



Published in final edited form as:

Hepatology. 2011 September 2; 54(3): 1082–1090. doi:10.1002/hep.24452.

Diagnostic Accuracy and Reliability of Ultrasonography for the Detection of Fatty Liver: A Meta-Analysis

Ruben Hernaez^{#1,2,3}, Mariana Lazo^{#1}, Susanne Bonekamp⁴, Ihab Kamel⁴, Frederick L. Brancati^{1,3,5}, Eliseo Guallar^{3,5,6}, and Jeanne M. Clark^{1,3,5}

¹Department of Medicine The Johns Hopkins School of Medicine, Baltimore, MD

²Department of Medicine, Washington Hospital Center, Washington, DC

³Department of Epidemiology, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

⁴Department of Radiology, The Johns Hopkins University School of Medicine, Baltimore, MD

⁵Welch Center for Prevention, Epidemiology, and Clinical Research, The Johns Hopkins University, Baltimore, MD

⁶Department of Cardiovascular Epidemiology and Population Genetics, National Center for Cardiovascular Research (CNIC), Madrid, Spain

These authors contributed equally to this work.

Abstract

Ultrasonography is a widely accessible imaging technique for the detection of fatty liver, but the reported accuracy and reliability have been inconsistent across studies. We aimed to perform a systematic review and meta-analysis of the diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver. We used MEDLINE and Embase from October 1967 to March 2010. Studies that provided cross-tabulations of ultrasonography versus histology or standard imaging techniques, or that provided reliability data for ultra-sonography, were included. Study variables were independently abstracted by three reviewers and double checked by one reviewer. Forty-nine (4720 participants) studies were included for the meta-analysis of diagnostic accuracy. The overall sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio of ultrasound for the detection of moderate-severe fatty liver, compared to histology (gold standard), were 84.8% (95% confidence interval: 79.5-88.9), 93.6% (87.2-97.0), 13.3 (6.4-27.6), and 0.16 (0.12-0.22), respectively. The area under the summary receiving operating characteristics curve was 0.93 (0.91-0.95). Reliability of ultrasound for the detection of fatty liver showed kappa statistics ranging from 0.54 to 0.92 for intrarater reliability and from 0.44 to 1.00 for interrater reliability. Sensitivity and specificity of ultrasound was similar to that of other imaging techniques

Copyright © 2011 by the American Association for the Study of Liver Diseases.

Address reprint requests to: Ruben Hernaez, M.D., Ph.D., Department of Medicine, The Johns School of Medicine, 2024 East Monument Street, Suite 2-600, Baltimore, MD, 21287. rhernae1@jhmi.edu; fax: 410-955-0476..

Potential conflict of interest: Nothing to report.

Additional Supporting Information may be found in the online version of this article.

(i.e., computed tomography or magnetic resonance imaging). Statistical heterogeneity was present even after stratification for multiple clinically relevant characteristics.

Conclusion—Ultrasonography allows for reliable and accurate detection of moderate-severe fatty liver, compared to histology. Because of its low cost, safety, and accessibility, ultrasound is likely the imaging technique of choice for screening for fatty liver in clinical and population settings.

Fatty liver is the accumulation of fat (i.e., macro-vesicular steatosis) within the hepatic parenchyma. Nonalcoholic fatty liver disease (NAFLD), the presence of fat infiltration in the liver in the absence of excessive alcohol consumption and other causes of liver disease, is the most common cause of fatty liver, with a prevalence as high as 30% in many populations.¹ NAFLD may lead to fibrosis,² cirrhosis,³ liver cancer,^{4,5} liver failure requiring liver transplant,⁶ and mortality⁷, and it is associated with type 2 diabetes, metabolic syndrome, and other cardiovascular risk factors.^{8,9} Although NAFLD represents a major public health challenge, its natural history and determinants are incompletely understood because of limitations in diagnostic technologies and because this condition is often asymptomatic until very late, severe complications occur. In addition, because of the risk of progression to more advanced stages, early noninvasive detection of fatty liver disease is clinically important.

Conventional B-mode ultrasonography is the most common technique used to assess the presence of fatty liver in clinical settings and population studies. However, several limitations of ultrasonography, including operator dependency, subjective evaluation, and limited ability to quantify the amount of fatty infiltration, have raised concerns. Indeed, some qualitative reviews^{10,11} have questioned the ability of ultrasound to reliably identify fatty liver, although no systematic review has performed a quantitative summary of available data on the diagnostic ability and reliability of ultrasound to identify fatty liver, compared to histology, the gold-standard.

The main aim of this meta-analysis was thus to systematically review and summarize the available literature on the diagnostic accuracy (i.e., sensitivity and specificity) and reliability of ultrasound to distinguish patients with and without fatty liver, defined as the presence of moderate to severe steatosis on liver biopsy (gold standard). As secondary aims, we sought to systematically review and summarize the diagnostic accuracy and reliability of different ultrasonographic parameters or criteria used to diagnose fatty liver (e.g., presence of liver-to-kidney contrast or scores summing a variety of parameters). And, finally, we planned to analyze the available literature on the diagnostic accuracy (i.e., sensitivity and specificity) of ultrasound to detect fatty liver, compared to other imaging techniques (i.e., magnetic resonance imaging [MRI] and computed tomography [CT]).

Patients and Methods

Data Sources and Search

Our search of PubMed and Embase included the term ultrasound and different combinations of fatty liver using free text and key words (Supporting Table 1). The period of the

electronic search extended from October 1967 through March 17, 2010, with no language restrictions. We also searched the reference lists of identified reviews and abstracted articles.

Study Selection

We included all studies that presented the following: (1) estimates of diagnostic accuracy (such as sensitivity or specificity), cross-tabulations, or correlations of B-mode ultrasonography to identify fatty liver against histology as the gold standard; (2) estimates of intra- or interrater reliability (such as kappa statistics or intraclass correlation coefficients) of ultrasound to identify fatty liver; and (3) comparisons of ultrasound to other imaging modalities (i.e., CT or MRI) to identify fatty liver.

We excluded studies that did not use ultrasound for evaluating fatty liver, studies that used ultrasound but did not study fatty liver (e.g., cirrhosis exclusively), and studies that evaluated ultrasound techniques not commonly used (e.g., Doppler, transient elastography contrast-enhanced ultrasound, artificial neural networks, or computer-aided readings, including histogram evaluation and fat quantification, using regions of interest). We also excluded studies using experimental conditions, studies performed in the operating room, studies performed in nonhumans, *in vitro* or *in vivo*, and articles that did not report original data (e.g., editorials, news, comments, guidelines, and reviews).

