Ager Michael (Orcid ID: 0000-0001-9924-4643)
Salonia Andrea (Orcid ID: 0000-0002-0595-7165)
Laguna Pilar (Orcid ID: 0000-0003-0906-4417)

Radiological features characterising indeterminate testes masses; A systematic review and meta-analysis

Michael Ager [a], Sarah Donegan [b], Luca Boeri [c], Javier Mayor de Castro [d], James F Donaldson [e, f], Muhammad Imran Omar [f], Konstantinos Dimitropoulos [e, f], Tharu Tharakan [a], Florian Janisch [g], Tim Muilwijk [h], Cathy Yuan [i], Catrin Tudur-Smith [b], Rien JM Nijman [j], Christian Radmayr [k], Andrea Salonia [l, *], Maria P Laguna Pes [m, *], Suks Minhas [a, *].

[a] Imperial College Healthcare NHS Trust, Dept. of Urology, London, United Kingdom, [b] Department of Health Data Science, University of Liverpool, United Kingdom, [c] Foundation IRCCS Ca' Granda, Ospedale Maggiore Policlinico Hospital, University of Milan, Dept. of Urology, Milan, Italy, [d] Hospital Gregorio Marañón, Dept. of Urology, Madrid, Spain, [e] Aberdeen Royal Infirmary, Dept. of Urology, Aberdeen, United Kingdom, [f] University of Aberdeen, Academic Urology Unit, Aberdeen, United Kingdom, [g] University Medical Center Hamburg-Eppendorf, Dept. of Urology, Hamburg, Germany, [h] University Hospitals Leuven, Dept. of Urology, Leuven, Belgium, [i] Division of gastroenterology, Department of Medicine, McMaster University, Hamilton, ON, Canada, [j] Martini Ziekenhuis, Dept. of Urology, Groningen, The Netherlands,

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/bju.15869

[k] Medizinische Universität Innsbruck, Dept. of Urology, Innsbruck,

Austria, [l] Vita-Salute San Raffaele University , Dept. of Urology, Milan, Italy,

[m] Istanbul Medipol University, Dept. of Urology, Istanbul, Turkey.

* Joint senior authors.

Corresponding Author:

Michael Ager,

Department of Urology

Charing Cross Hospital, Imperial College NHS Trust

Fulham Palace Road, W6 8RF, London, UK

Email: m.ager@nhs.net

Abstract: Context: The use of scrotal ultrasound (SUS) has increased the detection rate of indeterminate testicular masses. Defining radiological characteristics that identify malignancy may reduce the number of men undergoing unnecessary radical orchidectomy.

Objective: To define which SUS or scrotal magnetic resonance imaging (MRI) characteristics can predict benign or malignant disease in pre or post pubertal males with indeterminate testicular masses.

Evidence Acquisition: This SR was conducted in accordance with Cochrane Collaboration guidance. Medline, Embase, Cochrane controlled trials and systematic reviews databases were searched from (1970 - March 26, 2021). Benign and malignant masses were classified using the reported reference test: i.e., histopathology, or 12 months progression-free radiological surveillance. Risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool (QUADAS – 2).

Evidence Synthesis: 32 studies were identified, including 1692 masses of which 28 studies and 1550 masses reported SUS features, 4 studies and 142 masses reported MRI features. Meta-analysis of different SUS B mode values in post pubertal men demonstrated size of ≤0.5cm had a significant lower OR of malignancy compared to masses >0.5cm ($p < 0.001$). Comparison of masses 0.6-1.0cm and masses >1.5cm also demonstrated a significant lower OR of malignancy ($p = 0.04$). No significanct difference was observed between masses of 0.6-1.0cm and 1.1-1.5cm. SUS in post pubertal men also had a statistically significant lower odds of malignancy for heterogenous masses vs. homogenous

masses ($p = 0.04$), hyperechogenic vs. hypoechogenic masses ($p < 0.01$), normal vs. increased enhancement ($p < 0.01$) and peripheral vs. central vascularity ($p < 0.01$), respectively.
There was limited data on pre pubertal SUS, pre pubertal MRI and post pubertal MRI.
Conclusions: This meta-analysis identifies radiological characteristics that have a lower odds of malignancy and may be of value in the management of the indeterminate testis mass.

1. Introduction

Testicular tumours include germ cell tumours (seminomas and non-seminomas), non-germ cell (sex cord stromal, such as Leydig or Sertoli cell) tumours, lymphoma and metastases (1). The standard treatment for suspicious masses diagnosed on scrotal ultrasonography (SUS) is radical orchidectomy (RO) (2–4). RO can however negatively impact fertility, endocrine function, sexual satisfaction as well as leading to anxiety and depression (5). This is important as approximately 30% of orchidectomies for non-palpable testicular masses in post-pubertal males and between 63-94% in pre pubertal males are histopathologically benign (6–8). The increasing application and widespread use of SUS and the limitations of SUS and scrotal MRI in their ability to accurately identify a testicular mass as either benign or malignant, represents a significant management dilemma for clinicians, particularly in the context of an indeterminate testicular mass (9). 'Indeterminate testicular masses' encompass a broad range of tumours, for which there is no consensus definition. It is acknowledged that imaging on scrotal ultrasound (SUS) or scrotal magnetic resonance imaging (MRI) is non diagnostic. No definitive radiological features to date distinguish benign or malignant disease. Konstantatou et al. describe the indeterminate mass as a focal, circumscribed, or ill-defined area within the testicular parenchyma with variable levels of echogenicity or doppler flow that would necessitate surgical, medical, or imaging follow up

[(10)]. Several studies describe indeterminate masses interchangeably i.e., incidental non-palpable or small intra testicular masses with size thresholds ranging from ≤0.5 cm - ≤2.0cm (11,12). Most studies also define indeterminate masses as those found in men with normal tumour markers (11–16).

