1.12.2011

Master of Public Health

Dissertation

Contrast-enhanced ultrasound versus contrast-enhanced CT for diagnosing liver metastases

MPH student

Radiographer
Martin Lund
Billeddiagnostisk afdeling, Regionshospitalet Randers
Skovlyvej 1
8930 Randers NØ
Danmark

Dissertation supervisors

1.
Christian Gluud, M.D., Dr. Med. Sci., Associate Professor
Head of Department,
Copenhagen Trial Unit,
Centre for Clinical Intervention Research,
Department 3344,
Rigshospitalet,
Copenhagen University Hospital,
Blegdamsvej 9, DK 2100 Copenhagen,
Denmark.

2.
Susanne Rosthøj, Post doc
Københavns Universitet
Det Sundhedsvidenskabelige Fakultet
Institut for Folkesundhedsvidenskab
Biostatistisk Afdeling
Øster Farimagsgade 5, opg. B
1014 København K
Danmark

Index

Preface

The content of this paper will in the nearest future be part of a Cochrane diagnostic test accuracy review (DTAR).

I would like to thank the Copenhagen Trial Unit for helping me in the process of writing this paper. First of all I would like to thank Christian Gluud for supervising me extremely well all the way. I would like to thank Dimitrinka Nikolova for her support and coordination of the writing process, and I would like to thank Sarah Louise Klingenberg for helping me establish the search strategy.

I would also like to thank Susanne Rosthøj for her help concerning the statistical calculations and the assessment of the results.

Last but not least I would like to thank Thomas Abramovitz Bjerre for helping me in the process of selecting the included studies.

Martin Lund

Abstract

Background

Diagnosing liver metastases using imaging modalities like ultrasound (US) or computed tomography (CT) is an important step, revealing the progression of cancer diseases, i.e. revealing the presence of metastases or not. These tests can be performed as contrast-enhanced ultrasound (CEUS) and contrast-enhanced computed tomography (CECT). The latter is by many considered the mainstay of detecting, diagnosing, and follow-up of patients with suspected liver metastases, or verified liver metastases. However, there is no consensus worldwide on these matters.

Objective

To determine the diagnostic accuracy of CEUS, in comparison with CECT, for detecting suspected liver metastases in patients.

Search methods

Searches have been conducted in *The Cochrane Hepato-Biliary Group Diagnostic Test Accuracy Studies Register, The Cochrane Library* (*Database of Abstracts of Reviews of Effects* (DARE), *Health Technology Assessment* (HTA), and *NHS Economic Evaluation Database* (NHSEED)), *MEDLINE* (OvidSP), *EMBASE* (OvidSP), *Science Citation Index Expanded* (ISI Web of Knowledge), and *ACP Journal Club* (EBSCO host) until the 11th of November 2011. No language, document type or animal type restrictions have been applied.

Selection criteria

The studies in this review are studies based on direct comparison of CEUS and CECT, in the detection and diagnosis of liver metastases, i.e., head-to-head comparisons.

Data collection and analysis

The titles and the abstracts to every study have been screened and the studies have been selected individually by two persons. The included studies have been assessed for methodological quality using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) items. The statistically analyses have been based on the sensitivity and specificity of every included study.

Results

3 studies were included, and those studies were tested with the exact McNemar test (marginal homogeneity). They showed homogeneity between CEUS and CECT concerning sensitivity ($p = 1.0$ in all 3 studies). Concerning the calculation of the specificity, it is not possible to test for marginal homogeneity in 1 of the 3 studies because both CEUS and CECT have 100 % specificity

(0/94 in both groups). In the remaining 2 studies no heterogeneity was found ($p >= 0.250$ and $p >= 0.0625$, respectively).

Conclusions

The results show no significant difference between CEUS and CECT in diagnostic accuracy for detecting suspected liver metastases in patients. For this reason CEUS may be considered just as valuable as CECT when it comes to detecting and diagnosing liver metastases. However, the results are based on only 3 included studies. Therefore, it is recommended to perform further diagnostic accuracy tests in this area.

Background

The International Commission on Radiological Protection (ICRP) emphasises three key principles of radiological protection, which are justification, optimisation and application of dose limits (ICRP 2007a).

- The principle of justification: any decision that alters the radiation exposure situation should do more good than harm.
- The principle of optimization of protection: the likelihood of incurring exposure, the number of people exposed, and the magnitude of their individual doses should all be kept as low as reasonably achievable, taking into account economic and societal factors.
- The principle of application of dose limits: the total dose to any individual from regulated sources in planned exposure situations other than medical exposure of patients should not exceed the appropriate limits specified by the ICRP Commission (ICRP 2007a).

Computed tomography (CT) is a diagnostic modality, which uses X-rays to develop the images that are interpreted by the radiologists. X-rays are potentially carcinogenic. It is, therefore, important that any clinician and radiologist consider if the benefits of a CT scan are higher than the risks, before they are referring a patient to a CT scan. In order to keep the X-ray doses as low as reasonable achievable, it is also necessary to check out the possibilities of using non-ionizing alternative diagnostic modalities, such as magnetic resonance imaging (MRI) or ultrasound (US) (ICRP 2007c).

The risk of being exposed to X-rays can be divided into two general categories: deterministic effects and stochastic effects (ICRP 2007b). The induction of deterministic effects leads to harmful tissue reactions, and is in generally characterized by a threshold dose. A dose above the threshold dose increases the severity of the injury, and impairs the capacity for tissue recovery. Stochastic effects in exposed individuals are involving cancer and heritable effects, due to mutation of somatic cells and to mutation of reproductive cells (germ cells), or both. There is evidence of cancer risk at doses about 100 milliSievert (mSv) or less, although with uncertainties. In the case of heritable diseases there is no direct evidence of radiation risk to humans, but experimental observations argue that this risk to future generations should be considered in the system of protection (ICRP 2007b).

When the question about dose is discussed, one will have to distinguish between absorbed dose in the tissue and effective dose in the tissue. Absorbed dose in the tissue is the energy deposited in the tissue or an organ per unit mass, and is measured in Gray (Gy). This is the basic quantity used for assessing the radiation risk to a tissue or to an organ. The effective dose is a calculated quantity, which takes into account the difference in radiosensitivity of different organs. Effective dose is measured in Sievert (Sv), and is used as an index, to compare relative radiation risk from different radiological procedures (ICRP 2000). Absorbed doses in tissues from CT are some of the highest in diagnostic radiology, and are in the range of 10 to 100 mGy. Because of the tendency to repeat the CT scans, the doses can often reach levels that may increase the cancer risk. In comparison to

conventional X-ray procedures, the effective doses from CT scans are many times higher. A single conventional X-ray procedure of the chest has an effective dose of approximately 0.02 mSv, and a CT scan of the chest has an effective dose of approximately 8 mSv, i.e., 400 times more. A CT scan of chest, abdomen and pelvis has an effective dose of approximately 28 mSv (ICRP 2000).

There are at the moment four different modalities, which are used to detect and diagnose liver metastases: contrast-enhanced computed tomography (CECT), contrast-enhanced ultrasound (CEUS), positron emission tomography/computed tomography (PET/CT) and magnetic resonance imaging (MRI).

CECT is by many considered the mainstay of detecting, diagnosing and follow-up of patients with suspected liver metastases, or with verified liver metastases. The advantages of CECT are that it is a very quick examination, the images are ready to interpretation in only few seconds/minutes and the images are covering not only the liver, but also all the surrounding tissues. It is possible to characterize focal liver lesions with patterns of enhancement, due to the use of contrast agents. The images are stored electronically and are available to any physician, who has an interest in looking at them. The disadvantages are the use of relatively high doses of ionizing radiation to the patients and the use of contrast agents with iodine, which are known to have certain high risk adverse effects, such as allergic reactions and in the worst case anaphylactic shock. Therefore, one of the contraindications to CECT is patients with former allergic reactions to contrast agents with iodine. Another contraindication is patients with insufficient kidney function because of the risk of kidney failure.

