee concluded that data on MRE performance is needed in populations that match the scope, from either subgroup analysis of existing studies, or further accuracy studies.

Several assumptions in the model need further consideration once more data is available

3.17 The EAG made several assumptions in its economic model, including that:

- after an initial negative result from an MRI test (less than 875 ms), a second MRI test would be done at 6 months
- the second MRI test at 6 months was 100% accurate (that is, no liver disease remains undetected after 6 months).

  Perspectum commented on the diagnostic assessment report that the pathway used in the model did not reflect its clinical advice for use of LiverMultiScan. It explained that, in line with its advice, people would only have a follow-up LiverMultiScan if their cT1 score was between 800 ms and 875 ms. People with a score below 800 ms would be discharged to primary care and their condition monitored every 3 years as per local guidance (see section 2.9). The EAG modelled a scenario in which people with LiverMultiScan results less than 800 ms had no further testing. ICERs remained above £200,000 per QALY gained. The EAG also noted that extending the time before a follow-up test would reduce the cost effectiveness of the MRI-based tests. This is because those with false-negative results from the first test would accrue disutility from undiagnosed liver disease over a longer period (see section 3.8). A stakeholder also questioned the assumption that the second test done at 6 months would be

100% accurate. The EAG acknowledged that this was a simplifying assumption that favoured the MRI-based tests. The committee recalled liver biopsies can have sampling errors (see section 3.1), but that accuracy estimates used in the model were from studies that used liver biopsy as a reference standard. This may be unfavourable towards the MRI tests. The committee concluded that several assumptions used in the model would need reconsidering when more clinical data is available. This would improve the reliability of the cost-effectiveness estimates.

The EAG's model is suitable for decision making

3.18 Stakeholders questioned values used in the economic model (see section 3.8, section 3.13 and section 3.17). The EAG acknowledged uncertainty in model inputs. It highlighted that it had done extensive analyses to investigate how much parameters would have to change by for its base-case ICER to become cost effective. Perspectum said that the EAG should have varied uncertain parameters together, rather than one at a time in threshold analyses. The company provided its preferred model parameters for biopsy complications, QALY loss associated with false-negative test results and biopsies. Perspectum ran analyses using these values and no cost for a second MRI scan using LiverMultiScan, based on the company proposing to pay for any follow-up scans done after an initial negative result. Using Perspectum's preferred inputs, and assuming no cost for a follow-up scan, LiverMultiScan had ICERs below £30,000 per QALY gained when testing for NASH or advanced NASH. The EAG noted that the cost of complications used came from a 2022 economic evaluation of LiverMultiScan for people with autoimmune hepatitis (see section 3.13), and that the discounted MRI cost did not factor in the general cost of an MRI in the NHS (£148.24 in the EAG's base case). The EAG's threshold analyses showed that the model parameters would have to change in its base case by what the EAG considered an unfeasibly large degree for LiverMultiScan to have an ICER under £30,000 per QALY gained. So, the EAG considered that the conclusions of its analyses were robust to uncertainty in these parameters. The committee agreed and concluded that although there was considerable uncertainty in the model's parameters, the EAG's model and accompanying analyses were suitable for decision making. It also recalled that several assumptions in the model were highly favourable to the MRI tests (see section 3.17).

More data would help determine LiverMultiScan's accuracy

3.19 There was limited data on test accuracy, with only 1 small study identified for LiverMultiScan that explicitly included the scope population (Eddowes et al. 2018). Perspectum stated that other relevant biopsy-paired evidence had been submitted and inappropriately excluded by the EAG. The EAG explained that evidence had been excluded because it reported populations that had been used in other studies (that were included in the EAG's report), or were not the focus of the assessment, or did not report validation against liver biopsy (see the diagnostic assessment report, addendum 2). The committee considered that the accuracy estimates used in the EAG's model from the Eddowes study were not particularly high (for example, for advanced NASH, 64% sensitivity and 62% specificity). Unpublished accuracy data from RADIcAL1 was provided by the company at consultation on draft guidance, which reported higher specificity (90%) for advanced NASH. However, this was from only 18 people. The committee considered that further data on test accuracy would be highly beneficial to help estimate true test accuracy. It noted that validation against biopsy may underestimate test performance because of issues with sampling bias (see section 3.1). There may also be issues with getting a biopsy result for a reference standard if this is not clinically indicated, or a person refuses biopsy. Clinical experts noted that LiverMultiScan outputs are correlated with fibro-inflammation (see section 2.9), and cannot provide information on hepatocyte ballooning, one of the histological components of NASH. The committee considered that studies showing how well LiverMultiScan results predicted later clinical events could be used as an alternative to assess test performance. Perspectum stated that such data is becoming available. One study was identified by the EAG for LiverMultiScan (Jayaswal et al. 2021). However, this study included multiple liver disease aetiologies (people with NAFLD, alcohol-related liver disease and viral hepatitis), and the study did not report results for NAFLD separately. The committee concluded that it would be beneficial to see further test accuracy data for LiverMultiScan compared with biopsy, or evidence for LiverMultiScan's prognostic ability for clinical outcomes.