The primary aim of this systematic review and meta-analysis was to define SUS and scrotal MRI characteristics that correctly identify malignancy in post pubertal males. Identifying these features may reduce the number of men who undergo unnecessary RO. The secondary aim was to define SUS and scrotal MRI characteristics that correctly identify malignancy in pre pubertal males.

2. Evidence acquisition and study eligibility assessment

A systematic review and meta-analysis was performed in accordance with Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement (17), the Cochrane Handbook for Systematic Reviews of Diagnostic test accuracy (18) and the key steps in conducting systematic reviews underpinning clinical practice guidelines from the EAU Guidelines Office [19). The protocol was registered with PROSPERO (CRD42020199339). Medline, Embase, the Cochrane database of systematic reviews and central register of controlled trials were systematically searched between 1970 and March 2021 for relevant English language publications. The published *a priori* protocol includes the search strategy (20).

Following deduplication, two reviewers (MA) and (JMdC) independently screened the abstracts and full texts for eligibility. Any disagreements were resolved by a third reviewer (JFD).

2.1. Types of study design

All types of studies including observational, or case-control studies were included. Studies reporting outcomes on <5 patients were excluded.

2.2. Target condition

Any pre- or post-pubertal male with an indeterminate testicular mass as defined by the authors. Studies examining men with a previous history of TGCT (testicular germ cell tumours) (e.g., indeterminate mass in a contralateral testis after previous orchidectomy)

were excluded except in cases with a mixed population which accounted for <10% of the cohort.

2.3. Index tests

SUS or MRI were the index test(s). Any of the following index test features were included:

1. Scrotal ultrasound scan: B mode (texture, echogenicity, size, calcification, margin), colour flow doppler (location and vascularity), contrast enhanced ultrasound (enhancement), alone or in combination as a multiparametric SUS (mpSUS).

2. Magnetic Resonance Imaging of the scrotum: diffusion weighted imaging (DWI), dynamic contrast enhancement (DCE), apparent diffusion coefficient (ADC), diffusion and magnetization tensor alone or in combination as a multiparametric MRI (mpMRI). Semi quantitative parameters of DCE on scrotal MRI, such as time to reach maximal peak enhancement, percentage of peak enhancement of contrast, wash in rate of contrast, volume transfer constant and rate constant, were also assessed where reported.

2.5. Reference standard

The reference standard for comparison was histopathological examination of testicular masses after RO, testes sparing surgery (TSS), or enucleation performed through a trans-inguinal approach. When surgery was not performed, regression or stability of the 'indeterminate testicular mass' over 12 months or more of radiological follow up was considered indicative of benign disease. Stability constituted no significant interval growth

as defined by the study authors. Patients who underwent surgery during the surveillance period were included in the cohort of histopathological outcomes only.

2.6. Data extraction

We were unable to extract 2x2 tables - i.e., true, and false positives (TP), (FP), true and false negatives (TN), (FN) - as in a standard diagnostic test review. This was because patients were not diagnosed as positive or negative by SUS or scrotal MRI. We therefore explored the risk of malignancy with the following radiological features: SUS B mode assessing; margins (regular or irregular); size ≤0.5, 0.6-1, 1.1-1.5, >1.5cm; echogenicity (hypoechoic, hyperechoic, isoechoic) microlithiasis (yes, no) and texture (heterogenous, homogenous), CEUS enhancement (increased, normal, decreased) and CFD doppler flow (mixed, central, peripheral, and no, low, high). Scrotal MRI assessed T1 and T2; margins (regular or irregular); size ≤0.5, 0.6-1, 1.1-1.5, >1.5cm; microlithiasis (yes, no) and texture (heterogenous, homogenous), DCE (increased, normal, decreased), vascularity (mixed, central, peripheral, and no, low, high).

For each feature category, we extracted the number of patients with malignant masses and the number with benign masses. When a study reported the number of masses rather than the number of patients (with some patients having >1 mass), we extracted this data.

Data was extracted by MA, LB, TT, FJ, TM, JMdC, independently into a data extraction form and data compared. Discrepancies in the data were resolved through discussion with a senior reviewer (JFD).

2.7. Data Analysis

Basic descriptive characteristics were presented. For each feature with two categories (e.g., heterogenous vs. homogenous), a meta-analytic logistic regression model was fitted with malignancy (yes/no) defined as the event of interest and the categories of the feature defined as two groups to be compared. Therefore, the odds ratio (OR) can be interpreted as the odds of malignancy in category 1 compared to the odds of malignancy in category 2. For each feature with more than two categories (e.g., peripheral vs. central vs. mixed doppler flow), a regression model was fitted for each pair of categories (rather than one model including all categories) to ensure results were based on the maximum possible number of studies due to missing data from various categories in some studies.

Heterogeneity was assessed using the chi-square test ($p<0.1$ indicating significant heterogeneity) and I^2 statistic. Fixed effects and random effects meta-analysis models were applied. Results from a random effects model were reported when heterogeneity was detected, otherwise a fixed effect model was used.

We carried out sensitivity analyses excluding studies that reported the number of masses rather than the number of patients as well as studies that did not distinguish the proportion of the study population who were pre or post pubertal as we recognise a different disease process exists in these two cohorts and therefore analysed these two groups independently (1,21).