CEUS is mostly used to support or verify findings on CT. The advantages of CEUS are that it is performed without the use of ionizing radiation and the examination is therefore considered to be of no harm to the human body. It is a relatively quick examination, which can be performed in about half an hour. Like CECT it is possible to characterize focal liver lesions with patterns of enhancement, due to the use of contrast agents. The contrast agents consist of micro bubbles (sulphur hexafluoride), and they are not known to cause any serious adverse effects. If necessary, it is possible to perform a biopsy of suspected lesions in the liver during the examination. The disadvantages are that the value of CEUS is very much dependent on the physician, who performs the scan, and even if it is possible to store the images this is not an assurance that other physicians are able to interpret the images.

PET/CT is like CEUS mostly used to support or verify findings on CT. The advantage of PET/CT is that it provides information of glucose uptake and metabolism of malignant cells and not only anatomic alterations like visible lesions. The disadvantages are the same as for CECT. Furthermore the patients are given fluro-18-deoxyglucose (FDG) in dosages, which has an effective dose at approximately 6-7 mSv. It is a very slow examination. To complete a full PET/CT one has to use around 3 hours all in all.

MRI is like CEUS and PET/CT mostly used to support or to verify findings on CT. The advantages of MRI are that it like CEUS is performed without the use of ionizing radiation, and the

examination is therefore considered to be of no harm to the human body. There are two different kinds of contrast agents, which can be used to detect liver metastases on MRI: gadolinium (Gd) and ferucarbotran (SuperParamagnetic Iron Oxide, SPIO). These contrast agents are not known to have serious adverse effects if they are used in small doses. However, patients with insufficient kidney function are known to be in high risk of adverse effects like nephrogenic systemic fibrosis if they are exposed to MRI gadolinium agents, especially if the gadolinium agents are used in high doses. The disadvantages are besides the just mentioned that MRI is a relatively slow examination. Approximately one hour is used to do a liver scan. Contraindications to a MRI examination are patients suffering from claustrophobia and patients with metal implants in the head or eyes.

Target condition being diagnosed

The target condition of this review is liver metastases. Liver metastases may originate from virtually any primary cancer. The most common primary cites are colon, stomach, pancreas, breast and lung. Colorectal cancer is the third leading cause of cancer death in the United States and Europe (Townsend 2009), and nearly a third, of all patients with colorectal cancer, gets metastases spread to the liver within five years of the diagnosis of colorectal cancer (Gurusamy 2010). Several Cochrane reviews of interventions have studied different ways of treatments of liver metastases that originate from colorectal cancer. Most of these treatments are not fully investigated or cannot jet be recommended (Nelson 2006; Al-asfoor 2008; Gurusamy 2009; Gurusamy 2009a; Townsend 2009; Gurusamy 2010).

Index test

The index test for this review is contrast-enhanced ultrasound (CEUS). As already stated US is not based on the use of X-rays, and US is not known to have any other adverse effects apart from those that arise from misdiagnosis. Therefore, if CEUS could be a diagnostic substitution to CECT, it will be avoided to expose the patients to ionizing radiation, and to the possible adverse effects from iodine contrast agents.

Alternative tests

CECT is the most commonly used diagnostic modality for detecting liver metastases. Another possibility is to perform a PET/CT scan (positron emission tomography). PET/CT generates images depicting the distributions of positron-emitting nuclides in patients. In order to get information about the exact location of lesions, there is a CT system incorporated in the PET system. A series of CT images is acquired over the same section of the patient as the PET scan, and the two scans are combined. MRI is also a possibility for diagnosing liver metastases. With the use of an external uniform magnetic field, all protons in the body of the patient are pointing in the same direction. By the use of radiofrequencies the protons are excited. They produce signals, when they forced by the

magnetic field return to relaxation. These signals are recorded and produce the MRI-images. See the background section for more information about the modalities.

Rationale

In accordance with the key principles of the ICRP about justification and optimization, it is an obligation to do a systematic review like the present. It is necessary to keep the radiation to patients as low as reasonably achievable, but at the same time it is very important to offer the patients the best possible modality to reveal the presence of liver metastases or not. The intension is to determine if CEUS may replace CECT in the detection and diagnosis of liver metastases, on the mentioned conditions.

As stated earlier, CECT is by many considered the mainstay of detecting, diagnosing and follow-up of patients with liver metastases. However, there is no consensus worldwide on this matter. Because of the advantages, which are already mentioned, CECT is the first choice in most patients. But why not take PET/CT and MRI into consideration as well? The answer to this question is that those modalities are not cut out to function as mainstay modalities in detecting and diagnosing liver metastases. They are both very slow modalities, PET/CT entails even higher ionizing radiation than CT, and MRI involves too many contraindications. This is the reason why CEUS is chosen as the index test and CECT as the comparator test in this review.

Objective

To determine the diagnostic accuracy of CEUS, in comparison with CECT, for detecting suspected liver metastases in patients.

Methods

Criteria for considering studies for this review

Types of studies

The studies for this review will be studies based on direct comparison of the index test and the comparator test, i.e., head-to-head comparisons. This is the strongest design, especially if it is a fully paired, direct comparison. Studies with not fully paired designs should be randomized, direct comparisons, where study participants are randomly allocated to receive either the index test or the comparator test (Bossuyt 2008).

All test results must be verified by the same reference standard in order to avoid differential verification bias. At the same time it is important that all test results, and not only a fraction of the test results, are verified by the reference standard, in order to avoid partial verification bias. It is

also very important to be aware of the fact that the reference standard, the comparator test and the index test cannot in anyway be the same because this leads to incorporation bias (Reitsma 2009). All studies that mix up the reference standard, the comparator test and the index test will be excluded.

If both the index test and the comparator test give negative results in the same patient, the patient should be subjected to adequate follow-up, for at least 3 months, as the reference standard for this group of patients.

It is important that the studies have a cross-sectional design to avoid disease progression bias. This means that the maximum acceptable delay between the execution of the index test and the comparator test has to be two weeks.

The reference standard, the index test and the comparator test should be performed blinded to each other. Since the dichotomous values diseased and not diseased are the results of all three tests, test review bias and diagnostic review bias are very likely to occur if the tests are not blinded to each other (Reitsma 2009).

However, blinding can have the side effect that clinical data, which is normally available in practice when tests are made and interpreted, will not be available due to the blinding. Preferable these data should be available, but not at the expense of the blinding.

Participants

All patients suspected with liver metastases at the primary diagnosing, or suspected in some other way will be included. Some studies may have restrictions on the inclusion of participants, and these restrictions will be evaluated in every study to reveal the possibilities of bias.

Index test

The index test for this review is CEUS.

Comparator test

The comparator test for this review is CECT.

Target condition

The target condition is liver metastases. Liver metastases are common in connection with e.g. colorectal and gastric cancers, and it is standardized to clarify the presence of liver metastases when a patient has been diagnosed with, e.g., colorectal cancer. The staging of cancer is important in order to define prognosis, and to plan the optimal treatment strategy. CECT is often the modality of first choice, when it has to be determined if a patient has liver metastases or not. Most of these

patients will be CECT scanned several times in order to follow the progress of the cancer disease. This means that in order to cure one cancer disease, the diagnostic procedures may potentially induce another cancer disease.

Reference standards

The reference standard should be biopsy test results, pathological examinations of surgically removed specimens, intra-operative ultrasound (IOUS), follow up, or a combination of all four. Preferable the reference standard should consist of laparotomy including palpation, IOUS and a biopsy/pathological examination of surgically removed specimen. This approach is considered to be the one with the highest sensitivity. The participants in the same study should all be evaluated by the same reference standard in order to avoid differential verification bias. The only exception from this will be participants who will only receive the follow-up for the minimum of 3 months, as the reference standard, because both the index test and the comparator test have given negative results.

Search methods for identification of studies

The search strategy for relevant studies was discussed and established in co-operation with The Cochrane Hepato-Biliary Group.