Data Extraction and Quality Assessment

Three investigators (R.H., M.L., and S.B.) independently reviewed the search results to determine article inclusion and perform data abstraction. Discrepancies were resolved by consensus. For each selected publication, we abstracted year of publication, country, inclusion criteria, histological definition of fatty liver (i.e., simple steatosis and steatohepatitis), number of participants undergoing ultrasound and comparison tests (if applicable), definitions of fatty liver used in the study, ultra-sonographic parameters evaluated, and reported measures of accuracy and reliability. For articles with no reported measure of accuracy, we estimated the sensitivity and specificity from the available data. We evaluated the quality of each article by applying modified Quality Assessment of Diagnostic Accuracy Studies (QUADAS)¹² and STAndards for the Reporting of Diagnostic accuracy studies (STARD) criteria.¹³

Study outcome was the presence of fatty liver as a dichotomous variable, using the specific criteria and definitions used in each study. For ultrasound, a few studies reported four categories, and we combined the normal/mild categories as absence of fatty liver, and the moderate/severe categories as presence of fatty liver. For histology, we used the presence of greater than or equal to 20%-30% fat infiltration to define fatty liver, except for Nagata et al. (< 10%), Guajardo-Salinas (>0%), and Soresi (>5%). We conducted secondary analyses on the diagnostic accuracy using lower levels of fat infiltration on histology as diagnostic criteria (i.e., <5%, 10%, and a 20%-30%).

Because number of ultrasonographic parameters have been used alone or in combination to diagnose fatty liver; if data were available, we evaluated the diagnostic accuracy of the following parameters: (1) parenchymal brightness, (2) liver-to-kidney contrast, (3) deep

beam attenuation, (4) bright vessel walls, and (5) gallbladder wall definition. Given that some studies reported or combined different histological findings, such as inflammation and fibrosis, we performed secondary analyses to study how accurate ultrasound was in identifying fatty infiltration with or without inflammation or fibrosis.

Data Synthesis and Analysis

Sensitivity and specificity of each study were summarized using the hierarchical summary receiver operating characteristics (ROC) curve approach.¹⁴ In this method, the relationship between logit-transformed sensitivity and specificity in each study is quantified by the log diagnostic odds ratio (OR) and the results are used to estimate a summary ROC curve.¹⁵ This method provides summary estimates of sensitivity and specificity, 95% confidence and prediction regions, and summary ROC curves, and it allows for multivariate analysis of between-study heterogeneity. Between-study heterogeneity was assessed by plots of the standardized logarithm of the diagnostic OR versus the inverse of the standard error and by the I^2 statistic, a parameter that describes the percentage of total variation across studies attributable to heterogeneity, rather than chance.¹⁶ We used clinically important variables to assess between-study heterogeneity and fit metaregression models. Publication bias was assessed visually using the effective sample size funnel plot and associated regression test of asymmetry.¹⁷ Statistical analyses were performed using the STATA commands, METANDI and MIDAS (StataCorp 2007, Stata Statistical Software, Release 10; StataCorp LP, College Station, TX).

Results

Our review included 49 studies of diagnostic accuracy comparing ultrasound to histology (Table 1; Supporting Fig. 1)¹⁸⁻⁶⁶ and five studies comparing ultra-sound to other radiological techniques (including three studies that reported three-way comparisons between ultrasonography, another imaging technique, and histology) (Table 2).⁶⁷⁻⁷¹ Nine of the 49 studies comparing ultrasound to histology also included data comparing each ultrasonographic parameter (e.g., liver-to-kidney contrast, deep beam attenuation, etc.) and histology.^{25,26,31,34,38,40,49,61,62} Finally, 22 studies provided data on intra- or interrater reliability (Supporting Table 2).^{S1-S22}

Meta-Analysis of Diagnostic Accuracy of Ultrasonography Versus Histology

Forty-nine studies, including 4720 participants, provided data on the diagnostic accuracy of ultrasound compared to histology as the gold standard. The weighted prevalence of histologically defined fatty liver across all studies was 31.8%, but the studies varied with respect to study population and location. Twenty-seven studies (55%) were conducted in a hospital setting or included a mixture of inpatients and outpatients. The indication for testing was suspicion of liver disease in 17 studies and known liver disease in 16 studies. The underlying liver disease was a combination of NAFLD and other pathologies in 36 studies and NAFLD only in eight studies. All studies included a representative spectrum of patients. Seventeen (35%) of the 49 studies did not report the method of ascertainment or used a different method of ascertainment in controls. Fewer than 50% of studies reported whether

the interpretation of the ultrasound had been done without knowledge of the results of the biopsy.

Overall sensitivity of ultrasound to detect moderate to severe histologically defined fatty liver from the absence of steatosis (n = 34 studies, 2815 participants) was 84.8% (95% confidence interval [CI]: 79.5-88.9), specificity was 93.6% (87.2-97.0), the positive likelihood ratio was 13.3 (6.4-27.6), the negative likelihood ratio was 0.16 (0.12-0.22), and the summary area under the ROC curve was 0.93 (0.91-0.95) (Figs. 1 and 2A). We further examined the lower cutoffs for the detection of histologically defined fat, and found that ultrasounds have a diagnostic accuracy for the detection of 10% of steatosis between 0.91 and 0.93 and specificity between 0.88 and 0.99 (Supporting Table 3).

Heterogeneity for the area under the summary ROC curve was substantial (I^2 , 98%; 95% CI: 97-99). In subgroup analyses, clinically relevant categories only explained a minor proportion of between-study heterogeneity (Supporting Fig. 2). There was no indication of publication or related biases (data not shown).

When ultrasound was used to differentiate the presence of histologically based fatty liver alone versus other pathological findings, such as hepatitis or fibrosis or normal liver (n = 29 studies), overall sensitivity was similar (87.2%; 95% CI: 77.8-93.0), but specificity was substantially lower (79.2%; 95% CI: 72.8-84.4). Correspondingly, the positive likelihood ratio was lower (4.2; 95% CI: 3.3-5.4), but the negative likelihood ratio was unchanged (0.16; 95% CI: 0.09-0.28). Overall, the summary area under the ROC curve was the same as that for determining fatty liver versus not (0.93; 95% CI: 0.91-0.95) (Fig. 2B).