All logistic regression analyses were performed in R using the meta-R package.

Models were fitted using the metandi command in Stata. Plots were produced using Review Manager 5.4 (22)

We reported a narrative synthesis of scrotal MRI and pre pubertal SUS studies as there were insufficient studies to perform a meta-analysis.

2.8. Assessment of risk of bias

Risk of bias (RoB) was assessed independently by MA, LB, TT, TM, and FJ using the Quality Assessment of Diagnostic Accuracy Studies-2 tool (QUADAS-2) and then compared the assessment results (23).

3. Evidence synthesis

3.1. Quantity of evidence identified

A total of 3380 abstracts were screened and 283 full text articles reviewed. 32 studies were eligible for inclusion (9–16,21,24–46). There were 22 studies reporting SUS outcomes in 1260 masses in post pubertal males, three SUS studies with 212 masses in a mixed population of pre- and post-pubertal males, three studies reporting SUS in 78 masses in pre-pubertal males and four studies reporting scrotal MRI outcomes in 142 masses in post-pubertal males. No study reported scrotal MRI outcomes for pre-pubertal males. Fig 1 PRISMA flow diagram of the selection process for the review.

Malignant tumours of the testes included 440 GCTs, 9 sex cord stromal tumours. 32 cases were either metastasis to the testis, sarcoma, or lymphoma. Benign masses included 235 Leydig cell tumours. Other benign masses included epidermoid cysts and fibrous pseudo tumours.

Ager Michael (Orcid ID: 0000-0001-9924-4643)
Salonia Andrea (Orcid ID: 0000-0002-0595-7165)
Laguna Pilar (Orcid ID: 0000-0003-0906-4417)

3.2. Characteristics and outcomes of included studies

32 studies were included reporting 1692 testicular masses. (Table 1).

Ager Michael (Orcid ID: 0000-0001-9924-4643)
Salonia Andrea (Orcid ID: 0000-0002-0595-7165)
Laguna Pilar (Orcid ID: 0000-0003-0906-4417)

3.3. Risk of bias assessment

We performed a risk of bias assessment using the QUADAS-2 tool for all included studies (n = 32), see Figure 2.

Patient selection

23 of 32 studies had a low risk of bias and 9 of 32 studies were found to have a high risk of bias for patient selection. Ayati et al. included 10 patients over a 2-year period (28). Whilst their inclusion criteria of non-palpable incidental masses are described, there is no clarity on exclusion criteria or the nature of patient sampling (i.e., random, or consecutive patients). Manganaro et al. only included patients with a histopathology result; excluding patients who may have only had radiological follow up, as well as omitting patients in whom the masses did not enhance on scrotal MRI (45). Similarly, Avci et al. excluded 2 of 11 patients who declined surgery (27). Six of the 32 studies did not specify their methods of patient selection and/or criteria for participant exclusion.

Index test

14 of 32 studies were found to have a low risk of bias whilst 17 of 32 studies were deemed to have an unclear risk of bias for the index test. Due to the retrospective nature of these studies, it was unclear if the index tests had been reported without prior knowledge of the reference test result. The study by Chang et al. was deemed to have a high risk of bias as the radiologist was not blinded to the reference test (histopathology) when they retrospectively reviewed the SUS (41).

Reference standard

Seven of 32 studies were found to have a low risk of bias whilst 24 of 32 studies were judged to have an unclear risk of bias for the reference standard. There was often no clear implementation of blinding of the reference test to the index test. Reginelli et al. had a high risk of bias as interpretation of the reference test was not blinded to the index test (16). Four studies had an unclear risk of applicability concern (10,32,33,42).

Flow and timing

31 of 32 studies were found to have a low risk of bias for flow and timing. Avci et al. was deemed to have a high risk of bias because not all patients received a reference test (27).

3.4. Synthesis of results

3.4.1 SUS B mode:

Size

Size of the testicular mass was assessed in ten studies which included 334 patients (14,15,24,25,27,28,30,37–39). We divided them into four size groups: ≤0.5cm, 0.6-1.0cm, 1.1-1.5cm and >1.5cm. Masses ≤0.5cm demonstrated a lower odds of malignancy compared to those 0.6-1.0cm (OR 0.20 95%CI [0.10; 0.40] p < 0.01) (14,15,24,25,27,28,30,37–39) (Figure 3). No significant difference was observed between masses 0.6-1.0cm and masses 1.1-1.5cm (15,25,28,37,39). Finally, comparison of masses 0.6-1.0cm and masses >1.5cm showed a significant lower odds of malignancy (OR 0.3 95%CI [0.09; 0.96], p = 0.04). A sensitivity analysis excluding studies that reported masses rather than patients and the three studies which included pre and post pubertal males was performed. Masses of ≤0.5cm still demonstrated a significant difference in odds of malignancy compared to those >0.6-1.0cm (OR 0.34 95%CI [0.14; 0.84], p = 0.02).

Heterogeneous vs. homogeneous

Three studies which included 51 patients reported on heterogenous and homogenous masses (9,11,16). Overall, 16 patients were reported to have heterogenous masses and 35 homogenous. A meta-analysis of the 3 studies found heterogenous masses had a lower OR of malignancy compared to homogenous masses (OR 0.15 95%CI [0.03;0.89] p = 0.04) (Figure 4). A sensitivity analysis after excluding the study of Shaaban et al. which reported

number of masses rather than number of patients still demonstrated that heterogenous masses had a lower OR of malignancy (OR 0.06 95%CI [0.01;0.60]).