Electronic searches

Searches was conducted until the 11th of November 2011, in *The Cochrane Hepato-Biliary Group Diagnostic Test Accuracy Studies Register* (Gluud 2010)*, The Cochrane Library* (*Database of Abstracts of Reviews of Effects* (DARE), *Health Technology Assessment* (HTA), and *NHS Economic Evaluation Database* (NHSEED)), *MEDLINE* (OvidSP), *EMBASE* (OvidSP), *Science Citation Index Expanded* (ISI Web of Knowledge), and *ACP Journal Club* (EBSCO host)(Royle 2003). No language, document type or animal type restrictions were applied. Search strategies with the time span of the searches are given in Appendix 1.

Data collection and analysis

The guidelines provided in the draft *Cochrane Diagnostic Reviewer's Handbook* (de Vet 2008, Reitsma 2009), was followed.

Selection of studies

Two persons assessed the same literature searches, and selected the relevant studies in accordance with the chosen inclusion criteria and the quality tool for the review. If this generated any disagreements about which studies to select for the review, then each disagreement was discussed separately, and the disagreement was solved by consensus.

Two persons blinded to each other screened the titles and the abstracts to every study, and selected the relevant ones individually. All the studies selected in this first reading were provided in full text for further assessment. These studies were read in full text by the same two persons, and each one individually picked out the relevant studies. Then they got together and in consensus chose the studies for inclusion.

Data extraction and management

Two persons independently completed a data extraction form on all included studies. The following data was retrieved:

1. General information: title, journal, year, publication status, and study design (prospective versus retrospective).
2. Sample size: number of participants meeting the criteria and total number diagnosed or scanned or referred to.
3. Baseline characteristics: baseline diagnosis, age, sex, race and disease severity (metastases or not).
4. The index test and the comparator test.
5. Reference standard.
6. Number of true positive (TP), true negative (TN), false positive (FP), false negative (FN).

Missing data

Primary authors were contacted for missing, or unclear data.

Assessment of methodological quality

The tool for assessment of methodological quality of the studies was the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (Whiting 2003; Reitsma 2009), (Appendix 2). 11 items of the original 14 items in the QUADAS tool were used in the assessment of the methodological quality. The 3 excluded QUADAS items relate to the quality of reporting rather than methodology (Reitsma 2009). Two persons assessed the relevant studies according to the QUADAS items. Appendix 2 is a list of the 11 QAUDAS items and contains definitions on when to answer yes and when to answer no to the questions within the QUADAS items (Appendix 2). When to answer unclear is not defined in appendix 2, but this definition should be used in those cases when the QUADAS item cannot be answered with a clear yes or a clear no. Disagreements between the two investigators were resolved by consensus.

Statistical analysis and data synthesis

The most important estimates in the studies, and in the review, are sensitivity and specificity of the tests. From sensitivity and specificity the positive predictive value and negative predictive value can be established, along with the likelihood ratios and the diagnostic odds ratios. The data from the selected studies were tabulated and graphically displayed in coupled forest plots, on a trial by trial

level. The sensitivity and specificity, the predictive values, the likelihood ratios and the diagnostic odds ratios were determined and specified with a 95 % confidence interval (CI). All analyses and plots were done using SAS version 9.2.

CECT and CEUS are both modalities depending on enhancement patterns concerning the diagnosis of liver metastases, or any focal liver lesion. This is why the diagnosis is not established by a certain threshold or any cut-off point. For this reason, and because only studies based on direct comparison are considered, the bivariate model was chosen as the statistical model for the meta-analysis (Reitsma 2005; Macaskill 2010). The studies have been plotted on a summary receiver operator characteristic diagram (SROC). By the establishment of a 95 % confidence region around the summary operating point, it is possible to illustrate the extent of statistical heterogeneity. Covariates can then be added the statistical model in order to do a meta-regression to investigate the reasons for heterogeneity. If only a few studies are included this is not an issue though, but if more than 15 studies are included then this approach, doing a meta-regression, can be considered.

The data found in the included studies were statistically investigated on a patient by patient level. This gave an overview concerning the ability of CEUS and CECT to detect patients with liver metastases, and provided knowledge about how many patients are diagnosed correctly by CEUS and CECT.

Investigation of heterogeneity

Heterogeneity can appear due to different reference standards, with different accuracy. It is expected that the sensitivity of a reference standard in detecting liver metastases, consisting of a laparotomy including palpation, (IOUS) and a biopsy, or a pathological examination of surgically removed specimen, is higher than a reference standard consisting of only biopsy results or IOUS.

Heterogeneity can also appear due to different ways of selecting the study populations. The accuracy of both the index test and the comparator test are expected to be higher in a study population with a high proportion of diseased patients, than in a study population with a small proportion of diseased patients.

Blinding of the reference standard, the index test and the comparator test is an important issue. If some studies are blinded and others are not this could very likely lead to heterogeneity. A subgroup analysis of blinded and not blinded studies was planned to be conducted to examine if blinding is an issue. However, no subgroup analysis was performed because the included studies do not differ on this matter (see: Methodological quality of the included studies).

If variability in test accuracy among the studies is observed, subgroup analyses was planned to be conducted concerning the use of different reference standards, and different ways of selecting the study populations. No subgroup analysis was performed, as the test accuracy among the studies did not reveal any variability (see: Findings)

The above mentioned sources of heterogeneity are considered the most important. However, since heterogeneity can arise from both patient characteristics, test methods, study design and a lot of other factors, it might be necessary to take the investigation of heterogeneity even further (Macaskill 2010). Depending on the number of included studies it could be an issue to perform a meta-regression to investigate heterogeneity. If more than 15 studies were included then this approach was planned to be considered. Now, only 3 studies were included so such analyses have to await future updates (see: Results of the search)

Sensitivity analysis

If uninterpretable test results appear in the future, sensitivity analyses will be conducted classifying these results as test negatives in one analysis, test positives in a second analysis and exclude them in a third analysis. This kind of sensitivity analysis can reveal whether the uninterpretable test results occur randomly or not, and in this sense show if these test results arise bias or not (Reitsma 2009).

If the included studies report a lot of withdrawals it would be a good idea to do a sensitivity analysis on this issue as well. This can be done by counting withdrawals as test positives in one analysis and as test negatives in another to see if this makes the test results biased.

If the quality assessment of the studies according to the QUADAS tool reveals big differences in the quality of the included studies, a sensitivity analyses will be performed in future updates to assess the effect on the overall results of studies of poor quality compared to the overall results of studies of good quality. This will only be an issue when the number of included studies is higher than for the present.

Results

Results of the search

A total number of 12204 Studies were identified in the 7 different databases. 284 studies were retrieved in full report, and out of those studies only 3 studies could finally be included in the review. The last search for studies was performed the 11th of November 2011

Individual details of the included studies are tabulated in Appendix 3.

Individual details of the most relevant excluded studies are tabulated in Appendix 4.

Methodological quality of included studies

The methodological quality of the included studies is good (Appendix 5).

Li 2007: 9 out of the 11 QUADAS items are fulfilled with a clear yes, and 2 items are considered as unclear. It is unclear if the time period between reference standard and index test was short enough to be reasonably sure that the target condition did not change between the two tests. It is also

unclear if the reference standard results were interpreted without knowledge of the results of the index test.

Mainenti 2010: 9 out 11 QUADAS items are fulfilled with a clear yes, one item is not fulfilled and one item is considered as unclear. The reference standard results were interpreted with the knowledge of the results of the index test, and it is unclear if the same clinical data was available when test results were interpreted as would be available in practice.

Rafaelsen 2011: 10 out 11 QUADAS items are fulfilled with a clear yes, and one item is considered as unclear. It is unclear if the reference standard results were interpreted without knowledge of the results of the index test.

It is not stated in any of the included studies if there were any uninterpretable or intermediate test results. However, since every patient in the study population is accounted for this item is considered as fulfilled in all three included studies.

Findings

The sensitivity and specificity were calculated with 95 % exact confidence intervals (CI) for CEUS and CECT separately for each study as well as pooled estimates (Table 1). Heterogeneity in the estimates was investigated by Fisher's exact test.