Meta-Analysis of Diagnostic Accuracy of Ultrasonography Components Versus Histology

There was a wide variation in ultrasound parameters evaluated for assessing fatty liver (data not shown). Of the 49 studies with histology as a gold standard, parenchymal brightness was used as an ultrasound diagnostic criterion in 43 (88%) studies, deep beam attenuation in 30 (61%), vessels in 28 (57%), liver-to-kidney contrast in 27 (55%), and gallbladder wall definition in 4 (8%) studies.

In studies where the accuracy of ultrasonographic parameters of fatty liver definition were evaluated individually, sensitivities of liver to kidney contrast, vessel wall brightness, and deep beam attenuation were 98% (75%-100%), 81% (70%-89%), and 59% (45%-72%), respectively. Specificity was similar for all components (range, 93%-95%) (Supporting Table 4).

Systematic Review of the Reliability of Ultrasonography

Twenty-two studies reported the reliability of ultrasound findings: kappa statistics (17 studies), coefficients of variation (three studies), percent disagreement (one study), and intraclass correlation coefficient (one study).^{S1-S22} Among studies reporting kappa statistics, the number of readers ranged from 1 to 15. The range of kappa values for intrarater evaluation was 0.54-0.92 (six studies) and for the interrater evaluation was 0.44-1.00 (14 studies). Studies reporting reliability measures for individual components reported similar results across components (Supporting Table 5).^{S1-S22}

Meta-Analysis of Diagnostic Accuracy of Ultrasonography Versus Other Imaging Techniques

We found five studies comparing ultrasound data to CT, MRI, or magnetic resonance spectroscopy (MRS) without histology, including a total of 215 adults. Ultrasound had an overall sensitivity of 93.6% (60.5-99.3), specificity of 80.1% (53.3-93.4), positive likelihood ratio of 4.71 (1.89-11.71), and negative likelihood ratio of 0.08 (0.01-0.56). Only three studies had ultrasound,^{56,65,66} another imaging technique, and histology (Supporting Table 6),^{S23-S25} and ultrasound had slightly better overall accuracy for detecting fatty liver, compared to other techniques.

Discussion

Our meta-analysis shows that ultrasound is an accurate, reliable imaging technique for the detection of fatty liver, as compared with histology, with a pooled sensitivity of 84.8%, a pooled specificity of 93.6% for detecting 20%-30% steatosis, and a summary area under the ROC curve of 0.93. Because ultrasound is relatively inexpensive and accessible, compared to other diagnostic techniques, our results suggest that ultrasound may be the imaging technique of choice for screening for the presence of fatty liver in clinical settings and, especially, population studies. The widespread use of ultrasound to detect fatty liver may help better identify the determinants and natural history of fatty liver disease in the general population and may help target interventions directed to reducing the complications associated with fatty liver. Indeed, though no U.S. Food and Drug Administration–approved therapy exists for fatty liver, lifestyle changes,⁷² vitamin E, and pioglitazone⁷³ have shown some efficacy.

We found a relatively large number of studies using ultrasound as the diagnostic method and liver biopsy as a gold standard, with a wide range of sensitivities (55%-100%) and specificities (26%-100%). These differences could be the result of a number of factors. First, technical quality and performance of the ultrasound varied across studies. We included studies conducted from 1979 to 2010; during this time, technological advances in ultrasound equipment have occurred and could potentially explain part of this variation. Second, the ultrasound criteria used to define fatty liver differed across studies. Third, although the majority of the studies included patients who underwent liver biopsy with some suspicion of liver disease, there was a wide range in severity of the underlying disease. Finally, the composition of the comparison group (i.e., normal liver or other liver disease, such as inflammation, fibrosis, or a combination of these) also differed across studies, adding to the heterogeneity. Despite these differences, our sensitivity analyses, stratified by publication year, setting, degree of steatosis, and diagnosis, among others, showed similar results and, therefore, allow the use of the pooled accuracy estimates. Similar factors may have also contributed to the variation of reliability estimates between studies, including prevalence of cases with steatosis in the study population, lack of standard protocol to perform the evaluation, and the use of different criteria.

The potential role of ultrasound in clinical settings and in population research is very important. In the current obesity epidemic, the prevalence of fatty liver disease, in particular NAFLD, is likely to increase, making it necessary to use practical tools for measuring the

burden of disease and tracking time trends. In the clinical context, the number of patients at risk for fatty liver disease is also increasing. There is thus a pressing need to have readily available, accurate methods to assess the presence of fatty liver, and ultrasound compares favorably to alternative noninvasive techniques. Liver enzymes, indirect markers of liver injury, have lower sensitivity (0.30-0.63) and specificity (0.38-0.63) than ultrasound.⁷⁴ Indeed, compared to liver enzymes, the use of ultrasound as a triage test, applied early on to determine which patients should undergo further testing, would likely reduce the number of false-positive results and thus decrease the burden of subsequent testing. Other imaging techniques (i.e., CT or MRI/MRS) have similar operating characteristics, but are more expensive, and CT involves radiation, and therefore, their widespread usefulness is limited.

Our systematic review had certain limitations. We did not include other ultrasound techniques (e.g., Doppler and histogram) that would have allowed a more objective quantification of fat. Also, we could not assess the accuracy of ultrasound for the whole range of fat accumulation and could not evaluate the performance of an ultrasound-based four-grade scale (i.e., normal, intermediate, moderate, and severe) in the detection of fatty liver. We did not have individual patient data, so we were not able to evaluate the performance of ultrasound in key patient subgroups (e.g., by body mass index or presence of subcutaneous fat thickness). Although we reported significant statistical heterogeneity, in multiple secondary analyses on the key clinical variables, our inferences remained unchanged.

Our review shows that though ultrasound is useful for identifying fatty liver, additional research is needed to better assess the performance of specific ultrasound criteria of individual parameters, in particular gallbladder and vessel wall definition, to accurately and reliably detect fatty liver. Some parameters may be more reliable and justify the use of a more focused ultra-sound examination. In addition, future studies assessing the accuracy of ultrasound should aim to refine the ultrasound protocol and assess the accuracy of a scoring system to improve its reliability.

