Original Research

What proportions of focal liver lesions detected by unenhanced ultrasound are inconclusive?

Iain Willits[1], Julie Burn[1], Helen Cole[1], Tim Hoare[1] and Andrew Sims[1,2]

[1]The Newcastle upon Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Newcastle upon Tyne, UK
[2]Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK
Corresponding author: Andrew Sims. Email: andrew.sims@nuth.nhs.uk

Abstract

In August 2012, the National Institute for Health and Care Excellence produced positive diagnostics guidance on the ultrasound contrast agent SonoVue®, but recommended further research involving an estimation of the proportion of unenhanced ultrasound scans reporting, but not characterising, focal liver lesions, particularly in cirrhotic livers. Patient records from the Radiology Information System of an acute hospital trust were progressively filtered based on categorical fields and keywords in the free text reports, to obtain ultrasound records including the liver that were appropriate for manual analysis. In total, 21,731 records referred from general practice or out-patient clinics were analysed. Patients described as having cirrhosis were analysed as a subgroup. After automatic exclusion of records considered likely to be negative, 5812 records were manually read and categorised as focal liver lesion inconclusive, benign or malignant. In the general practice cohort of 9175 records, 746 reported the presence of one or more focal liver lesions, with 18.4% (95% CI 15.7% to 21.3%) of these records mentioning an inconclusive focal liver lesion. In the out-patient cohort of 12,556 records, 1437 reported one or more focal liver lesions, and 29.4% (95% CI 26.9% to 32.0%) of these were inconclusive. Cirrhosis was reported in 10.8% of the out-patient scans that also reported a focal liver lesion, and 47.4% (95% CI 39.3% to 55.6%) of these scans had an inconclusive focal liver lesion, compared with 27.3% (95% CI 24.9% to 29.8%) that were inconclusive in non-cirrhotic livers (odds ratio 2.4; 95% CI 1.7 to 3.4). This retrospective study indicates that unenhanced ultrasound scans, in which a focal liver lesion is detected, are frequently inconclusive, with the probability of an inconclusive scan being greater in out-patient than general practice referrals. Inconclusive focal liver lesions were also reported in greater proportions of cirrhotic than non-cirrhotic livers. The results of this research will inform future updates of National Institute for Health and Care Excellence diagnostics guidance.

Keywords: Liver, ultrasound, SonoVue®, inconclusive, cirrhosis, National Institute for Health and Care Excellence

Introduction

Ultrasound examinations of the abdomen routinely interrogate the liver to identify abnormal morphology or hepatic pathology. Occasionally, ultrasound of the liver may reveal the presence of a focal liver lesion (FLL), which can be either malignant or benign. However, it is not always possible for unenhanced ultrasound to correctly identify the nature of an FLL, and further characterization may be necessary to determine the appropriate patient management. Typically, FLLs that require further characterization are investigated using contrast-enhanced ultrasound (CEUS), contrast-enhanced computed tomography (CECT), or contrast-enhanced magnetic resonance imaging (CEMRI).[1]

In August 2012, a contrast agent, SonoVue® sulphur hexafluoride microbubbles (Bracco Imaging S.p.A), for CEUS, was assessed and given a positive recommendation by the National Institute for Health and Care Excellence (NICE) for use with liver imaging (NICE diagnostics guidance number 5).[2] The evidence used to formulate the recommendations was derived from a systematic review and an associated cost utility model, which incorporated three economic scenarios to assess the cost-effectiveness of SonoVue®. That work has since been published as a Health Technology Assessment (HTA).[3] The scenarios were: (i) detection of metastases following diagnosis of colorectal cancer; (ii) incidental detection of FLLs in people with non-specific indications; and (iii) cirrhosis surveillance, which incorporated an adaptation of a previously published model.[4] Cirrhosis surveillance has been shown to be positively associated with increased survival rates by a recent

systematic review and meta-analysis of observational studies.[5]