Echogenicity

Echogenicity was reported in 10 studies which included 446 patients (10,11,13,16,25,29,32,33,35,36). We compared the radiological characteristics as hyperechoic vs. hypoechoic, hypoechoic vs. isoechoic and hyperechoic vs. isoechoic. Hyperechoic vs. hypoechoic masses included 385 patients: 51 hyperechoic and 334 hypoechoic, respectively. Hyperechoic masses had a lower OR of malignancy vs hypoechoic masses (OR 0.26 95%CI [0.11; 0.58] $p < 0.01$) (10,16,29,32,33,35,36) (Supplementary Figure 1). Hyperechoic masses also had a lower OR of malignancy vs. isoechoic masses. Of a total of 82 patients and five studies analysed, the OR was 0.25 95%CI [0.07; 0.83], $p = 0.02$ (27,28,33–35). Of a further 379 patients analysed, there was no significant difference in OR of malignancy between hypoechoic and isoechoic masses (10,11,13,16,25,29,32,35,36).

3.4.1.2 SUS colour flow doppler (CFD):

Doppler flow

Four studies which included 197 patients reported US Doppler flow (16,32,35,36). These were reported as either having peripheral flow, central flow, or mixed flow. We compared these as peripheral vs. central flow, mixed vs. central flow and peripheral vs. mixed flow. The meta-analysis demonstrated peripheral flow had a lower OR of malignancy compared to

either central (OR 0.09 95%CI [0.04;0.20] p < 0.01) (Supplementary Figure 2) or mixed flow (OR 0.17 95%CI [0.04;0.70] p = 0.01) (Supplementary Figure 3) (14,28,31,32). Three studies reporting on 83 patients did not demonstrate a significant difference between masses with mixed or central doppler flow (16,35,36). Sensitivity analysis excluding studies reporting number of masses rather than number of patients in the peripheral vs. central doppler flow cohort did not alter the outcome with peripheral doppler flow still demonstrating a lower OR of malignancy (OR 0.11 95%CI [0.04;0.3])p < 0.01.

Vascularity

Three studies which included 79 patients reported on vascularity on SUS. Each study compared different parameters (i.e., reduced vs. increased vascularity; normal vs. increased vascularity and normal vs. reduced vascularity) (11,12,26). As a result, a meta-analysis was not conducted. Shaaban et al. reported on 20 patients comparing the odds of malignancy between reduced and increased vascularity (11). None of the 11 patients with reduced vascularity had a malignant pathology whereas two of the nine with increased vascularity did. The OR was 0.13 but was not statistically significant (13). Auer et al. assessed the OR of malignancy in masses with normal vs. increased vascularity on SUS (26). Four of 42 masses with normal vascularity were malignant on histology compared to eight of 13 with increased vascularity. There was no significant difference between the two groups (22). Finally, Soh et al. compared normal vs. low vascularity. Only four patients were reported in this study and results were not significant (12).

3.4.1.3 Contrast enhanced ultrasound (CEUS):

Two studies which included 87 patients reported a comparison of normal vs. increased enhancement of testicular masses (27,34). Normal enhancement demonstrated a lower odds of malignancy vs. increased enhancement (OR 0.14 95%CI [0.005; 0.44], $p < 0.01$) (27,34). 74 patients across two studies compared increased vs. reduced enhancement however results were not significant (34,36). Only one study compared non enhancement vs. increased enhancement or non-enhancement vs. reduced enhancement (36) and a further study normal vs. reduced enhancement (34). Of those, Schwarze et al. reported on 46 patients comparing non enhancement vs. increased enhancement. This single study showed an OR of 0.09 95%CI [0.01;0.98] favouring benign pathology if the testicular mass was non enhancing (36). The same paper also reported on non-enhancement vs. reduced enhancement in a total of six patients. Neither OR was statistically significant. Luzurier et al. are the only group who reported on normal vs. reduced enhancement in indeterminate masses (34). Of 10 masses with normal enhancement, five were malignant whereas all 8 of 8 masses with reduced enhancement were malignant. The result however was not statistically significant (34).

3.4.2 Scrotal MRI

Khanna et al. compared scrotal mpMRI to differentiate benign sex cord stromal tumours from malignant (non-stromal and stromal) testicular neoplasms (43). Tumour size, T1 and T2

signal intensity, diffusion restriction, apparent diffusion coefficient and dynamic contrast enhancement were assessed. Malignant masses were more likely to be larger ($p < 0.01$) and demonstrate heterogenous enhancement patterns ($p < 0.02$). However, no cut offs were reported for size (43).

Manganaro et al. explored the role of DCE, DWI and semiquantitative and quantitative parameters to differentiate benign and malignant indeterminate masses of the testes (44). They assessed 47 patients comparing time to peak, percentage of peak enhancement, wash in rate, volume transfer constant and rate constant. Time to peak enhancement was shorter in the benign group compared to malignant ($p < 0.05$). They also reported higher values of percentage peak enhancement, wash in rate, volume transfer constant and rate constant with benign masses compared to malignant (45). Sanharawi et al. also compared quantitative and semiquantitative parameters derived from dynamic contrast enhancement on scrotal MRI (46). Overall, 31 patients were analysed comparing benign, burnt out and malignant tumours. Benign tumours demonstrated a shorter time to peak enhancement ($p = 0.0003$) with a higher peak ($p < 0.0001$) and higher initial enhancement slope ($p < 0.0001$) compared to burnt out or malignant masses (44). Benign masses also demonstrated a higher transfer constant ($p < 0.0001$) and rate constant ($p < 0.0001$) in comparison to other masses.