We also identified relatively few studies comparing the accuracy of ultrasound against other noninvasive imaging techniques and alternative testing strategies (e.g., including a combination of imaging and liver enzymes), and, therefore, could not conduct comparative analyses of different techniques and/or different testing strategies. Future studies are warranted to answer those questions, including about the accuracy of MRS extensively used in epidemiological studies, in which no large comparison with histology is available. However, the existing data do not provide evidence that other techniques are superior to ultrasound to detect the presence of fatty liver, although they may be useful in experimental settings, where a more precise quantification of liver fat is needed. Further studies are needed contrasting the diagnostic performance of ultrasound in persons with different degrees of adiposity to histology.

We also found that there is a need for improved study quality and reporting. Only five of the studies evaluating the diagnostic accuracy of ultrasound also assessed its reliability. These few studies support that ultrasound has a good intra- and interrater reliability and is comparable to the reliability of biopsy data,⁷⁵ but future studies using ultrasound should

include detailed reliability data. In addition, a number of studies did not provide details of the ultrasound protocol and/or assessment. This information is important to ensure proper replication and comparability between studies. Finally, few studies clearly reported the ultrasound reviewers masked to participant's characteristics and histological findings, and therefore, there was the risk of bias.

In conclusion, our meta-analysis shows that liver ultrasonography is an accurate, reliable tool to detect moderate to severe fatty liver, with sensitivity and specificity of 84.8% and 93.6%, respectively. These findings, together with the relatively low cost and lack of radiation exposure, support the use of ultrasound as the imaging technique of choice for screening for fatty liver in clinical settings and population studies. More research is needed to assess the long-term prognostic significance of ultrasound findings as well as the diagnostic implications of improvements in ultrasound technology and of more detailed quantification of liver fat.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by the American Diabetes Association Mentor-Based Postdoctoral Fellowship Program (7-07-MN-08; to R.H. and M.L.), National Institute of Diabetes and Digestive and Kidney Diseases grant 1RO1DK083393-01A1 (to J.M.C.), and K24-DK62222 P60 DK079637 (to F.L.B.).

Abbreviations

CI	confidence interval
CT	computed tomography
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
NAFLD	nonalcoholic fatty liver disease
N/R	not reported
OR	odds ratio
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
ROC	receiver operating characteristics
STARD	STAndards for the Reporting of Diagnostic accuracy studies

References

1. Lazo M, Clark JM. The epidemiology of nonalcoholic fatty liver disease: a global perspective. *Semin Liver Dis.* 2008; 28:339–350. [PubMed: 18956290]
2. Harrison SA, Torgerson S, Hayashi PH. The natural history of nonalcoholic fatty liver disease: a clinical histopathological study. *Am J Gastroenterol.* 2003; 98:2042–2047. [PubMed: 14499785]

3. Clark JM, Diehl AM. Nonalcoholic fatty liver disease: an underrecognized cause of cryptogenic cirrhosis. *JAMA*. 2003; 289:3000–3004. [PubMed: 12799409]
4. Ascha MS, Hanouneh IA, Lopez R, Tamini TA-R, Feldstein AF, Zein NN. The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *HEPATOLOGY*. 2010; 51:1972–1978. [PubMed: 20209604]
5. Bugianesi E, Leone N, Vanni E, Marchesini G, Brunello F, Carucci P, et al. Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. *Gastroenterology*. 2002; 123:134–140. [PubMed: 12105842]
6. Charlton M. Cirrhosis and liver failure in nonalcoholic fatty liver disease: molehill or mountain? *HEPATOLOGY*. 2008; 47:1431–1433. [PubMed: 18393323]
7. Adams LA, Lymp JF, St Sauver J, Sanderson SO, Lindor KD, Feldstein A, et al. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology*. 2005; 129:113–121. [PubMed: 16012941]
8. Marchesini G, Brizi M, Bianchi G, Tomassetti S, Bugianesi E, Lenzi M, et al. Nonalcoholic fatty liver disease: a feature of the metabolic syndrome. *Diabetes*. 2001; 50:1844–1850. [PubMed: 11473047]
9. Targher G, Day CP, Bonora E. Risk of cardiovascular disease in patients with nonalcoholic fatty liver disease. *N Engl J Med*. 2010; 363:1341–1350. [PubMed: 20879883]
10. Adams LA, Talwalkar JA. Diagnostic evaluation of nonalcoholic fatty liver disease. *J Clin Gastroenterol*. 2006; 40(Suppl 1):S34–S38. [PubMed: 16540765]
11. Wieckowska A, Feldstein AE. Diagnosis of nonalcoholic fatty liver disease: invasive versus noninvasive. *Semin Liver Dis*. 2008; 28:386–395. [PubMed: 18956295]
12. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol*. 2006; 6:9. [PubMed: 16519814]
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med*. 2003; 138:W1–W12. [PubMed: 12513067]
14. Harbord RM, Whiting P, Sterne JA, Egger M, Deeks JJ, Shang A, Bachmann LM. An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary. *J Clin Epidemiol*. 2008; 61:1095–1103. [PubMed: 19208372]
15. Gatsonis C, Paliwal P. Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. *AJR Am J Roentgenol*. 2006; 187:271–281. [PubMed: 16861527]
16. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002; 21:1539–1558. [PubMed: 12111919]
17. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. 2005; 58:882–893. [PubMed: 16085191]
18. Gosink BB, Lemon SK, Scheible W, Leopold GR. Accuracy of ultrasonography in diagnosis of hepatocellular disease. *AJR Am J Roentgenol*. 1979; 133:19–23. [PubMed: 110064]
19. Foster KJ, Dewbury KC, Griffith AH, Wright R. The accuracy of ultra-sound in the detection of fatty infiltration of the liver. *Br J Radiol*. 1980; 53:440–442. [PubMed: 7388276]
20. Youssef MM. Ultrasonography liver disease diagnosis. *Ann Clin Res*. 1980; 12(Suppl 26):1–47. [PubMed: 7469379]
21. Debongnie JC, Pauls C, Fievez M, Wubin E. Prospective evaluation of the diagnostic accuracy of liver ultrasonography. *Gut*. 1981; 22:130–135. [PubMed: 7215943]
22. Spuhler A, Pösl H, Sander R, Götz U. Ultrasonography in the diagnosis of fatty liver (author's transl) [in German]. *Leber Magen Darm*. 1981; 11:15–20. [PubMed: 6452560]
23. Pirovino M, Grauer W, Götz A, Huber M, Altorfer J, Maranta E, et al. Importance of ultrasonography in the diagnosis of diffuse parenchymal disorders of the liver [in German]. *Schweiz Med Wochenschr*. 1982; 112:525–526. [PubMed: 7079689]
24. Pamilo M, Sotaniemi EA, Suramo I, Lahde S, Arranto AJ. Evaluation of liver steatotic and fibrous content by computerized tomography and ultrasound. *Scand J Gastroenterol*. 1983; 18:743–747. [PubMed: 6669938]