An important parameter used specifically in the cirrhosis surveillance model was the proportion of 'inconclusive' unenhanced ultrasound scans; NICE defined an inconclusive scan as 'an unenhanced ultrasound scan in which a FLL is detected, but not characterised'. The manufacturer of SonoVue® provided an estimate that 43% of ultrasound scans were inconclusive, and therefore require further diagnostic investigation. This estimate aligned with a previous study by Strobel et al.,[6] which compared the diagnostic accuracy of unenhanced Doppler ultrasound with Doppler ultrasound with CEUS (contrast medium used not SonoVue®) in patients with known FLLs and with a high suspicion of malignancy (cirrhosis status not reported). However, there was a considerable degree of uncertainty regarding the 43% estimate as applied to UK clinical practice, because of the methodology employed, the technology used and the population recruited in the study. A different estimate of this parameter could substantially affect the cost-effectiveness estimates of using SonoVue® in the cirrhosis cohort using existing economic models.[3]

To address this uncertainty, NICE recommended that further research should be undertaken on 'the percentage of unenhanced ultrasound scans that are inconclusive, particularly in people with cirrhosis'.[2] To achieve this, NICE commissioned further research through their research facilitation framework, using an external assessment centre. The assessment centre is free to undertake the research using the most appropriate methodology, but the work is required to be practical and rapid, in order to quickly inform updates of the guidance. This paper reports the work undertaken by the assessment centre (the authors of this manuscript).

The aim of this study was to answer NICE's research question to estimate the probability that an unenhanced abdominal ultrasound scan report mentions an uncharacterized FLL, given that one or more FLLs are detected during the scan. We define this as the proportion of inconclusive scans to match the NICE definition of 'inconclusive'.[2] The setting of the study was a retrospective review of Radiology Information System (RIS) records of patients from a hospital trust who received a general ultrasound examination of the abdomen or a specific scan of the liver. Additionally, the cirrhosis status of patients, in whom an FLL was reported, was recorded, to allow for subgroup analysis to directly answer NICE's research question.

Materials and methods
Data source and extraction of records

The study was placed on the Newcastle upon Tyne Hospitals NHS Foundation Trust (NUTH) clinical effectiveness register as an audit study. Full ethical approval was not necessary because the primary purpose of the study was to retrospectively measure service provision without using identifiable patient information. An extract from the RIS was provided in spreadsheet format by a radiology data manager. Using categorical fields (appointment date, examination and source of referral), potentially relevant records were extracted for further analysis. Each record included

these three categorical fields, a pseudonymised patient identifier (unique to each patient) and a free text report field. Records were included if the appointment date was between 1 April 2009 and 31 March 2012, the examination name contained the terms 'liver' or 'abdo' (abdomen or abdominal) and the scan modality was ultrasound.

Two distinct cohorts of patients were identified and included in the study; these were patients referred directly from their General Practitioner (GP cohort) and patients referred from secondary care as out-patients (OP cohort). The GP cohort was considered to be more likely to represent patients in whom an incidental detection of an FLL might be made; that is they would be predominantly naïve to abdominal ultrasound and other imaging modalities. Those referred as out-patients, however, may have received previous abdominal imaging, with a prior clinical report available to the ultrasound operator, or may be undergoing periodic screening for primary liver cancer (e.g. on a surveillance programme following development of cirrhosis or infection with hepatitis B or C) or metastatic disease (following diagnosis of an extra-hepatic solid malignancy). Records from in-patient and accident and emergency referrals were not retrieved, as these were considered to be likely to represent a particularly complex case-mix. Other than these restrictions, all patient reports were considered potentially relevant, without further exclusions.