Sensitivity (CEUS) p = 0.6803
Specificity (CEUS) p = 0.0020
Sensitivity (CECT) p = 1.0000
Specificity (CECT) p = 0.0717

These p-values show that there is significant heterogeneity concerning the specificity of CEUS.

To summarize the diagnostic ability of CEUS and CECT, the diagnostic odds ratio (DOR), positive and negative predicted values (PPV and NPV) as well as positive and negative likelihood ratios (LR+ and LR-) were computed (Table 2; Table 3; Table 4). These values were computed for each study separately and in case of homogeneity, pooled estimates were also determined.

Heterogeneity in the estimates of DOR was assessed by the Breslow-Day test, and because of homogeneity being plausible, the overall DOR was calculated by the Mantel-Haenszel method. Heterogeneity in the estimates of PPV and NPV was assessed by Fisher's exact test whereas heterogeneity of LR+ and LR- was investigated by a likelihood ratio test. For those measured demonstrating homogeneity across studies, the pooled estimates was calculated by pooling the tables from the individual studies. All estimates were provided with 95 % CI, exact intervals were provided for all estimates except for LR+ and LR-.

DOR (CEUS) p = 0.0767
DOR (CECT) p = 0.4384

PPV (CEUS) p = 0.0176
NPV (CEUS) p = 0.3922
PPV (CECT) p = 0.0184
NPV (CECT) p = 0.3818

Significant heterogeneity is here found in both PPV (CEUS) and PPV (CECT).

In one of the studies (Li 2007) all non-diseased patients were correctly classified as non-diseased by both CEUS and CECT. In the tables based on this study for which there are no observations in one of the cells, the DOR, LR+ and LR- cannot be determined

LR+/- (CEUS) p = 0.0039
LR+/- (CECT) p = 0.1254

The LR+/- (CEUS) estimate is significant concerning heterogeneity.

The performance of CEUS and CECT was compared by comparing the specificities and the sensitivities for each study by the exact McNemar test (marginal homogeneity). The information on a patient level is not available, only the marginals of the tables of sensitivity, respectively specificity, for CEUS versus CECT are known. The smallest p-value was determined based on all the p-values for the possible tables with these margins.

All 3 included studies showed homogeneity between CEUS and CECT concerning sensitivity (p = 1.0 in all 3 studies), and the pooled estimate as well showed homogeneity (p = 1.0). When it comes to the calculation of the specificity, it is not possible to test for marginal homogeneity in Li 2007 because both CEUS and CECT have 100 % specificity (0/94 in both groups). Concerning the remaining 2 studies no heterogeneity was found between CEUC and CECT (Mainenti 2010: p >= 0.250, Rafaelsen 2011: >= 0.0625).

Two-sided p-values < 0.05 were considered significant. Calculations were performed using SAS version 9.2.

The results concerning sensitivity and specificity are graphically displayed as a SROC curve (figure 1), and as coupled forest plots (figure 2 and figure 3).

The pooled estimates, which it is possible to compare because of no heterogeneity, consolidate the fact that there is no difference in diagnostic accuracy between CEUS and CECT.

Pooled estimates:

	CEUS	**CECT**
Sensitivity	88.10 % (95 % CI: 74.37 – 96.02)	85.71 % (95 % CI: 71.46 – 94.57)
Diagnostic odds ratio (DOR)	168.1160 (95 % CI: 48.4458 – 583.3940)	192.9177 (95 % CI: 63.4239 – 586.8016)
Negative predictive value (NPV)	98.64 % (95 % CI: 96.85 – 99.56)	98.36 % (95 % CI: 96.47 – 99.40)

The estimates and the 95 % CI show a high amount of overlap which indicates that there is no difference in the pooled estimates.

Discussion

The search strategy has provided 12204 hits, but only 3 studies fulfilled the inclusion criteria and could be included in the systematic review. This indicates that the search for studies has been broad enough, but the question is if the selection of studies may have been too selective? The QUADAS tool has revealed that the 3 included studies are of high quality. This gives the result of this review a high amount of validity. On the other hand 3 included studies are not enough to make clear conclusions. However, to include low quality studies creates bias and this is the reason why, e.g., studies which mix up the reference standard, the comparator test and the index test have been excluded.

Only 1 of the 3 included studies (Rafaelsen 2011) contained the required data to establish information about sensitivity and specificity. The last 2 included studies (Li 2007; Mainenti 2010) did not contain the required data, but these data were provided on request. The authors of 7 other studies have been contacted as well. One author did not want to provide his data, another author was not able to provide the required data, and the last 5 authors did not reply to the request for their data. Potentially this review could have contained 3 times the amount of included studies if all needed data had been available. This is of course a weakness of this review, but every author has been contacted 3 times from April 2011 to November 2011 without any success.

Since the number of included studies is only 3, it has not been an option to perform a meta-regression to look for heterogeneity. This is the reason why other ways of testing for heterogeneity have been performed. The results of those tests have not given any reason to perform any subgroup analysis. If a search for data some time later makes it possible to include more studies, a new investigation for heterogeneity will be performed.

The bivariate model has not been used in this review so far because of only 3 included studies. All the statistical analyses will be reconsidered if it at some time later is possible to include more studies.

No sensitivity analysis has been performed either, but the need of sensitivity analyses will as well be considered once again if the number of included studies at some time later gets higher.

Conclusions

The results show no significant difference between CEUS and CECT in diagnostic accuracy for detecting suspected liver metastases in patients. For this reason CEUS may be considered just as valuable as CECT when it comes to detecting and diagnosing liver metastases. However, the results are based on only 3 included studies. Absence of evidence for a difference must not be considered evidence of absence of a difference. Therefore, it is recommended to perform further diagnostic accuracy tests in this area. Furthermore, investigators ought to be required by law to share their data. The present situation where investigators take ownership of their data not wishing to share them is unsustainable and unethical.

References to studies

Additional references

Al-asfoor 2008

Al-asfoor A, Fedorowicz Z, Lodge M. Resection versus no intervention or other surgical interventions for colorectal cancer liver metastases. Cochrane Database of Systematic Reviews 2008, Issue 2. Art. No.: CD006039. DOI: 10.1002/14651858.CD006039.pub4.

Bossuyt 2008

Bossuyt PM, Leeflang MM. Chapter 6: Guide to the contents of a Cochrane review and protocol. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008]. The Cochrane Collaboration, 2008.

de Vet 2008

de Vet HCW, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for Studies. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated September 2008].. The Cochrane Collaboration, 2008.

Gluud 2010

Gluud C, Nikolova D, Klingenberg SL, Alexakis N, Als-Nielsen B, Colli A, et al. Cochrane Hepato-Biliary Group. About The Cochrane Collaboration (Cochrane Review Groups (CRGs)). 2010, Issue 12. Art. No.: LIVER.

Gurusamy 2009

Gurusamy KS, Ramamoorthy R, Sharma D, Davidson BR. Liver resection versus other treatments for neuroendocrine tumours in patients with resectable liver metastases. Cochrane Database of Systematic Reviews 2009, Issue 2. Art. No.: CD007060. DOI: 10.1002/14651858.CD007060.pub2.

Gurusamy 2009a

Gurusamy KS, Pamecha V, Sharma D, Davidson BR. Palliative cytoreductive surgery versus other palliative treatments in patients with unresectable liver metastases from gastro-entero-pancreatic neuroendocrine tumours. Cochrane Database of Systematic Reviews 2009, Issue 1. Art. No.: CD007118. DOI: 10.1002/14651858.CD007118.pub2.

Gurusamy 2010

Gurusamy KS, Ramamoorthy R, Imber C, Davidson BR. Surgical resection versus non-surgical treatment for hepatic node positive patients with colorectal liver metastases. Cochrane Database of Systematic Reviews 2010, Issue 1. Art. No.: CD006797. DOI: 10.1002/14651858.CD006797.pub2.