25. Yajima Y, Ohta K, Narui T, Abe R, Suzuki H, Ohtsuki M. Ultrasono-graphical diagnosis of fatty liver: significance of the liver-kidney contrast. *Tohoku J Exp Med.* 1983; 139:43–50. [PubMed: 6220488]
26. Sanford NL, Walsh P, Matis C, Baddeley H, Powell LW. Is ultrasonography useful in the assessment of diffuse parenchymal liver disease? *Gastroenterology.* 1985; 89:186–191. [PubMed: 3891495]
27. Berrut C, Curati W, de Gautard R, Widmann JJ, Godin N, Loizeau E. The role of ultrasonography in the diagnosis of diffuse liver disease [in German]. *Schweiz Med Wochenschr.* 1986; 116:215–218. [PubMed: 3513304]
28. Cusumano S, De Mori E, Buschini M. Clinical statistics on the agreement of echographic and bioptic findings in diffuse liver diseases. *Minerva Med.* 1986; 77:889–892. [PubMed: 3523300]
29. Needleman L, Kurtz AB, Rifkin MD, Cooper HS, Pasto ME, Goldberg BB. Sonography of diffuse benign liver disease: accuracy of pattern recognition and grading. *AJR Am J Roentgenol.* 1986; 146:1011–1015. [PubMed: 3515875]
30. Saverymuttu SH, Joseph AE, Maxwell JD. Ultrasound scanning in the detection of hepatic fibrosis and steatosis. *Br Med J (Clin Res Ed).* 1986; 292:13–15.
31. Tam KM, Wu JS. Ultrasonographic diagnosis of fatty liver [in Chinese]. *Taiwan Yi Xue Hui Za Zhi.* 1986; 85:45–53. [PubMed: 3519848]
32. Coulson IH, McKenzie J, Neild VS, Joseph AE, Marsden RA. A comparison of liver ultrasound with liver biopsy histology in psoriatics receiving long-term methotrexate therapy. *Br J Dermatol.* 1987; 116:491–495. [PubMed: 3555595]
33. Forsberg L, Floren CH, Hederstrom E, Prytz H. Ultrasound examination in diffuse liver disease. Clinical significance of enlarged lymph nodes in the hepato-duodenal ligament. *Acta Radiol.* 1987; 28:281–284. [PubMed: 2958032]
34. Sato Y, Ohta Y, Fujiwara K, Oka H. Loss of echoes from the gallbladder wall in diagnosis of fatty infiltration of the liver. *Scand J Gastroenterol.* 1987; 22:1160–1164. [PubMed: 3324295]
35. Savarino V, Magnolia MR, Scalabrini P. Are sonographic and radionuclide investigations alternative or complementary in diagnosing liver disease? A comparison between these methods and laparobiopsy. *Ital J Gastroenterol.* 1987; 19:5–9.
36. Celle G, Savarino V, Picciotto A, Magnolia MR, Scalabrini P, Dodero M. Is hepatic ultrasonography a valid alternative tool to liver biopsy? Report on 507 cases studied with both techniques. *Dig Dis Sci.* 1988; 33:467–471. [PubMed: 3280274]
37. Lössner C, Cuno S, Kleine S, Kleine FD. Value of ultrasound tomography in the diagnosis and follow-up of fatty liver [in German]. *Dtsch Z Verdau Stoffwechselkr.* 1988; 48:22–26. [PubMed: 3286217]
38. Saitoh S, Nagamine T, Takagi H, Sekiguchi T, Uehara M, Yuasa K, et al. Diagnosis of fatty liver —comparison of ultrasonographic and computed tomographic findings to histological features [in Japanese]. *Nippon Shokakibyō Gakkai Zasshi.* 1988; 85:2658–2665. [PubMed: 3072429]
39. Yang PM, Huang GT, Lin JT, Sheu JC, Lai MY, Su IJ, et al. Ultrasonography in the diagnosis of benign diffuse parenchymal liver diseases: a prospective study. *Taiwan Yi Xue Hui Za Zhi.* 1988; 87:966–977. [PubMed: 3071569]
40. Ferrari F, Fagioli Zucchi A, Saloni E, Marini M, Frosini G. Variability of sonographic changes in diffuse hepatic diseases. *Ital J Gastroenterol.* 1989; 21:155–158.
41. Nishimura N. Non-invasive quantitative assessment of hepatic fatty change by the hepatic ultrasound velocity and attenuation coefficient. *Jpn J Med Ultrason.* 1989; 16:31–40.
42. Joseph AEA, Saverymuttu SH. Ultrasound in the assessment of diffuse parenchymal liver disease. *Clin Radiol.* 1991; 44:219–221. [PubMed: 1959295]
43. Bloom SL, Ireland A, Ryley NG, Chapman RW, Lindsell D. Correlation between hepatic ultrasound and histology. *Eur J Gastroenterol Hepatol.* 1992; 4:39–42.
44. Castellano G, González A, Colina F, Rodríguez S, Muñoz MT, Sánchez F, et al. Diagnostic value of the hepatic echo-histogram in chronic hepatopathy [in Spanish]. *Rev Esp Enferm Dig.* 1993; 84:373–380. [PubMed: 8129991]