Analysis and classification of records

The process used to analyse and classify records is shown in Figure 1. There were three stages: (i) exclusion of records that were out of scope; (ii) identification of records that had negative detection of FLLs and (iii) classification of the remaining records in which FLL was detected. Due to the large number of records, the first two steps were automated using scripts written in the R programming language.[7]

Exclusion of records that were out of scope

Using categorical fields, records of scans were removed as follows: those which were duplicates; those for which the study name field did not indicate an ultrasound scan of the abdomen or liver; those not within the specified date range; and those not in the GP or OP cohorts (identified by the source of referral field). Using the free text report field, records were excluded if the text did not make reference to the liver or a related adjective (i.e. did not contain any words starting with 'liver', 'lver', 'livr', 'hepat' or 'lob'). Records remaining after this first step were considered to be ones that were likely to have been actively scanned for the detection of FLLs of the liver. Quality assurance of this filtering stage was conducted by one researcher manually reading 200 records randomly selected from the exclusions.

Automatic identification of negative findings

Firstly, records were excluded if they did not contain a relevant diagnostic, pathological, or related word (Table 1). To aid the choice of these terms, a random sample of 200 records was analysed to identify standard text strings or stored report templates used in the reporting of results.
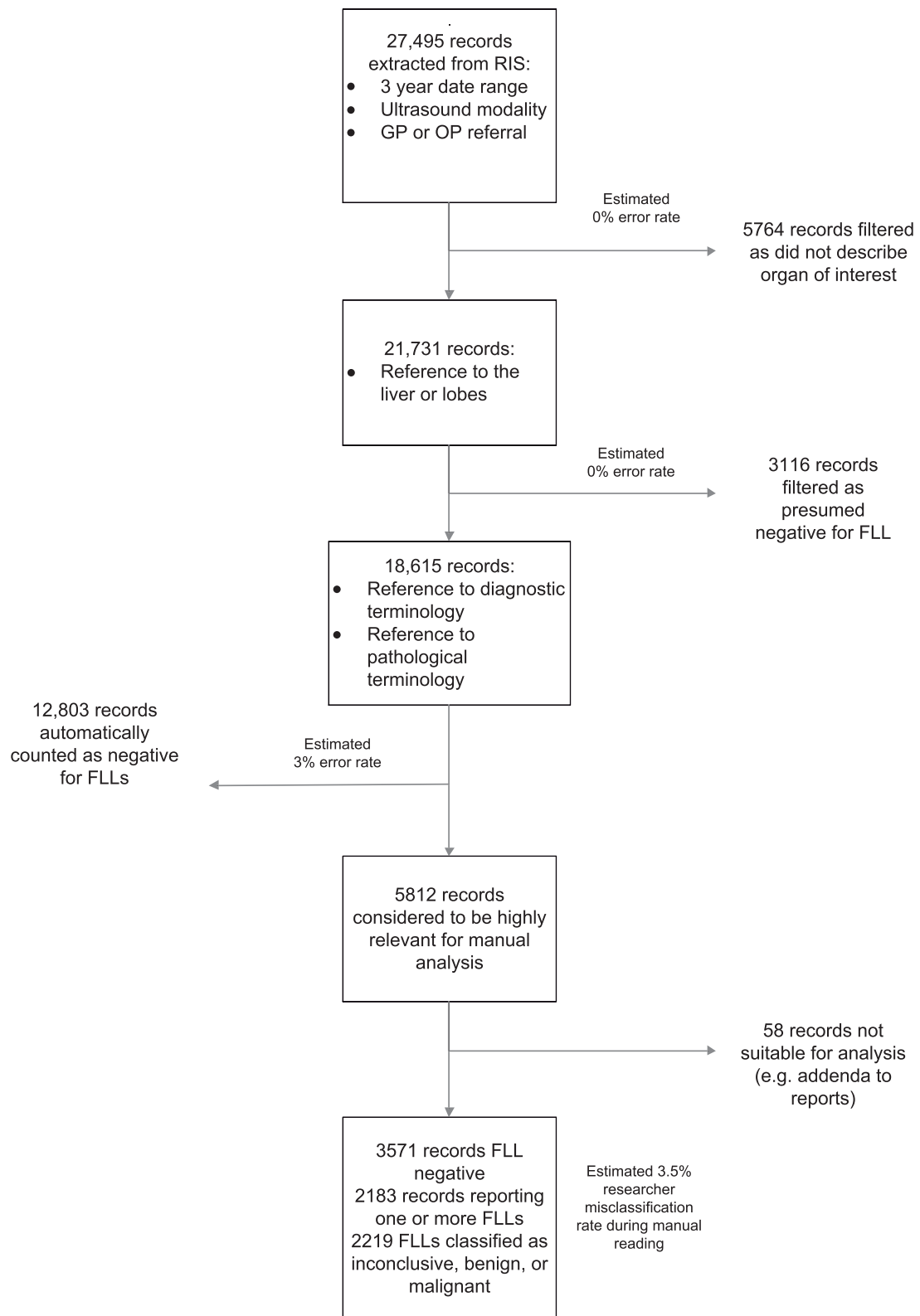
Figure 1 Flow chart illustrating progressive exclusion, automatic read and manual read of RIS reports during analysis of ultrasound records