In a further study, Manganaro et al. evaluated the role of contrast enhanced scrotal MRI in identifying Leydig cell tumours in males with indeterminate testicular masses (45). They reported a sensitivity of 89.47%, specificity 95.65% for Leydig cell tumours compared to a sensitivity of 95.65% and 80.95% specificity for malignant masses (45). Overall, the diagnostic accuracy of identifying Leydig cell tumours compared to malignant masses was quoted at 93% (45).

3.4.3 Pre-pubertal SUS B mode

Three studies which included 77 patients with 78 masses reported on prepubertal males (21,41,42). All these studies were based on SUS B mode and/or CFD characteristics. Hoag et al. presented a case series of seven patients with benign pathology, six of whom were diagnosed on testis sparing surgery and the final showing resolution on radiological surveillance (42). Of these, four were >1.5cm, two 1.1cm - ≤1.5cm and one 0.6cm and ≤1.0cm. Chang et al. evaluated the role of clinical and sonographic features differentiating testicular teratomas and epidermoid cysts in 18 prepubertal males presenting with 19 testicular masses. Reported features included microlithiasis (yes/no), vascularity (yes/no) and solid/cystic or mixed and size in addition to age and pre-operative alpha foetoprotein (AFP) levels. They performed a receiver operating characteristics (ROC) analysis to obtain optimal cut off values. They demonstrated a significance in age ($p = 0.008$); however, they did not find any sonographic features that differentiated immature teratomas (41). They proposed children < 8 months an AFP level of 23ng/ml and 2.5 cm mass be considered for surgical intervention (41). Sensitivity and specificity were 100% and 89.5%, respectively (41).

3.5 Discussion

3.5.1 Principal findings

A total of 32 studies including 1692 masses were eligible for inclusion in this systematic review to determine which SUS or scrotal MRI characteristics can differentiate benign and malignant disease in indeterminate masses of the testes. These studies defined indeterminate masses as ≤2.0cm on SUS or scrotal MRI. Where size was not reported, a definition of a small testicular mass or non-palpable mass or focal mass incidentally detected was accepted. There was large variation among studies reporting on the different ultrasound radiological features, either per patient or per mass. There were also a limited number of studies on pre-pubertal males and scrotal MRI characteristics for any age cohort.

Our data demonstrates that specific radiological features on SUS can be diagnostic in differentiation of benign and malignant disease when faced with an indeterminate mass of the testes (Table 2). Mass size is an important factor, and this study demonstrates that masses ≤0.5cm have a significant lower OR of malignancy in comparison to any masses >0.5cm. Masses >0.5-≤1.0cm also had a significant difference compared to those >1.5cm. There however was no significance between masses of >0.5cm - ≤1.0 cm and >1.0cm - ≤1.5cm, or masses >1.0cm - ≤1.5cm and those >1.5cm. Intuitively it would be expected that the odds of malignancy should be lower with smaller masses.

In addition to size, heterogenous masses, hyper echogenicity, peripheral doppler flow and non-enhancement on SUS had a significant lower OR of malignancy. We did not find any significant demonstrable difference in OR of malignancy with other SUS features, such as microlithiasis or contour of masses. There were limited SUS studies reporting on vascularity.

Auer et al. did however demonstrate a lower odds of malignancy in indeterminate masses with a normal vascularity compared to those with increased vascularity (26).

Only four studies (i.e., 122 patients) reported on scrotal MRI outcomes. There was significant variation in reporting of quantitative characteristics. Of note two studies - Manganaro et al. and Sanharawi et al. - both reported on DCE outcomes of time to peak enhancement and the maximal peak enhancement. In both studies their results favoured benign disease with indeterminate masses demonstrating a shorter time to peak enhancement and a higher maximal peak (44,46).

Only 3 studies reported on indeterminate masses in 77 pre-pubertal males (21,41,42). None of these papers reported on any significant findings with regards to either SUS or scrotal MRI characteristics that could distinguish between benign and malignant disease.

3.5.2 Implications for clinical practice

In the presence of a normal contralateral testis, radical inguinal orchiectomy is the standard of care for all males diagnosed with a testicular mass suspicious for malignancy or an indeterminate testicular mass, despite a significant risk of over treatment (7,8). However, this must be balanced with the risks of under treatment (e.g., surveillance with the potential effects of delayed treatment on survival and the psychological burden of disease surveillance). Both the EAU and AUA recommend the use of TSS in masses ≤2.0cm, for patients wishing to preserve gonadal function or indeterminate masses on imaging with normal tumour markers or solitary testes (3,4,47,48). These guidelines suggest that TSS and frozen section (FSE) is reliable and has an accurate concordance with final histology (3,4,47,48). Therefore, TSS could be considered for the indeterminate testicular mass

provided the patient is fully counselled regarding immediate or delayed risk of orchidectomy which include a high risk of local recurrence and the need for completion RO if malignant pathology is found on final histopathological diagnosis (7,49).

Size alone is one of the main objective radiological features reported on US and has least inter-observer variability compared to other radiological features (e.g., heterogenicity). For masses less than 0.5cm where OR of malignancy is low, as suggested from the results in this study, patients could be offered less radical treatment i.e., TSS, enucleation or potentially surveillance. This is in accordance with the findings by Berney et al. where up to 94% of masses less than 0.5 cm were found to be benign at the time of RO. Numbers in this study however were small. In their study only 1 of 16 patients with a mass <0.5cm was found to have a germ cell tumour. On further review, this was found to be a partially regressed tumour with an 18mm area of granulation inflammation which would have appeared larger on SUS. 15 of the 16 patients in this cohort underwent RO (7). TSS or possibly even surveillance may have been a preferential option for this cohort of patients.