45. Nagata O, Uto K, Sugihara H, Nohara Y, Mishima Y, Yoshida M, et al. Evaluation study of measurement method in ultrasonic diagnosis of fatty change of the liver. Based on trace histogram examination. *Jpn J Med Ultrason*. 1993; 20:9–16.
46. Cardi M, Muttillio IA, Amadori L, Petroni R, Mingazzini P, Barillari P, et al. Superiority of laparoscopy compared to ultrasonography in diagnosis of widespread liver diseases. *Dig Dis Sci*. 1997; 42:546–548. [PubMed: 9073137]
47. Dietrich CF, Wehrmann T, Zeuzem S, Braden B, Caspary WF, Lembcke B. Analysis of hepatic echo patterns in chronic hepatitis C [in German]. *Ultraschall Med*. 1999; 20:9–14. [PubMed: 10226340]
48. Graif M, Yanuka M, Baraz M, Blank A, Moshkovitz M, Kessler A, et al. Quantitative estimation of attenuation in ultrasound video images: correlation with histology in diffuse liver disease. *Invest Radiol*. 2000; 35:319–324. [PubMed: 10803673]
49. Mathiesen UL, Franzén LE, Aselius H, Resjö M, Jacobsson L, Foberg U, et al. Increased liver echogenicity at ultrasound examination reflects degree of steatosis but not of fibrosis in asymptomatic patients with mild/moderate abnormalities of liver transaminases. *Dig Liver Dis*. 2002; 34:516–522. [PubMed: 12236486]
50. Lakomy EA, Stalke P, Sikorska K, Pawlak A, Michalski Z, Witczak-Malinowska K, et al. Liver steatosis in chronic hepatitis C patients. *Ann Acad Med Gedanensis*. 2004; 34:183–189.
51. Hepburn MJ, Vos JA, Fillman EP, Lawitz EJ. The accuracy of the report of hepatic steatosis on ultrasonography in patients infected with hepatitis C in a clinical setting: a retrospective observational study. *BMC Gastroenterol*. 2005; 5:14. [PubMed: 15829009]
52. Kim SH, Lee JM, Kim JH, Kim KG, Han JK, Lee KH, et al. Appropriateness of a donor liver with respect to macrosteatosis: application of artificial neural networks to US images—initial experience. *Radiology*. 2005; 234:793–803. [PubMed: 15665225]
53. Palmentieri B, De Sio I, La Mura V, Masarone M, Vecchione R, Bruno S, et al. The role of bright liver echo pattern on ultrasound B-mode examination in the diagnosis of liver steatosis. *Dig Liver Dis*. 2006; 38:485–489. [PubMed: 16716779]
54. Riley TR III, Mendoza A, Bruno MA. Bedside ultrasound can predict nonalcoholic fatty liver disease in the hands of clinicians using a prototype image. *Dig Dis Sci*. 2006; 51:982–985. [PubMed: 16783524]
55. Hamaguchi M, Kojima T, Itoh Y, Harano Y, Fujii K, Nakajima T, et al. The severity of ultrasonographic findings in nonalcoholic fatty liver disease reflects the metabolic syndrome and visceral fat accumulation. *Am J Gastroenterol*. 2007; 102:2708–2715. [PubMed: 17894848]
56. Lee JY, Kim KM, Lee SG, Yu E, Lim YS, Lee HC, et al. Prevalence and risk factors of non-alcoholic fatty liver disease in potential living liver donors in Korea: a review of 589 consecutive liver biopsies in a single center. *J Hepatol*. 2007; 47:239–244. [PubMed: 17400323]
57. Perez NE, Siddiqui FA, Mutchnick MG, Dhar R, Tobi M, Ullah N, et al. Ultrasound diagnosis of fatty liver in patients with chronic liver disease: a retrospective observational study. *J Clin Gastroenterol*. 2007; 41:624–629. [PubMed: 17577120]
58. Salueña I, Ortega L, Devesa MJ, López-Alonso G, Taxonera C, Díaz-Rubio M, Ladero JM. Utility of liver biopsy in the etiologic diagnosis of biochemical liver abnormalities of unknown cause [in Spanish]. *Gastroenterol Hepatol*. 2007; 30:325–330. [PubMed: 17662214]
59. Chen CH, Lin ST, Yang CC, Yeh YH, Kuo CL, Nien CK. The accuracy of sonography in predicting steatosis and fibrosis in chronic hepatitis C. *Dig Dis Sci*. 2008; 53:1699–1706. [PubMed: 17939048]
60. de Moura AA, Cotrim HP, Barbosa DB, de Athayde LG, Santos AS, Bitencourt AG, et al. Fatty liver disease in severe obese patients: diagnostic value of abdominal ultrasound. *World J Gastroenterol*. 2008; 14:1415–1418. [PubMed: 18322958]
61. Ahmed L, Salama H, Ahmed R, Hamdy S, Al-Akel W, Shafi SA, et al. Non-invasive fibrosis seromarkers as a predictor of liver fibrosis in chronic hepatitis C and/or non-alcoholic steatohepatitis. *Arab J of Gastroenterol*. 2009; 10:14–20. [PubMed: 24842131]
62. Dasarathy S, Dasarathy J, Khyami A, Joseph R, Lopez R, McCullough AJ. Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study. *J Hepatol*. 2009; 51:1061–1067. [PubMed: 19846234]