In addition, selection of terms was informed by a survey that was sent out to local sonographers. Quality assurance of this filtering stage was conducted by one researcher manually reading 200 random records excluded from this stage. Secondly, records were automatically counted as 'FLL not detected' if they contained words or phrases in the free text that were suggestive of the record being negative for FLLs (Table 2). Quality assurance of this step was conducted by one researcher manually reading a random sample of 400 of the records counted as 'FLL not detected'.

Table 1 Terms used to indicate possible FLL

Term
'character', 'malig', 'benign', 'staging', 'CEUS', 'sonovue', 'contrast', 'imaging'
'HCC', 'carcinoma', 'cyst', 'hyperplasia', 'FNH', 'adenoma', 'focal', 'fcal', 'foal'
'FLL', 'lesion', 'lsion', 'leision', 'leison', 'leson', 'haemang', 'hemang', 'pathology'
'metast', 'abnormality', 'mass', 'parenchyma', 'disease', 'change', 'inconc'
'indetermin', 'abscess', 'sparing', 'infiltration', 'neoplasm', \\<CT*, 'MRI')

*\\<CT represents 'CT' at the beginning of a word, so was intended to capture a CT scan
Records containing none of these terms were identified as negative for FLL

Table 2 Terms used to identify records as negative for FLLs (i.e. no FLLs detected). Two character spaces were allowed between all words

Number	Term
1	'No focal lesion'
2	'No liver lesion'
3	'No focal liver lesion'
4	'No hepatic focal lesion'
5	'No focal hepatic lesion'
6	'liver [is] normal in size, contour and reflectivity'
7	'Liver and spleen appear normal in size, contour and reflectivity'
8	'No intrahepatic biliary duct dilatation or focal lesions seen'
9	'No intrahepatic biliary duct dilatation or focal liver lesion'
10	'No evidence of focal liver lesions or intrahepatic biliary duct dilatation'
11	'No suspicious focal [hepatic/liver] lesion'
12	'no evidence of focal liver lesions'
13	'no evidence of focal abnormality'
14	'liver appears normal in size, contour and reflectivity'
15	'liver and spleen appear normal in size, contour and reflectivity'
16	'Normal ultrasound appearances of the liver'
17	'Normal appearance[s] of [the] liver'
18	'liver, [list of other organs*] normal'
19	'Liver of normal size, contour and reflectivity'
20	'No abnormality [is] seen in [the] liver'
21	'No focal hepatic abnormalit[y or ies]'
22	'Normal liver size, contour and texture with no evidence of focal'
23	'No focal hepatic change'
24	'Normal size and ultrasound appearances of [the] liver'
25	'No gross focal lesion[s] or biliary dilatation'
26	'Normal liver,'[†]
27	'Normal appearance[s] of [the] liver'[†]

*For example 'liver, spleen, pancreas, gallbladder, biliary tree, aorta and both kidneys normal'
[†]Report started with this phrase
Phrases in square brackets were optional

Manual analysis and classification

Manual analysis consisted of one researcher reading the remaining reports and classifying them into one of the following categories: FLL not detected; FLL detected but not characterised (i.e. inconclusive or indeterminate); FLL detected and characterised as malignant (metastatic or primary); and FLL detected and characterised as benign. Cysts, focal fatty sparing, focal fatty infiltration, focal nodal hyperplasia and haemangiomas were all individually categorised and collectively classified as benign. Patients who were identified as having cirrhosis were manually marked in the data set, to allow for subsequent subgroup analysis. A random sample of 200 report fields was extracted and independently read by an experienced consultant radiologist as a further quality assurance step.