TSS should also be considered for masses between 0.6-1.0cm. This would avoid the need for RO in up to 69% of cases where histology for indeterminate masses ≤ 1.0cm were found to be benign (7). Presence of other characteristics demonstrating a lower OR of malignancy (i.e., heterogenous masses with hyper echogenicity, normal enhancement, and peripheral vascularity) may further support this management strategy (Table 2).

Based on the results of this meta-analysis, indeterminate masses >1cm should be treated by surgical intervention either by TSS (if technically feasible and with negative tumour markers) or RO. The presence of additional radiological features such as hypo echogenicity,

homogeneity, central doppler flow and increased enhancement may further aid diagnostic value.

Our data suggests that a criteria-based approach based on SUS characteristics assessing size, heterogeneity, echogenicity, doppler flow and enhancement may aid in clinical decision-making and differentiate between indeterminate testicular masses that may be managed with surveillance (<0.5cm where 90% of masses are benign) or those with a higher risk of malignancy in whom obtaining histology is mandated. In those in whom surveillance is undertaken, we would recommend a stringent imaging protocol with a threshold of a change and/ or development of the aforementioned SUS characteristics as an indication to perform surgical intervention (4). It is also best practice that all these cases are discussed in a specialist multi-disciplinary team meeting (50).

3.5.3 How the review compares to previous reviews/guidelines

There are no systematic reviews investigating test accuracy measures of indeterminate testicular masses. However, the EAU 2021 and AUA 2019 testicular cancer guidelines suggests that TSS can be offered where feasible in cases of small or indeterminate testicular masses on imaging with negative tumour markers to avoid overtreatment and preserve testicular function, however no established and unequivocal size or radiological criteria exist for which masses can be safely monitored (3,4). Multiple studies have suggested that size is an important factor in indeterminate testicular masses, with up to 69% of ≤ 1.0cm and up to 94% of masses ≤0.5cm quoted as benign (6,7). Some studies suggest hyperechoic non-vascular masses with normal tumour markers reduce the risk of underlying malignancy,

whereas hypoechoic and vascular masses are more likely to be malignant (4,31) and this meta-analysis supports this data.

3.5.4 Future research

This systematic review and meta-analysis has demonstrated SUS characteristics which can be used to stratify those patients with indeterminate testicular masses who may benefit from intervention but also reduce the risk of overtreatment with radical surgery for benign disease. Further studies are required to not only assess the role of scrotal MRI either as a diagnostic test or as an adjunct with SUS and other more novel imaging modalities such as CEUS. Importantly, we have been unable to draw firm conclusions regarding the difference between masses >0.5cm ≤1.0cm and those >1.0cm with regards to OR of malignancy. A large prospective study exploring size as a continuous variable would help determine the optimum size threshold to exclude malignancy. This with the different SUS characteristics may help facilitate our further understanding of the risk of malignancy in this group of patients.

3.5.5 Strengths and limitations

The clinical question and outcomes were developed in conjunction with the EAU Sexual and Reproductive Health, Testicular Cancer and Paediatric urology panels on this important and difficult clinical scenario. The review was performed in accordance with PRISMA guidelines and Cochrane methodology. The searches were carried out and additional sources of

studies such as the reference lists were explored by the reviewers to ensure a comprehensive review of the literature. Most of the studies included demonstrated a low risk of bias and low concerns for applicability to the review question.

Our study however does have significant limitations. The screening data period was large with interobserver variability and likely differences in the interpretation of results due to improvements in SUS and scrotal MRI capabilities. The number of studies reporting on some imaging modalities were too small to draw any meaningful conclusion.

Conclusion

This study provides important and cogent information on radiological features that may aid clinicians in managing patients with indeterminate masses of the testis.

Patient Summary

We reviewed the available evidence on scrotal ultrasound and magnetic resonance imaging findings that may predict whether indeterminate masses of the testis are cancerous or not. These findings may be of benefit in avoiding radical surgery in patients who may be placed on imaging surveillance or undergo testis sparing surgery where technically feasible.

Conflicts of interest:

N/A

Author Contributions:

Michael Ager – Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization

Sarah Donegan – Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization

Luca Boeri - Data Curation, Writing - Review & Editing

Javier Mayor de Castro – Validation, Investigation, Data Curation, Writing - Review & Editing

James F Donaldson – Conceptualization, Methodology, Validation, Formal analysis, Data Curation, Writing - Review & Editing, Supervision

Muhammad Imran Omar – Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing

Konstantinos Dimitropoulos – Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing

Tharu Tharakan - Data Curation

Florian Janisch - Data Curation

Tim Muilwijk - Data Curation

Cathy Yuan- Resources

Catrin Tudur-Smith – Methodology, Formal analysis, Writing - Review & Editing

Rien JM Nijman – Conceptualization, Writing - Review & Editing

Christian Radmayr – Conceptualization, Writing - Review & Editing

Andrea Salonia – Conceptualization, Validation, Methodology, Writing - Review & Editing, Supervision

Maria P Laguna Pes – Conceptualization, Validation, Methodology, Writing - Review & Editing, Supervision

Suks Minhas – Conceptualization, Validation, Methodology, Writing - Review & Editing, Supervision

References:

1. Huang DY, Sidhu PS. Focal testicular lesions: Colour Doppler ultrasound, contrast enhanced ultrasound and tissue elastography as adjuvants to the diagnosis. British Journal of Radiology. 2012 Nov 1;85(SPEC. ISSUE 1).
2. Jones HR, Vasey A. Part 1: Testicular cancer-management of early disease. The Lancet Oncology. 2003;4:730–7.
3. Algaba F, Bokemeyer C, Boormans J, Fischer S, Fizazi K, Gremmels H, et al. Testicular Cancer EAU Guidelines on. 2021.
4. Stephenson A, Eggener SE, Bass EB, Chelnick DM, Daneshmand S, Feldman D, et al. Diagnosis and Treatment of Early-Stage Testicular Cancer: AUA Guideline. Journal of Urology. 2019 Aug 1;202(2):272–81.
5. Soon JA, Anton A, Torres J, Lawrence R, Parente P, McKendrick J, et al. Exploring the spectrum of late effects following radical orchidectomy for stage I testicular seminoma: a systematic review of the literature. Vol. 27, Supportive Care in Cancer. Springer Verlag; 2019. p. 373–82.
6. Metcalfe PD, Farivar-Mohseni H, Farhat W, McLorie G, Khoury A, Bägli DJ. Pediatric testicular tumors: Contemporary incidence and efficacy of testicular preserving surgery. Journal of Urology. 2003;170(6 I):2412–6.
7. Scandura G, Verrill C, Protheroe A, Joseph J, Ansell W, Sahdev A, et al. Incidentally detected testicular lesions <10 mm in diameter: can orchidectomy be avoided? BJU International. 2018 Apr 1;121(4):575–82.
8. Abboudi H, Malde S, Mchaourab A, Eddy B, Shrotri N. Nonpalpable Testicular Masses—Should We Be Worried? Open Journal of Urology. 2013;03(07):281–6.

9. Passarella M, Usta MF, Bivalacqua TJ, Hellstrom WJG, Davis R. Testicular-sparing surgery: a reasonable option in selected patients with testicular lesions. 2003.

10. Konstantatou E, Fang C, Romanos O, Derchi LE, Bertolotto M, Valentino M, et al. Evaluation of intratesticular lesions with strain elastography using strain ratio and color map visual grading: Differentiation of neoplastic and nonneoplastic lesions. Journal of Ultrasound in Medicine. 2019 Jan 1;38(1):223–32.

11. Shaaban MS. Use of strain sonoelastography in differentiation of focal testicular lesions. Egyptian Journal of Radiology and Nuclear Medicine. 2017 Jun 1;48(2):485–91.

12. Soh E, Berman LH, Grant JW, Bullock N, Williams M v. Ultrasound-guided core-needle biopsy of the testis for focal indeterminate intratesticular lesions. European Radiology. 2008;18(12):2990–6.

13. Müller T, Gozzi C, Akkad T, Pallwein L, Bartsch G, Steiner H. Management of incidental impalpable intratesticular masses of ≤5 mm in diameter. BJU International. 2006 Nov;98(5):1001–4.

14. Pastore AL, Palleschi G, Maceroni P, Manfredonia G, Autieri D, Cacciotti J, et al. Correlation between semiquantitative sonoelastography and immunohistochemistry in the evaluation of testicular focal lesions. Cancer Imaging. 2014 Dec 1;14(1).

15. Galosi AB, Fulvi P, Fabiani A, Servi L, Filosa A, Leone L, et al. Testicular sparing surgery in small testis masses: A multinstitutional experience. Archivio italiano di urologia, andrologia : organo ufficiale [di] Societa italiana di ecografia urologica e nefrologica. 2016 Dec 30;88(4):320–4.

16. Reginelli A, D'Andrea A, Clemente A, Izzo A, Urraro F, Scala F, et al. Does multiparametric US improve diagnostic accuracy in the characterization of small testicular masses? Gland Surgery. 2019 Sep 1;8:S136–41.

17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. The BMJ. 2015 Oct 28;351.

18. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Chapter 10 Analysing and Presenting Results [Internet]. Available from: http://srdta.cochrane.org/.

19. Knoll T, Omar MI, Maclennan S, Hernández V, Canfield S, Yuan Y, et al. Key Steps in Conducting Systematic Reviews for Underpinning Clinical Practice Guidelines: Methodology of the European Association of Urology. European Urology. 2018 Feb 1;73(2):290–300.

20. Mayor de Castro J, Ager M, Donaldson J, Dimitropoulos K, Imran Omar M, Yuhong Yuan C. PROSPERO International prospective register of systematic reviews Citation [Internet]. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020199339

21. Karmazyn B, Weatherly DL, Lehnert SJ, Cain MP, Fan R, Jennings SG, et al. Characteristics of testicular tumors in prepubertal children (age 5–12 years). Journal of Pediatric Urology. 2018 Jun 1;14(3):259.e1-259.e6.

22. The Cochrane Collaboration. Review Manager (RevMan) [Computer program]. Version 5.4. 2020.

23. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies [Internet]. 2011. Available from: www.annals.org

24. Goddi A, Sacchi A, Magistretti G, Almolla J, Salvadore M. Real-time tissue elastography for testicular lesion assessment. European Radiology. 2012 Apr;22(4):721–30.

25. Ates F, Malkoc E, Zor M, Demirer Z, Alp BF, Basal S, et al. Testis-Sparing Surgery in Small Testicular Masses Not Suspected to Be Malignant. Clinical Genitourinary Cancer. 2016 Feb 1;14(1):e49–53.

26. Auer T, de Zordo T, Dejaco C, Gruber L, Pichler R, Jaschke W, et al. Value of multiparametric US in the assessment of intratesticular lesions. Radiology. 2017 Nov 1;285(2):640–9.