63. Guajardo-Salinas GE, Hilmy A. Prevalence of nonalcoholic fatty liver disease (NAFLD) and utility of FIBROspect II to detect liver fibrosis in morbidly obese Hispano-American patients undergoing gastric bypass. *Obes Surg*. 2010; 20:1647–1653. [PubMed: 19957049]
64. Soresi M, Giannitrapani L, Florena AM, La Spada E, Di Gesaro V, Rappa F, et al. Reliability of the bright liver echo pattern in diagnosing steatosis in patients with cryptogenic and HCV-related hypertransaminasaemia. *Clin Radiol*. 2009; 64:1181–1187. [PubMed: 19913127]
65. Yamashiki N, Sugawara Y, Tamura S, Kaneko J, Matsui Y, Togashi J, et al. Noninvasive estimation of hepatic steatosis in living liver donors: usefulness of visceral fat area measurement. *Transplantation*. 2009; 88:575–581. [PubMed: 19696642]
66. Lee SS, Park SH, Kim HJ, Kim SY, Kim MY, Kim DY, et al. Noninvasive assessment of hepatic steatosis: prospective comparison of the accuracy of imaging examinations. *J Hepatol*. 2010; 52:579–585. [PubMed: 20185194]
67. Scatarige JC, Scott WW, Donovan PJ, Siegleman SS, Sanders RC. Fatty infiltration of the liver: ultrasonographic and computed tomographic correlation. *J Ultrasound Med*. 1984; 3:9–14. [PubMed: 6694259]
68. Pacifico L, Celestre M, Anania C, Paolantonio P, Chiesa C, Laghi A. MRI and ultrasound for hepatic fat quantification: relationships to clinical and metabolic characteristics of pediatric nonalcoholic fatty liver disease. *Acta Paediatr*. 2007; 96:542–547. [PubMed: 17306008]
69. Pozzato C, Radaelli G, Dall'Asta C, Verduci E, Villa A, Villa C, et al. MRI in identifying hepatic steatosis in obese children and relation to ultrasonography and metabolic findings. *J Pediatr Gastroenterol Nutr*. 2008; 47:493–499. [PubMed: 18852643]
70. Edens MA, van Ooijen PM, Post WJ, Haagmans MJF, Kristanto W, Sijens PE, et al. Ultrasonography to quantify hepatic fat content: validation by (1)h magnetic resonance spectroscopy. *Obesity (Silver Spring)*. 2009; 17:2239–2244. [PubMed: 19461588]
71. Mancini M, Prinster A, Annuzzi G, Liuzzi R, Giacco R, Medaqli C, et al. Sonographic hepatic-renal ratio as indicator of hepatic steatosis: comparison with (1)H magnetic resonance spectroscopy. *Metabolism*. 2009; 58:1724–1730. [PubMed: 19716568]
72. Lazo M, Solga SF, Horska A, Bonekamp S, Diehl AM, Brancati FL, et al. The effect of a 12-month intensive lifestyle intervention on hepatic steatosis in adults with type 2 diabetes. *Diabetes Care*. 2010; 33:2156–2163. [PubMed: 20664019]
73. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steato-hepatitis. *N Engl J Med*. 2010; 362:1675–1685. [PubMed: 20427778]
74. Neuschwander-Tetri BA, Clark JM, Bass NM, Van Natta ML, Unalp-Arida A, Tonascia J, et al. Clinical, laboratory, and histological associations in adults with nonalcoholic fatty liver disease. *HEPATOLOGY*. 2010; 52:913–924. [PubMed: 20648476]
75. Juluri R, Vuppalanchi R, Olson J, Unalp A, Van Natta ML, Cummings OW, et al. Generalizability of the nonalcoholic steatohepatitis clinical research network histologic scoring system for nonalcoholic fatty liver disease. *J Clin Gastroenterol*. 2011; 45:55–58. [PubMed: 20505526]

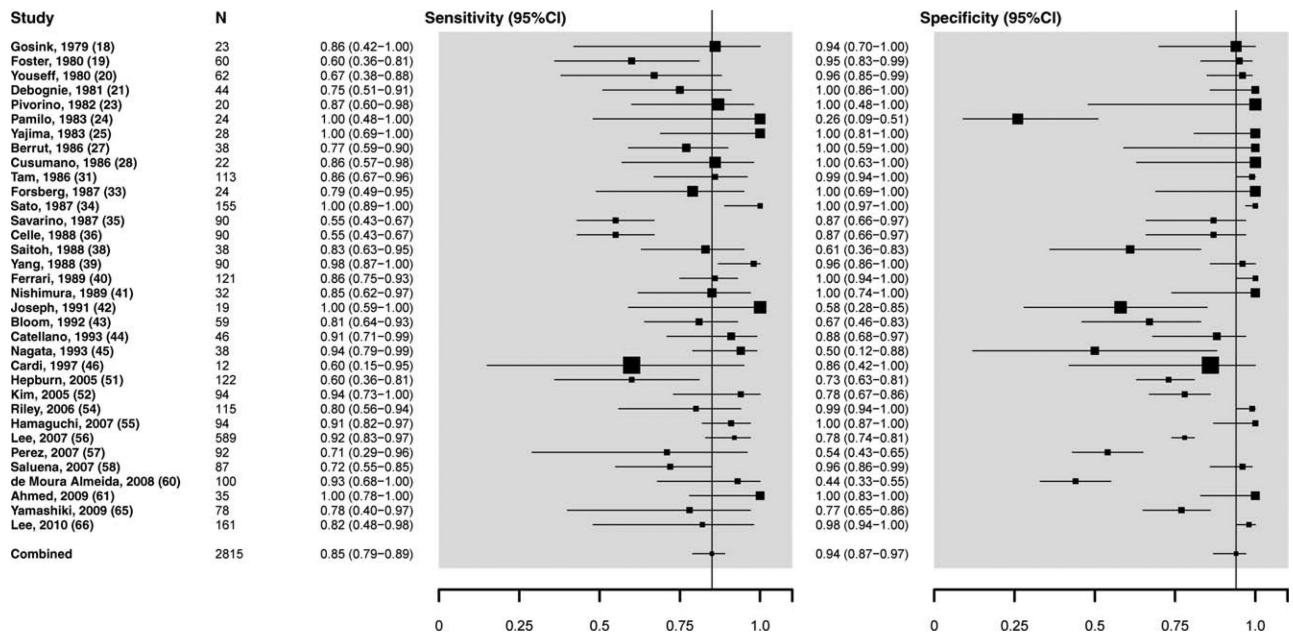


Fig. 1. Overall sensitivity and specificity of ultrasound to detect moderate-to-severe histologically defined fatty liver from the absence of steatosis.

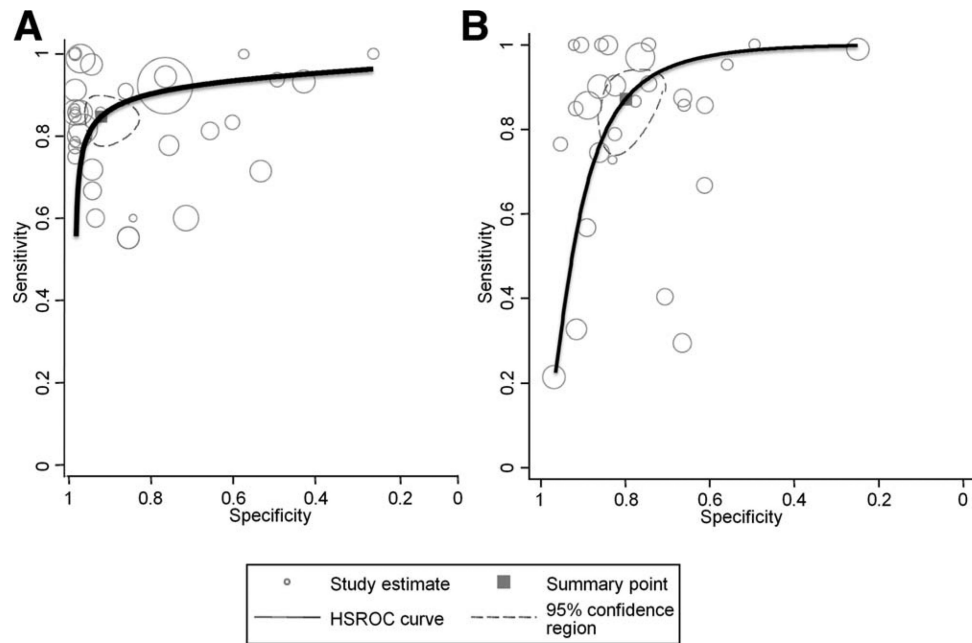


Fig. 2. Summary receiver-operating characteristic (ROC) curve plots showing test accuracy of ultrasound compared to histology to distinguish between presence versus absence of steatosis (A), and presence of steatosis versus everything else (B).