Statistical analysis

Confidence intervals of simultaneous multinomial proportions were calculated using the method of Sison and Glaz[8] with 95% significance level. Associations between the presence of cirrhosis and having an inconclusive scan were tested with Fisher's Exact Test.

Results

Analysis and classification of records

A total of 504,362 potentially relevant records were retrieved from the RIS. After removing duplicates, 27,495 records were extracted which were in the date range of interest, were for ultrasound scans and were for patients referred by their GP or from out-patient (OP) clinics. From analysis of the free text report fields of these 27,495 records (Figure 1), 5764 were excluded by the application of a text filter because they did not contain a word relating to the organ of interest. A sample of 200 records excluded by this filter was manually reviewed and none were found to have been excluded incorrectly. This left 21,731 records of relevance to this study (9175 in the GP cohort and 12,556 in the OP cohort).

A total of 3116 records were identified as negative using a text filter (Table 1), i.e. there was no mention of liver pathology. A sample of 200 records, chosen at random from excluded records, was reviewed manually and none were found to have been excluded incorrectly. This left 18,615 records which potentially mentioned a FLL.

A total of 12,803 records were identified as having negative findings of FLLs, using automated text analysis (Table 2). From a sample of 400 records identified as being negative for FLL, 12 (3%) were considered to have been inappropriately excluded by the automatic algorithm. All of these sampled records, upon manual reading for quality assurance, were found to contain text indicating a benign FLL was present, but also contained text indicating a negative finding in another organ (e.g. no focal lesion of the kidney). However, this discrepancy was accepted as a reasonable error rate for this step.

This left a total of 5812 records that required manual review of the report field to determine whether an FLL was detected and then characterised. On comparing the

Table 3 Totals and proportions of inconclusive, malignant and benign ultrasound scans reporting FLLs from GP and OP referred cohorts

Cohort		Inconclusive FLL	Malignant FLL	Benign FLL*	Total
GP	(n)	137	11	598	746
	(% FLLs)	18.4 (15.7–21.3)	1.5 (0.0–4.4)	80.2 (77.5–83.1)	100.0
OP	(n)	423	107	907	1437
	(% FLLs)	29.4 (26.9 to 32.1)	7.4 (4.9 to 10.1)	63.1 (60.6–65.7)	100.0

*Including cysts
Proportions are expressed as percentages (95% confidence interval)

manual read by researcher and experienced radiologist, and following arbitration on discordant pairs, it was found that 193/200 records were correctly classified by the researcher, giving an error rate of less than 4% for this step. From the misclassification error rates estimated from the three samples of records used for quality assurance purposes, the estimated total number of misclassified reports was 587 (2.7% of the 21,731 records studied).

Proportion of inconclusive scans

In the GP group, there were 9175 records of ultrasound scans with reports which referred to the liver or lobes and, of these, 746 records reported the presence of an FLL (8.1%; 95% CI 7.6% to 8.7%). The proportions of inconclusive, malignant and benign FLL scans are reported in Table 3, with 18.4% (95% CI 15.7% to 21.3%) of scans being classified as inconclusive. Of the benign FLLs, 252 were cysts; if these were not classified as FLLs, the proportion of inconclusive scans was 27.7% (95% CI 23.7% to 31.8%). Only 11 malignant lesions were positively identified in records from the GP cohort, with 10 of these being diagnosed as metastases.