ICRP 2000

Rehani MM, Bongartz G, Golding SJ, Gordon L, Kalender W, Murakami T, et al. Managing patient dose in computed tomography, 2. How high are the doses? Annals of the ICRP 2000;30(4):19-24.

ICRP 2007a

Alexakhin RM, Cousins C, González AJ, Menzel H, Pentreath RJ, Shandala N, et al. Publication 103, The 2007 Recommendations of the International Commission on Radiological Protection, Preface, executive summary and glossary. Annals of the ICRP 2007;37(2-4):9-34.

ICRP 2007b

Alexakhin RM, Cousins C, González AJ, Menzel H, Pentreath RJ, Shandala N, et al. Publication 103, The 2007 Recommendations of the International Commission on Radiological Protection, Chapters 3 and 4. Annals of the ICRP 2007;37(2-4):49-79.

ICRP 2007c

Alexakhin RM, Cousins C, González AJ, Menzel H, Pentreath RJ, Shandala N, et al. Publication 103, The 2007 Recommendations of the International Commission on Radiological Protection, Chapter 7. Annals of the ICRP 2007;37(2-4):125-32.

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 23.12.2010.

Nelson 2006

Nelson RL, Freels S. Hepatic artery adjuvant chemotherapy for patients having resection or ablation of colorectal cancer metastatic to the liver. Cochrane Database of Systematic Reviews 2006, Issue 4. Art. No.: CD003770. DOI: 10.1002/14651858.CD003770.pub3.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. Journal of clinical epidemiology 2005;58(10):982-90. [PubMed: 16168343]

Reitsma 2009

Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2009.

Royle 2003

Royle P, Milne R. Literature searching for randomized controlled trials used in Cochrane reviews: rapid versus exhaustive searches. International Journal of Technology Assessment in Health Care 2003;19(4):591-603.

Townsend 2009

Townsend A, Price T, Karapetis C. Selective internal radiation therapy for liver metastases from colorectal cancer. Cochrane Database of Systematic Reviews 2009, Issue 4. Art. No.: CD007045. DOI: 10.1002/14651858.CD007045.pub2.

Whiting 2003

Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Medical Research Methodology 2003;3:25. [PubMed: 14606960]

Included studies

Li 2007

Li R, Guo Y, Hua X, He Y, Ding J, Guo A, et al. Characterization of focal liver lesions: comparison of pulse-inversion harmonic contrast-enhanced sonography with contrast-enhanced CT. Journal of clinical ultrasound : JCU 2007;35(3):109-17. [PubMed: 17295272]

Mainenti 2010

Mainenti PP, Mancini M, Mainolfi C, Camera L, Maurea S, Manchia A, Tanga M,Persico F, Addeo P, D'Antonio D, Speranza A, Bucci L, Persico G, Pace L,Salvatore M. Detection of colo-rectal liver metastases: prospective comparison of contrast enhanced US, multidetector CT, PET/CT, and 1.5 Tesla MR with extracellular and reticulo-endothelial cell specific contrast agents.. Abdominal Imaging 2010 October;35(5):511-521. [PubMed: 19562412]

Rafaelsen 2011

Rafaelsen SR, Jakobsen A. Contrast-enhanced ultrasound vs multidetector-computed tomography for detecting liver metastases in colorectal cancer: a prospective, blinded, patient-by-patient analysis. Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland 2011;13(4):420-5. [PubMed: 20412096]

Excluded studies

Dietrich 2006

Dietrich CF, Kratzer W, Strobe D, Danse E, Fessl R, Bunk A, et al. Assessment of metastatic liver disease in patients with primary extrahepatic tumors by contrast-enhanced sonography versus CT and MRI. World journal of gastroenterology : WJG 2006;12(11):1699-705. [PubMed: 16586537]

Hatanaka 2008

Hatanaka K, Kudo M, Minami Y, Maekawa K. Sonazoid-enhanced ultrasonography for diagnosis of hepatic malignancies: comparison with contrast-enhanced CT. Oncology 2008;75 Suppl 1:42-7. [PubMed: 19092271]

Janica 2007

Janica JR, Lebkowska U, Ustymowicz A, Augustynowicz A, Kamocki Z, Werel D, et al. Contrast-enhanced ultrasonography in diagnosing liver metastases. Medical science monitor : international medical journal of experimental and clinical research 2007;13 Suppl 1:111-5. [PubMed: 17507895]

Larsen 2007

Larsen LP, Rosenkilde M, Christensen H, Bang N, Bolvig L, Christiansen T, et al. The value of contrast enhanced ultrasonography in detection of liver metastases from colorectal cancer: a prospective double-blinded study. European journal of radiology 2007;62(2):302-7. [PubMed: 17194561]

Larsen 2009

Larsen LP, Rosenkilde M, Christensen H, Bang N, Bolvig L, Christiansen T, et al. Can contrast-enhanced ultrasonography replace multidetector-computed tomography in the detection of liver metastases from colorectal cancer? European journal of radiology 2009;69(2):308-13. [PubMed: 18068925]

Moriyasu 2009

Moriyasu F, Itoh K. Efficacy of perflubutane microbubble-enhanced ultrasound in the characterization and detection of focal liver lesions: phase 3 multicenter clinical trial. AJR. American journal of roentgenology 2009;193(1):86-95. [PubMed: 19542399]

Muhi 2010

Muhi A, Ichikawa T, Motosugi U, Sou H, Nakajima H, Sano K, et al. Diagnosis of colorectal hepatic metastases: Contrast-enhanced ultrasonography versus contrast-enhanced computed tomography versus superparamagnetic iron oxide-enhanced magnetic resonance imaging with diffusion-weighted imaging. Journal of magnetic resonance imaging : JMRI 2010;32(5):1132-40. [PubMed: 21031519]

Quaia 2006

Quaia E, D'Onofrio M, Palumbo A, Rossi S, Bruni S, Cova M. Comparison of contrast-enhanced ultrasonography versus baseline ultrasound and contrast-enhanced computed tomography in metastatic disease of the liver: diagnostic performance and confidence. European radiology 2006;16(7):1599-609. [PubMed: 16552507]

Tables and figures

Table 1

Li 2007	**CEUS**	**CECT**
Sensitivity	93.33 % (95 % CI: 68.05 – 99.83)	86.67 % (95 % CI: 59.54 – 98.34)
Specificity	100 % (95 % CI: 96.15 – 100)	100 % (95 % CI: 96.15 – 100)
Mainenti 2010	**CEUS**	**CECT**
Sensitivity	83.33 % (95 % CI: 35.88 – 99.58)	83.33 % (95 % CI: 35.88 – 99.58)
Specificity	85.71 % (95 % CI: 67.33 – 95.97)	96.43 % (95 % CI: 81.65 – 99.91)
Rafaelsen 2011	**CEUS**	**CECT**
Sensitivity	85.71 % (95 % CI: 63.66 – 96.95)	85.71 % (95 % CI: 63.66 – 96.95)
Specificity	97.60 % (95 % CI: 94.85 – 99.11)	95.60 % (95 % CI: 92.26 – 97.78)
Pooled estimates	**CEUS**	**CECT**
Sensitivity	88.10 % (95 % CI: 74.37 – 96.02)	85.71 % (95 % CI: 71.46 – 94.57)
Specificity	Cannot be reported because of heterogeneity.	96.77 % (95 % CI: 94.43 – 98.32)

Table 2

Li 2007	**CEUS**	**CECT**
Diagnostic odds ratio (DOR)	Cannot be determined.	Cannot be determined.
Mainenti 2010	**CEUS**	**CECT**
Diagnostic odds ratio (DOR)	30 (95 % CI: 2.7385 – 328.6443)	135 (95 % CI: 7.1980 – 2531.9361)
Rafaelsen 2011	**CEUS**	**CECT**
Diagnostic odds ratio (DOR)	244 (95 % CI: 56.3126 – 1057.2407)	130.3636 (95 % CI: 33.3415 – 509.7149)
Pooled estimates	**CEUS**	**CECT**
Diagnostic odds ratio (DOR)	168.1160 (95 % CI: 48.4458 – 583.3940)	192.9177 (95 % CI: 63.4239 – 586.8016)