Table 1

Characteristics of the 44 Studies of Diagnostic Accuracy Comparing Ultrasound to histology, sorted by publication year^(*)

Author, year (reference)	Country	Setting	Indication Ultrasound/Standard	Liver Disease	N
Gosink, 1979(18)	USA	Hospital	Suspicion liver disease	Mixed	23
Foster, 1980(19)	UK	Hospital	Suspicion liver disease	Mixed	60
Youssef, 1980(20)	Finland	Mixed (inpatient/outpatient)	Mixed	Mixed	62
Debongnie, 1981(21)	Belgium	N/R	Suspicion liver disease	Mixed	44
Spuhler, 1981(22)	Germany	Hospital	N/R	Mixed	310
Pirovino, 1982 (23)	Switzerland	Hospital	Known liver disease	Mixed	20
Pamilo, 1983 (24)	Finland	N/R	Suspicion liver disease	Mixed	24
Yajima, 1983 (25) *	Japan	N/R	Known liver disease	Mixed	28
Sanford, 1985(26) *	Australia	Hospital	Known liver disease	Mixed	125
Berrut, 1986 (27)	Switzerland	Hospital	Suspicion liver disease	N/R	38
Cusumano, 1986 (28)	Italy	N/R	Suspicion liver disease	Mixed	22
Needleman, 1986 (29)	USA	Hospital	Known liver disease	Mixed	96
Saverymuttu, 1986 (30)	UK	Outpatient clinic	Suspicion liver disease	Mixed	90
Tam, 1986 (31) *	Taiwan	Hospital	N/R	Mixed	113
Coulson, 1987 (32)	UK	Outpatient clinic	Other	Mixed	49
Forsberg, 1987 (33)	Sweden	N/R	Mixed	Mixed	24
Sato, 1987 (34) *	Japan	Mixed (inpatient/outpatient)	Known liver disease	Mixed	155
Savarino, 1987 (35)	Italy	Mixed (inpatient/outpatient)	Known liver disease	Mixed	90
Celle, 1988 (36)	Italy	Hospital	Known liver disease	Mixed	90
Lossner, 1988 (37)	Germany	Hospital	Suspicion liver disease	Mixed	187
Saitoh, 1988 (38) *	Japan	Hospital	N/R	Mixed	38
Yang, 1988 (39)	Taiwan	Hospital	Known liver disease	Mixed	90
Ferrari, 1989 (40) *	Italy	N/R	Known liver disease	Mixed	121
Nishimura, 1989 (41)	Japan	Mixed (inpatient/outpatient)	Known liver disease	Mixed	32
Joseph, 1991 (42)	UK	Outpatient clinic	Suspicion liver disease	Mixed	19
Bloom, 1992 (43)	UK	Mixed (inpatient/outpatient)	Known liver disease	Mixed	59
Castellano, 1993 (44)	Spain	Mixed (inpatient/outpatient)	Mixed	Mixed	46
Nagata, 1993 (45)	Japan	Hospital	N/R	Mixed	38
Cardi, 1997 (46)	Italy	N/R	Suspicion liver disease	Mixed	12
Kim, 2005 (52)	Korea	Hospital	Other	NAFLD	94
Palmentieri, 2006 (53)	Italy	N/R	Suspicion liver disease	Mixed	208
Riley, 2006 (54)	USA	Outpatient clinic	Known liver disease	Mixed	115
Hamaguchi, 2007 (55)	Japan	Hospital	Suspicion liver disease	NAFLD	94
Lee, 2007(56)	Korea	Hospital	Other	NAFLD	589
Perez, 2007 (57)	USA	Hospital	Known liver disease	Mixed	92
Saluena, 2007 (58)	Spain	N/R	Suspicion liver disease	NAFLD	87

Author, year (reference)	Country	Setting	Indication Ultrasound/Standard	Liver Disease	N
Chen, 2008 (59)	Taiwan	N/R	Known liver disease	No NAFLD	108
De Moura Almeida, 2008 (60)	Brazil	Hospital	Suspicion liver disease	NAFLD	100
Ahmed, 2009 (61) *	Egypt	Outpatient clinic	Known liver disease	Mixed	35
Dasarathy, 2009 (62) *	USA	Outpatient clinic	Suspicion liver disease	Mixed	73
Guajardo-Salinas, 2009 (63)	USA	Outpatient clinic	Other	NAFLD	102
Soresi, 2009 (64)	Italy	N/R	Mixed	Mixed	150
Yamashiki, 2009 (65)	Japan	Hospital	Health screening	NAFLD	78
Lee, 2010 (66)	Korea	Outpatient clinic	Health screening	NAFLD	161

* n = These studies provided data of the accuracy of individual ultrasound parameters compared to histology.

Table 2

Characteristics of the Five Studies of Diagnostic Accuracy Comparing Ultrasound to Another Imaging Technique, Sorted by Publication Year

Author, year (reference)	Country	Setting	Indication Ultrasound	Liver Disease	Standard	N
Scatarige, 1984 (67)	USA	Mixed (inpatient/outpatient)	Known liver disease	Mixed	CT	94
Pacifico, 2007 (68)	Italy	N/R	Suspicion liver disease	NAFLD	MRI	100
Pozzato, 2008 (69)	Italy	Hospital	N/R	NAFLD	MRI	60
Edens, 2009 (70)	Netherland	General population	Other	NAFLD	MRS	18
Mancini, 2009 (71)	Italy	Outpatient clinic	Other	NAFLD	MRS	40