In the OP group, there were 12,556 records of ultrasound scans with reports which referred to the liver or lobes, with 1437 reports identifying an FLL (11.4%, 95% CI 10.9% to 12.0%). The proportion of inconclusive, malignant and benign FLL scans are reported in Table 3, with 29.4% (95% CI 26.9% to 32.1%) being inconclusive. Of the benign FLLs, 460 were cysts; if these were not classified as FLLs, the proportion of inconclusive scans was 43.3% (95% CI 40.0% to 46.7%). A total of 107 scans were classified as malignant in the OP group accounting for 7.4% (95% CI 4.9% to 10.1%). Of these, 91 (85%) were classified as metastatic.

A direct comparison of the GP and OP reports was not made, because the study population scanned were considered to be too heterogeneous for meaningful evaluation.

Sub-group analysis of cirrhotic livers

From the GP cohort, only two of the 746 reports in which an FLL was identified also mentioned the presence of cirrhosis. In the OP cohort, 154 reports of the 1437 in which an FLL was identified also mentioned cirrhosis (10.8%; 95% CI 9.2% to 12.5%). The numbers of reports that mentioned FLLs in the OP cohort, grouped by scan finding, are shown in Table 4. Of the patients who had an identified FLL and also

Table 4 Numbers of reports that mentioned FLLs in the out-patient cohort, grouped by scan finding (FLL conclusive or inconclusive) and cirrhotic status

Scan finding	Have cirrhosis	Do not have cirrhosis	Total
Inconclusive	73	350	423
Conclusive (benign or malignant)	81	933	1014
Total	154	1283	1437

had a diagnosis of cirrhosis, 47.4% (95% CI 39.3% to 55.6%) of scans were reported to be inconclusive. In patients who had an FLL identified but cirrhosis was not mentioned, 27.3% (95% CI 24.9% to 29.8%) were classified as inconclusive. The probability of having an inconclusive scan was significantly associated with the presence of cirrhosis, with an odds ratio of 2.4 (95% CI 1.7% to 3.4%).

Discussion
Main study findings

Inconclusive test results are a frequent, but often overlooked, feature of many diagnostic investigations.[9] In this study, we reported the results of a large retrospective study which had the primary aim of understanding what proportion of unenhanced ultrasound liver scans, in which an FLL was detected, were inconclusive. This study was designed and implemented to inform a future update of the NICE diagnostic guidance (DG5) of SonoVue® sulphur hexafluoride microbubbles ultrasound contrast agent,[2] which had originally estimated that 43% of FLLs were inconclusive in a cirrhosis surveillance cohort, aligned with the study by Strobel et al.[6] To the best of our knowledge, this is the first time a large retrospective study has been used to quantify the proportion of inconclusive scans in which an FLL was detected by ultrasound of the liver. The methodology developed to measure this outcome has been extensively quality assured and appears robust, with an estimated combined accuracy rate of automatic filtering and manual analysis of around 97%.

In our study, around 18% of scans of livers with identified FLLs in the GP cohort and around 29% in the OP cohort were described as inconclusive by the operator. These figures included counting cysts as benign lesions.[10] However, although ultrasound is often considered to be the modality of choice for detecting liver cysts, cystic lesions may not always be interpreted as an FLL in practice,[10,11] which was verified by our survey of sonographers. Using an

alternative assumption that cysts are not classified by sonographers as FLLs, the proportion of inconclusive scans was 28% (GP) and 43% (OP), respectively. The latter figure is close to the estimate used in the HTA to inform NICE guidance.[3] However, although the proportion of inconclusive scans was the same in this case, the relative proportions of benign and malignant scans were different, again illustrating the uncertainties concerning lesion definition and the population the analysis is drawn from.