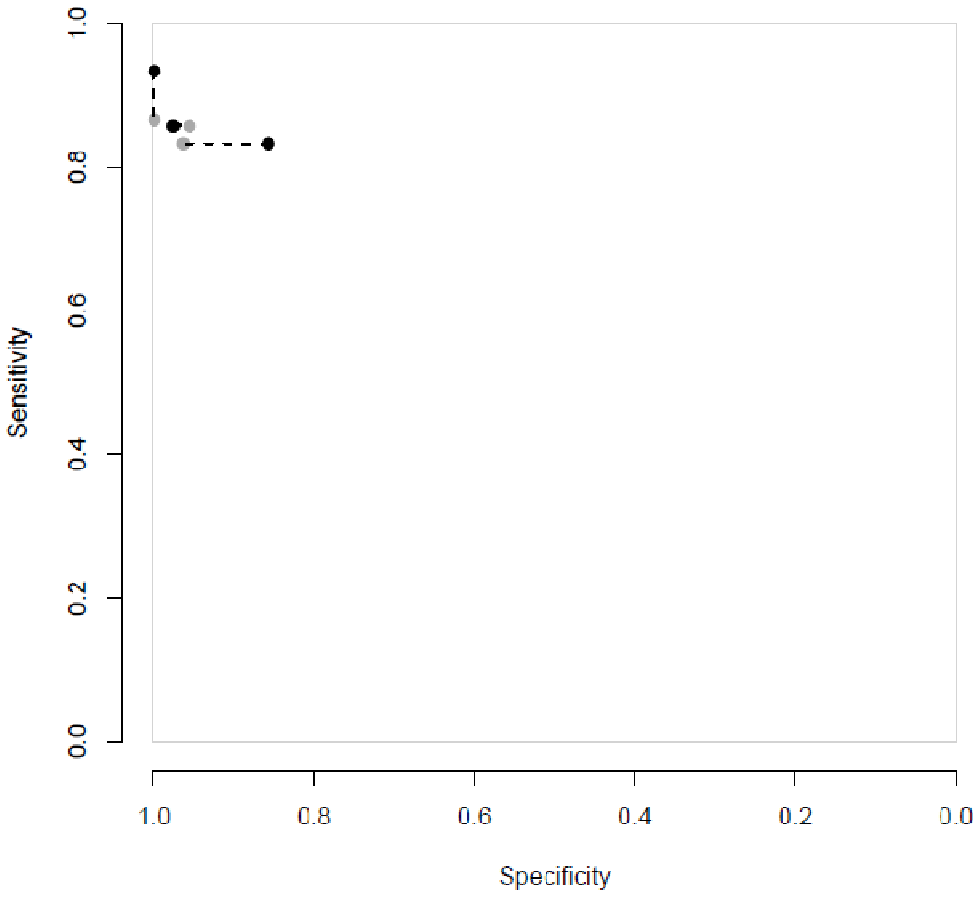
Table 3

Li 2007	**CEUS**	**CECT**
Positive predictive value (PPV)	100 % (95 % CI: 96.15 - 100)	100 % (95 % CI: 96.15 - 100)
Negative predictive value (NPV)	98.95 % (95 % CI: 94.27 – 99.97)	97.92 % (95 % CI: 92.68 – 99.75)
Mainenti 2010	**CEUS**	**CECT**
Positive predictive value (PPV)	55.56 % (95 % CI: 21.20 – 86.30)	83.33 % (95 % CI: 65.88 – 99.58)
Negative predictive value (NPV)	96 % (95 % CI: 79.65 – 99.60)	96.43 % (95 % CI: 81.65 – 99.91)
Rafaelsen 2011	**CEUS**	**CECT**
Positive predictive value (PPV)	75 % (95 % CI: 53.29 – 90.23)	62.07 % (95 % CI: 42.26 – 79.31)
Negative predictive value (NPV)	98.79 % (95 % CI: 96.49 – 99.75)	98.76 % (95 % CI: 96.42 – 99.74)
Pooled estimates	**CEUS**	**CECT**
Positive predictive value (PPV)	Cannot be reported because of heterogeneity.	Cannot be reported because of heterogeneity.
Negative predictive value (NPV)	98.64 % (95 % CI: 96.85 – 99.56)	98.36 % (95 % CI: 96.47 – 99.40)

Table 4

Li 2007	**CEUS**	**CECT**
Positive likelihood ratio (LR+)	Cannot be determined.	Cannot be determined.
Negative likelihood ratio (LR-)	Cannot be determined.	Cannot be determined.
Mainenti 2010	**CEUS**	**CECT**
Positive likelihood ratio (LR+)	5.8333 (95 % CI: 2.1996 – 15.4698)	23.3333 (95 % CI: 3.2945 – 165.2587)
Negative likelihood ratio (LR-)	0.1944 (95 % CI: 0.0323 – 1.1711)	0.1728 (95 % CI: 0.0288 – 1.0358)
Rafaelsen 2011	**CEUS**	**CECT**
Positive likelihood ratio (LR+)	35.7143 (95 % CI: 15.8950 – 80.2460)	19.4805 (95 % CI: 10.6526 – 35.6242)
Negative likelihood ratio (LR-)	0.1464 (95 % CI: 0.0513 – 0.4174)	0.1494 (95 % CI: 0.0524 – 0.4262)
Pooled estimates	**CEUS**	**CECT**
Positive likelihood ratio (LR+)	Cannot be reported because of heterogeneity.	26.5714 (95 % CI: 15.0250 – 46.9911)
Negative likelihood ratio (LR-)	Cannot be reported because of heterogeneity.	0.1476 (95 % CI: 0.0704 – 0.3097)

Figure 1



Black dots = contrast-enhanced ultrasound (CEUS) and gray dots = contrast-enhanced computed tomography (CECT).

Figure 2
Forest plot contrast-enhanced ultrasound (CEUS): sensitivity, specificity and pooled estimates.