More data is needed to assess MRE performance at set thresholds

3.20 No studies were identified using MRE in the scope population (see section 3.7),

and no studies assessed MRE using the thresholds that the company stated for advanced fibrosis or cirrhosis. Before the second committee meeting, the company commented that it does not have stated manufacturer cut-offs. The committee concluded that more data on the test accuracy of MRE is needed using prespecified thresholds set by the company, done in a population not used to derive these thresholds (external validation). Data showing prognostic ability for clinical outcomes could be another measure of test performance.

There is considerable uncertainty about how LiverMultiScan results would affect care in the NHS

3.21 Clinical experts said it is not clear what care decisions LiverMultiScan results would affect, or how people may adhere to lifestyle advice or interventions based on results. The clinical impact of a NASH diagnosis is uncertain (see section 3.2), and there was little evidence for the effect of LiverMultiScan results on clinical management (see section 3.5). The only available direct evidence for the impact of LiverMultiScan outputs on the number of liver biopsies was of poor quality (see section 3.6). The committee recognised that future approvals of drug treatments for NASH may make the role of LiverMultiScan and NASH diagnosis clearer. Clinical experts also commented that it is unclear how LiverMultiScan can distinguish fibrosis from inflammation (see section 3.19). If this is not possible, a positive result could be because of either fibrosis or inflammation, and it was not clear how the technique could identify NASH alone. Further tests may be needed to make decisions about care. The committee concluded that there is considerable uncertainty about how LiverMultiScan results would affect care in the NHS and how people follow lifestyle advice, and that these are key elements of determining the cost effectiveness of the test.

Research considerations

Research should represent how the tests would be used in practice and include people who could particularly benefit from testing

3.22 The committee specified that further research should be in a population representative of how the tests would be used in practice, and should include people who could particularly benefit from testing. It mentioned that MRI-based testing may particularly benefit people with a high BMI for whom other liver tests such as transient elastography may not work as well (see section 3.1). People who would decline an offered biopsy were highlighted by a company as a group who could particularly benefit from access to more non-invasive test options (see section 3.11). Stakeholders noted that the risks of biopsy are higher for children (see section 3.1) and that more non-invasive test options could reduce risk.

4 Recommendations for further research

4.1 Further research is recommended on:

- the impact of LiverMultiScan test results on decisions about care, such as the decision to do a liver biopsy, or on adherence to lifestyle interventions (see section 3.5)
- the test accuracy of LiverMultiScan compared with biopsy, or the prognostic ability of LiverMultiScan to predict clinical outcomes (see section 3.19).

4.2 Further research is recommended on:

- the impact of magnetic resonance elastography (MRE) test results on decisions about care, such as the decision to do a liver biopsy, or on adherence to lifestyle interventions (see section 3.5)
- the test accuracy of MRE to assess the level of fibrosis or presence of cirrhosis compared with biopsy at thresholds specified by the company, or the prognostic ability of MRE to predict clinical outcomes (see section 3.20).

5 Implementation

NICE intends to develop tools, in association with relevant stakeholders, to help organisations put this guidance into practice.

In addition, NICE will support this guidance through a range of activities to promote the recommendations for further research. The research proposed will be considered for developing specific research study protocols as appropriate. NICE will also incorporate the research recommendations in section 4 into its guidance research recommendations database and highlight these recommendations to public research bodies.

6 Diagnostics advisory committee members and NICE project team

Committee members

This topic was considered by the diagnostics advisory committee, which is a standing advisory committee of NICE.

Committee members are asked to declare any interests in the test to be assessed. If it is considered there is a conflict of interest, the member is excluded from participating further in that assessment.

The minutes of each committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Additional specialist committee members took part in the discussions for this topic:

Specialist committee members

Dr Raneem Albazaz
Consultant radiologist, St James's University Hospital Leeds

Heather Boult
Lay specialist committee member

Dr Pinelopi Manousou
Consultant hepatologist, Imperial College NHS Trust

Prof Alastair O'Brien
Professor and honorary consultant hepatologist, University College London Hospital and Royal Free Hospitals

Dr Jeremy Shearman
Consultant gastroenterologist and hepatologist, Warwick Hospital

NICE project team

Each diagnostics assessment is assigned to a team consisting of a technical analyst (who acts as the topic lead), a technical adviser and a project manager.

Jacob Grant
Topic lead

Thomas Walker
Technical adviser

Donna Barnes
Project manager (until April 2022)

Harriet Wilson
Project manager (from May 2022)

ISBN: 978-1-4731-4948-9

Accreditation