Whilst there is very little information in the literature concerning the proportion of FLLs that are inconclusive in unenhanced ultrasound scans, we did identify a randomized controlled trial by Trillaud et al.,[12] in which 134 patients with a previously detected FLL had the lesion characterised as malignant, benign, or indeterminate using unenhanced ultrasound, CEUS, CECT, CEMRI, and/or histopathology. The final diagnosis (gold standard) was adjudicated using a combination of imaging and histopathological results, and clinical information. The proportion of inconclusive FLLs was specifically recorded for all modalities, including baseline ultrasonography, and this was measured as 60.6% (estimated 95% CI of 52.1% to 69.1%), which is higher than observed in the present study. The reasons for these differing results are not known, but speculatively are likely to be related to the population selection and the definition of FLL. For instance, although cysts were recorded as a category of benign FLL in this study, none of the patients were reported as having one in the RCT.

An important factor likely to influence the proportion of inconclusive scans is the population of patients that are scanned. In our study, we included patients referred from GP and OP cohorts, and found that they were different, with GP referrals less likely to yield inconclusive results. The reasons for this are likely to be because the case-mix was different in each group, with some patients in the OP group being on liver surveillance programmes, and likely to have been scanned several times (this was a retrospective study of individual scans, not patients). Scans resulting from GP referrals on the other hand are more likely to be de novo referrals and be more representative of true incidental findings of FLLs and their characterisation.

One aspect of the case-mix that was investigated was the presence of cirrhosis, which was much more prevalent in the OP cohort, as expected (i.e. the GP cohort showed better parenchymal health). Sub-analysis of this group showed the FLLs were more than twice as likely (OR 2.4) to be classified as inconclusive in cirrhotic livers as in non-cirrhotic livers. This result is likely to reflect the known features of cirrhosis, which modifies the liver parenchymal architecture, and causes distortion of ultrasound attenuation.[1] Additionally, unenhanced ultrasound may not be capable of differentiating hepatocellular carcinoma from other benign lesions,[13] particularly regenerative nodules.[14] The use of CEUS to characterise FLLs in patients with cirrhosis is thus a particularly useful application of the technology.[13]

Advantages and limitations of study design

A retrospective review of radiology reports was selected as the most appropriate study design to quantify the proportion of inconclusive unenhanced ultrasound scans, because this was primarily a measurement rather than a comparative study, and more complex study designs would not be necessary. As there was no comparator, confounding was not an issue. This study type has several advantages compared with other observational studies, not least because the data already exists, so it can be accessed with little or no ethical concerns, as long as the data remains anonymised. An alternative approach would be a prospective study, but this carries the risk of cognitive bias and inducing an operator-related Hawthorne effect.[15] Additionally, this study had the advantage of being relatively quick to perform, which was a requirement of NICE who have need of rapid research facilitation in order to inform updates of guidance. As the database is fully inclusive of all NHS patients who receive ultrasound in a single defined hospital trust, there is also little risk of selection bias contributing to this measurement. Thus, it could be considered that this was a pragmatic study reflecting real-life practice, and the methodology employed in this study could be used on data from other hospitals and for other conditions.

In this case, the RIS database contains a wealth of information and more detailed analysis, or follow-up analysis, remains an option. For instance, a future study could follow up individual patients in RIS to identify misdiagnoses (e.g. lesions identified as metastases later found to be haemangioma), and therefore estimate the diagnostic accuracy of unenhanced ultrasound.

Nevertheless, there are also clear limitations to this methodology, and the risk of several types of bias.[16] As there were too many primary records to feasibly read and categorise manually, a method of automatically analysing report fields was devised to exclude reports that were out of scope and to identify reports suggesting that no FLL had been detected. Hence the only relevant records which were manually read were those which could not be eliminated automatically (Figure 1). However, quality assurance indicated that approximately 3% of records (which were all benign lesions) were inappropriately counted as 'FLL not detected', which means the proportion of inconclusive scans, and proportion of benign lesions in particular, in the whole dataset was likely to have been underestimated by this amount. Additionally, although patient inclusion was universal, due to limited access to all data fields (to preserve anonymity), it was not possible to stratify patients to clinical conditions or other characteristics of interest, or to investigate pre-specified factors not recorded by RIS. There was also potential interpretation bias on behalf of the researcher. For instance, often, on reading RIS reports, there was clearly a degree of subjectivity when classifying the categorical data, although the quality assurance undertaken indicated that this was unlikely to be a significant problem.