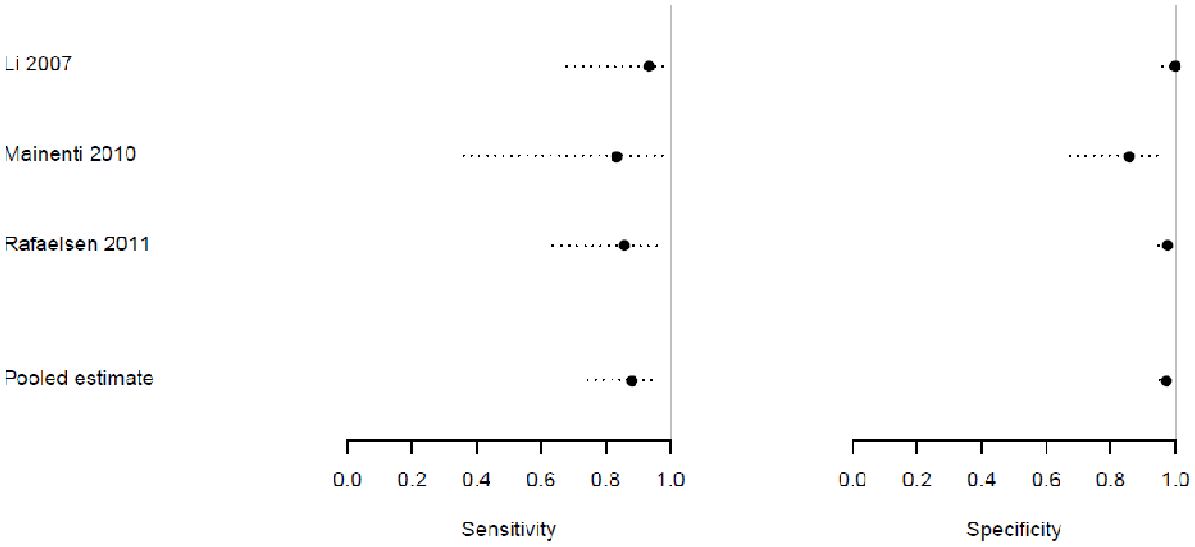
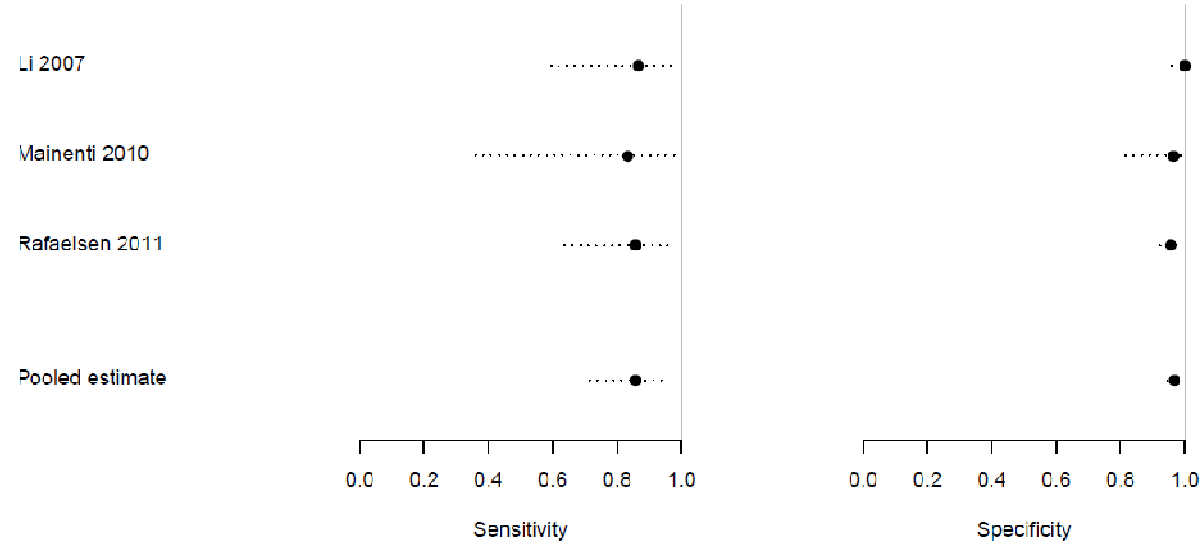


Figure 3
Forest plot contrast-enhanced computed tomography (CECT): sensitivity, specificity and pooled estimates.



Appendices

1 Search strategies

Database	Time span	Search strategy
The Cochrane Hepato-Biliary Group Controlled Trials Register	Updated search performed the 11th of November 2011	('computed tomograph*' OR CT OR MSCT OR CECT) AND (ultrasound OR ultrasonograph* OR US OR CEUS) AND ((liver OR hepat*) AND (metasta* OR secondar* OR spread OR advanced))
The Cochrane Hepato-Biliary Group Diagnostic Test Accuracy Studies Register	Updated search performed the 11th of November 2011	('computed tomograph*' OR CT OR MSCT OR CECT) AND (ultrasound OR ultrasonograph* OR US OR CEUS) AND ((liver OR hepat*) AND (metasta* OR secondar* OR spread OR advanced))
The Cochrane Library	Updated search performed the 11th of November 2011	#1 MeSH descriptor Tomography, X-Ray Computed explode all trees #2 computed tomograph* OR CT OR MSCT OR CECT #3 (#1 OR #2) #4 MeSH descriptor Ultrasonography explode all trees #5 ultrasound OR ultrasonograph* OR US OR CEUS #6 (#4 OR #5) #7 MeSH descriptor Liver Neoplasms explode all trees #8 (liver OR hepat*) AND (metasta* OR secondar* OR spread OR advanced) #9 (#7 OR #8) #10 (#3 AND #6 AND #9)
MEDLINE (Ovid SP)	1950 to the 11th of November 2011	1. exp Tomography, X-Ray Computed/ 2. (computed tomograph* or CT or MSCT or CECT).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] 3. 1 or 2 4. exp Ultrasonography/ 5. (ultrasound or ultrasonograph* or US or CEUS).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] 6. 4 or 5 7. exp Liver Neoplasms/

		8. ((liver or hepat*) and (metasta* or secondar* or spread or advanced)).mp. [mp=protocol supplementary concept, rare disease supplementary concept, title, original title, abstract, name of substance word, subject heading word, unique identifier] 9. 7 or 8 10. 3 and 6 and 9
EMBASE (Ovid SP)	1980 to the 11th of November 2011	1. exp COMPUTED TOMOGRAPHY SCANNER/ or TOMOGRAPHY/ 2. (computed tomograph* or CT or MSCT or CECT).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 3. 1 or 2 4. exp ULTRASOUND/ 5. exp echography/ 6. (ultrasound or ultrasonograph* or US or CEUS).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 7. 4 or 5 or 6 8. exp liver metastasis/ 9. ((liver or hepat*) and (metasta* or secondar* or spread or advanced)).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer] 10. 8 or 9 11. 3 and 7 and 10
Science Citation Index Expanded (ISI Web of Knowledge)	1900 to the 11th of November 2011	# 4 #3 AND #2 AND #1 # 3 TS=((liver or hepat*) and (metasta* or secondar* or spread or advanced)) # 2 TS=(ultrasound or ultrasonograph* or US or CEUS) # 1 TS=(computed tomograph* or CT or MSCT or CECT)
ACP Journal Club (EBSCO host)	2002 to the 11th of November 2011	S11 S1 and S4 and S7 and S10 S10 S8 or S9 S9 TX ((liver or hepat*) and (metasta* or secondar* or spread or advanced)) S8 SU liver metastasis S7 S5 or S6 S6 TX ultrasound OR ultrasonograph* OR US OR CEUS S5 SU ultrasonography S4 S2 or S3

		S3 TX computed tomograph* OR CT OR MSCT OR CECT S2 SU tomography S1 JN "ACP Journal Club"

2 QUADAS items

QUADAS item	Yes	No
1. Was the spectrum of patients representative of the patients who will receive the test in practice?	If the patients are suspected of having liver metastases.	If the patients (or some of the patients) are not suspected of having liver metastases.
2. Is the reference standard likely to classify the target condition correctly?	If the reference standard consists of biopsy test results, pathological examinations of surgically removed specimens, intra-operative ultrasound (IOUS), follow up, or a combination of all four.	If the reference standard consists of none of those described under "yes".
3. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	If the period is no longer than two weeks for both the index test, the comparator test and the reference standard to be executed on the same patient (this does not count for the reference standard, if the reference standard is follow up).	If the period is longer than two weeks, and the reference standard is not follow up.
4. Did the whole sample or a random selection of the sample, receive verification using the intended reference standard?	If all patients in the studies based on direct comparison have received the reference standard. If a random selection of patients, who received the index test or the comparator test in the studies with not fully paired designs, also received the reference standard.	If the patients do not receive the reference standard in accordance with the description under "yes".

5. Did patients receive the same reference standard irrespective of the index test result?	If all patients received the same reference standard (this does not count for patients who undergo to follow up, because they have both a negative index test result and a negative comparator test result).	If the patients did not receive the same reference standard (this does not count for patients who undergo to follow up, because they have both a negative index test result and a negative comparator test result).
6. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	If the index test and/or the comparator test do not form part of the reference standard in any way	If the reference standard formally include the results of the index test and/or the comparator test.
7. Were the reference standard results interpreted without knowledge of the results of the index test?	If the reference standard results are interpreted blind to the results of the index test and the comparator test.	If the reference standard results are not interpreted blind to the results of the index test and the comparator test.
8. Were the index test results interpreted without knowledge of the results of the reference standard?	If the index test results and the comparator test results are interpreted blind to the results of the reference standard.	If the index test results and the comparator test results are not interpreted blind to the results of the reference standard.
9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	If relevant clinical data is available without ruining the blinding.	If no relevant clinical data is available, or if too much data is available and the blinding is ruined.
10. Were uninterpretable / intermediate test results reported?	If the number of uninterpretable test results is stated.	If the number of uninterpretable test results is not stated.
11. Were withdrawals from the study explained?	If all patients are accounted for.	If not all patients are accounted for.

3 Characteristics of included studies

Li 2007

<table>
<tr><th>Heading</th><th>Content</th></tr>
<tr><td>Study ID</td><td>Li R, Guo Y, Hua X, He Y, Ding J, Guo A, et al. Characterization of focal liver lesions: comparison of pulse-inversion harmonic contrast-enhanced sonography with contrast-enhanced CT. Journal of clinical ultrasound : JCU 2007;35(3):109-17. [PubMed: 17295272]</td></tr>
<tr><td>Clinical features and settings</td><td>Study population: patients admitted to Southwest Hospital, Third Military Medical University, Chongqing 400038, China.

Purpose: to compare the efficacy of CEUS, for the characterization of focal liver lesions, with that of CECT.</td></tr>
<tr><td>Participants</td><td>109 patients (37 women, 72 men; age range 18-79 years, mean age 46 +/- 12 years).</td></tr>
<tr><td>Study design</td><td>Prospective study. 109 patients (37 women, 72 men; age range 18-79 years, mean age 46 +/- 12 years) with focal liver lesions, including 61 hepatocellular carcinomas, 15 liver metastases, 5 cholangiocellular carcinomas, 12 hemangiomas, 5 regenerative nodules, 3 adenomas, 3 focal nodular hyperplasias, 4 focal necroses and 1 angiomyolipoma.

Only the data concerning the patients with liver metastases is considered for this review. This data is kindly provided on request by Rui Li, MD.

Metastases, CEUS:

<table>
<tr><td>True Positive (TP):

14</td><td>False Positive (FP):

0</td></tr>
<tr><td>False Negative (FN):

1</td><td>True Negative (TN):

94</td></tr>
</table>

Metastases, CECT:

<table>
<tr><td>True Positive (TP):

13</td><td>False Positive (FP):

0</td></tr>
</table>
</td></tr>
</table>

	False Negative (FN): 2	True Negative (TN): 94
Target condition and reference standard	The target condition is focal liver lesions. The reference standard is histopathologic diagnosis by means of surgical resection in 56 cases, and percutaneous needle biopsy with an 18-gauge needle in 53 cases.	
Index and comparator tests	Index test: CEUS Comparator test: CECT	
Follow up	Cross-sectional design. No follow up is reported. No uninterpretable tests reported.	

Mainenti 2010

Heading	**Content**
Study ID	Mainenti PP, Mancini M, Mainolfi C, Camera L, Maurea S, Manchia A, Tanga M,Persico F, Addeo P, D'Antonio D, Speranza A, Bucci L, Persico G, Pace L,Salvatore M. Detection of colo-rectal liver metastases: prospective comparison of contrast enhanced US, multidetector CT, PET/CT, and 1.5 Tesla MR with extracellular and reticulo-endothelial cell specific contrast agents.. Abdominal Imaging 2010 October;35(5):511-521. [PubMed: 19562412]
Clinical features and settings	Study population: patients with histological proven diagnosis of colo-rectal adenocarcinoma and scheduled for surgery at University of Naples "Federico II", Via Pansini 5, 80131 Naples, Italy. Purpose: to compare CEUS, CECT, 1.5 Tesla MR with extra-cellular (Gd-enhanced) and intra-cellular (SPIO-enhanced) contrast agents and PET/CT, in the detection of hepatic metastases from colorectal cancer.
Participants	34 patients with histological proven diagnosis (the histological specimen was obtained during conventional colonoscopy) of colorectal adenocarcinoma (20 men and 14 women; age range, 29 – 81 years; mean age: 63 years).
Study design	Prospective study. Only the data concerning CEUS and CECT is considered for this review.

<table>
<tr><td></td><td>This data is kindly provided on request by Pier Paolo Mainenti.

Metastases, CEUS:
<table><tr><td>True Positive (TP):

5</td><td>False Positive (FP):

4</td></tr><tr><td>False Negative (FN):

1</td><td>True Negative (TN):

24</td></tr></table>
Metastases, CECT:
<table><tr><td>True Positive (TP):

5</td><td>False Positive (FP):

1</td></tr><tr><td>False Negative (FN):

1</td><td>True Negative (TN):

27</td></tr></table></td></tr>
<tr><td>Target condition and reference standard</td><td>The target condition is hepatic metastases from colorectal cancer.

Reference standard: all the patients underwent surgery. At laparotomy, the liver was evaluated by means of bimanual palpation and intraoperative-US (IOUS). In all patients positive for the presence of metastases, biopsy or resection of at least one lesion was performed.</td></tr>
<tr><td>Index and comparator tests</td><td>Index test: CEUS

Comparator test: CECT</td></tr>
<tr><td>Follow up</td><td>All the patients underwent abdomino-pelvic CECT performed 6 months and 12 months after the surgery (1) to assess the size of the lesions not biopsied, (2) to evaluate the development of new hepatic metastases.</td></tr>
</table>

Rafaelsen 2011

<table>
<tr><th>Heading</th><th>Content</th></tr>
<tr><td>Study ID</td><td>Rafaelsen SR, Jakobsen A. Contrast-enhanced ultrasound vs multidetector-computed tomography for detecting liver metastases in colorectal cancer: a prospective, blinded, patient-by-patient analysis. Colorectal disease : the official journal of the Association of Coloproctology of Great Britain and Ireland 2011;13(4):420-5. [PubMed: 20412096]</td></tr>
<tr><td>Clinical features and settings</td><td>Study population: surgically treated patients with histopathological confirmed primary colorectal cancer.

Purpose: to compare the sensitivity and specificity of CEUS and CECT in the detection of liver metastases in patients with colorectal cancer.</td></tr>
<tr><td>Participants</td><td>271 surgically treated patients with primary colorectal cancer [146 men and 125 women, median age 68 (range: 34-91) years].</td></tr>
<tr><td>Study design</td><td>Prospective study.

Metastases, CEUS:

<table>
<tr><td>True Positive (TP):

18</td><td>False Positive (FP):

6</td></tr>
<tr><td>False Negative (FN):

3</td><td>True Negative (TN):

244</td></tr>
</table>

Metastases, CECT:

<table>
<tr><td>True Positive (TP):

18</td><td>False Positive (FP):

11</td></tr>
<tr><td>False Negative (FN):

3</td><td>True Negative (TN):

239</td></tr>
</table>
</td></tr>
<tr><td>Target condition and reference standard</td><td>The target condition is hepatic metastases from colorectal cancer.

Reference standard: Intra-operative evaluation of the liver, including inspection, bimanual palpation and intra-operative ultrasound (IOUS). When liver resection was performed , the pathological examination</td></tr>
</table>

	contributed to the assessment.
Index and comparator tests	Index test: CEUS Comparator test: CECT
Follow up	Any new metastases detected beyond 3 months after CEUS and CECT were not considered synchronous liver metastases.

4 Characteristics of excluded studies

	Reason for exclusion
Dietrich 2006	CECT is used as the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Hatanaka 2008	CEUS and CECT are used as parts of the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Janica 2007	CECT is used as the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Larsen 2007	CEUS and CECT are used as parts of the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Larsen 2009	CEUS and CECT are used as parts of the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Moriyasu 2009	CEUS and CECT are used as parts of the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Muhi 2010	CECT is used as the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.
Quaia 2006	CECT is used as the reference standard. Differential verification is also an issue because not all of the patients were verified by the same reference standard.

5 Assessment of methodological quality

QUADAS item	Li 2007	Mainenti 2010	Rafaelsen 2011
1	yes	yes	yes
2	yes	yes	yes
3	unclear	yes	yes
4	yes	yes	yes
5	yes	yes	yes
6	yes	yes	yes
7	unclear	no	unclear
8	yes	yes	yes
9	yes	unclear	yes
10	yes	yes	yes
11	yes	yes	yes