Conclusion

The results of this retrospective study indicate that, depending on the study population and the definition of FLL, between 18 and 43% of FLLs that are detected are not

diagnostically characterised by operators using unenhanced ultrasound. Higher proportions of inconclusive scans were observed in scans taken from people referred from OP departments compared with GP referrals, and higher proportions in those with cirrhosis. These groups in particular will routinely need additional imaging technologies to provide a definitive diagnosis, and CEUS is likely to be a useful option for this.

DECLARATIONS

ACKNOWLEDGEMENTS

REFERENCES

1. Claudon M, Dietrich CF, Choi BI, et al. Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver–update 2012: a WFUMB-EFSUMB initiative in cooperation with representatives of AFSUMB, AIUM, ASUM, FLAUS and ICUS. *Ultraschall in Med* 2013;**34**:11–29

2. National Institute for Health and Clinical Excellence. *SonoVue (sulphur hexafluoride microbubbles) – contrast agent for contrast enhanced ultrasound imaging of the liver.* London: NICE, 2012

3. Westwood M, Joore M, Grutters J, et al. Contrast-enhanced ultrasound using SonoVue (sulphur hexafluoride microbubbles) compared with contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging for the characterisation of focal liver lesions and detection of liver metastases: a systematic review and cost-effectiveness analysis. *Health Tech Assess* 2013;**17**:1–243

4. Thompson Coon J, Rogers G, Hewson P, et al. Surveillance of cirrhosis for hepatocellular carcinoma: systematic review and economic analysis. *Health Tech Assess* 2007;**11**:1–206

5. Singal AG, Pillai A, Tiro J. Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis. *PLoS Med* 2014;**11**:e1001624

6. Strobel D, Hoefer A, Martus P, Hahn EG, Becker D. Dynamic contrast-enhanced power Doppler sonography improves the differential diagnosis of liver lesions. *Intl J Colorectal Dis* 2001;**16**:247–56

7. R Core Team. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2013. See http://www.R-project.org/ (last checked 23 November 2014)

8. Sison C, Glaz J. Simultaneous confidence intervals and sample size determination for multinomial proportions. *J Am Stat Assoc* 1995;**90**:366–9

9. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. *BMJ* 2013;**346**:f2778

10. Alobaidi M, Shirkhoda A. Benign focal liver lesions: discrimination from malignant mimickers. *Curr Probl Diagn Radiol* 2004;**33**:239–53

11. Mortele KJ, Ros PR. Cystic focal liver lesions in the adult: differential CT and MR imaging features. *Radiographics* 2001;**21**:895–910

12. Trillaud H, Bruel JM, Valette PJ, et al. Characterization of focal liver lesions with SonoVue-enhanced sonography: international multicenter-study in comparison to CT and MRI. *World J Gastroenterol* 2009;**15**:3748–56

13. Wong GL, Xu HX, Xie XY. Detection of focal liver lesions in cirrhotic liver using contrast-enhanced ultrasound. *World J Radiol* 2009;**1**:25–36

14. Rapaccini GL, Pompili M, Caturelli E, et al. Focal ultrasound lesions in liver cirrhosis diagnosed as regenerating nodules by fine-needle biopsy. Follow-up of 12 cases. *Dig Dis Sci* 1990;**35**:422–7

15. Kennedy MT, Ong JC, Mitra A, Harty JA, Reidy D, Dolan M. The use of weekly departmental review of all orthopaedic intra-operative radiographs in order to improve quality, due to standardized peer expectations and the ''Hawthorne effect''. *Surgeon* 2013;**11**:10–3

16. Kelly S, Berry E, Roderick P, et al. The identification of bias in studies of the diagnostic performance of imaging modalities. *Br J Radiol* 1997;**70**:1028–35