27. Avci A, Erol B, Eken C, Ozgok Y. Nine cases of nonpalpable testicular mass: An incidental finding in a large-scale ultrasonography survey. International Journal of Urology. 2008 Sep;15(9):833–6.

28. Ayati M, Ariafar A, Jamshidian H, Soleimani A, Ghasemi F, Nowroozi MR. Management of Nonpalpable Incidental Testicular Masses: Experience with 10 Cases.

29. Cayetano-Alcaraz AA, Halilovic M, Olavarria-Sayavedra N, Quiñones-Capistrán CA, Arceo-Olaiz R, Castillejos-Molina RA. Diagnostic accuracy and complications of frozen section examination of equivocal malignant testicular masses. Revista Mexicana de Urologia. 2018;78(2):128–34.

30. Colpi GM, Carmignani L, Nerva F, Guido P, Gadda F, Castiglioni F. Testicular-sparing microsurgery for suspected testicular masses. BJU International. 2005 Jul;96(1):67–9.

31. Dell'atti L, Fulvi P, Galosi AB. Are ultrasonographic measurements a reliable parameter to choose non-palpable testicular masses amenable to treatment with sparing surgery? JBUON. 2018;23(2):439–43.

32. Isidori AM, Pozza C, Gianfrilli D, Giannetta E, Lemma A, Pofi R, et al. Differential diagnosis of nonpalpable testicular lesions: Qualitative and quantitative contrast-enhanced US of benign and malignant testicular tumors. Radiology. 2014 Nov 1;273(2):606–18.

33. Li Q, Vij A, Hahn PF, Xiang F, Samir AE. The value of active ultrasound surveillance for patients with small testicular lesions. Ultrasound Quarterly. 2017;33(1):23–7.

34. Luzurier A, Maxwell F, Correas JM, Benoit G, Izard V, Ferlicot S, et al. Qualitative and quantitative contrast-enhanced ultrasonography for the characterisation of non-palpable testicular tumours. Clinical Radiology. 2018 Mar 1;73(3):322.e1-322.e9.

35. Rocher L, Ksouri A, Maxwell F, Bresson B, Hindawi G, Balasa C, et al. Testicular tumors: A diagnostic challenge of imaging. Vol. 106, Bulletin du Cancer. John Libbey Eurotext; 2019. p. 875–86.

36. Schwarze V, Marschner C, Sabel B, de Figueiredo GN, Marcon J, Ingrisch M, et al. Multiparametric ultrasonographic analysis of testicular tumors: a single-center experience in a collective of 49 patients. Scandinavian Journal of Urology. 2020 May 3;54(3):241–7.

37. Toren PJ, Roberts M, Lecker I, Grober ED, Jarvi K, Lo KC. Small Incidentally Discovered Testicular Masses in Infertile Men-Is Active Surveillance the New Standard of Care? Journal of Urology. 2010 Apr;183(4):1373–7.

38. Schwen ZR, Liu JL, Gabrielson AT, Patel HD, Gupta M, Rowe SP, et al. Testicular ultrasound underestimates the size of small testicular masses: a radiologic–pathologic correlation study. World Journal of Urology. 2021 Sep 1;39(9):3399–405.

39. Staudacher N, Tulchiner G, Bates K, Ladurner M, Kafka M, Aigner F, et al. Organ-sparing surgery in testicular tumor: Is this the right approach for lesions ≤ 20 mm? Journal of Clinical Medicine. 2020 Sep 1;9(9):1–11.

40. Gentile G, Rizzo M, Bianchi L, Falcone M, Dente D, Cilletti M, et al. Testis sparing surgery of small testicular masses: Retrospective analysis of a multicenter cohort. Journal of Urology. 2020 Apr 1;203(4):760–6.

41. Chang MY, Shin HJ, Kim HG, Kim MJ, Lee MJ. Prepubertal testicular teratomas and epidermoid cysts: Comparison of clinical and sonographic features. Journal of Ultrasound in Medicine. 2015 Oct 1;34(10):1745–51.

42. Hoag NA, Afshar K, Youssef D, Masterson JST, Murphy J, MacNeily AE. Cystic intratesticular lesions in pediatric patients. Journal of Pediatric Surgery. 2013 Aug;48(8):1773–7.

43. Khanna M, Abualruz AR, Yadav SK, Mafraji M, Al-Rumaihi K, Al-Bozom I, et al. Diagnostic performance of multi-parametric MRI to differentiate benign sex cord stromal tumors from malignant (non-stromal and stromal) testicular neoplasms. Abdominal Radiology. 2021 Jan 1;46(1):319–30.

44. Manganaro L, Saldari M, Pozza C, Vinci V, Gianfrilli D, Greco E, et al. Dynamic contrast-enhanced and diffusion-weighted MR imaging in the characterisation of small, non-palpable solid testicular tumours. European Radiology. 2018 Feb 1;28(2):554–64.

45. Manganaro L, Vinci V, Pozza C, Saldari M, Gianfrilli D, Pofi R, et al. A prospective study on contrast-enhanced magnetic resonance imaging of testicular lesions: distinctive features of Leydig cell tumours. European Radiology. 2015 Dec 1;25(12):3586–95.

46. Sanharawi I el, Correas JM, Glas L, Ferlicot S, Izard V, Ducot B, et al. Non-palpable incidentally found testicular tumors: Differentiation between benign, malignant, and burned-out tumors using dynamic contrast-enhanced MRI. European Journal of Radiology. 2016 Nov 1;85(11):2072–82.

47. Matei DV, Vartolomei MD, Renne G, Tringali VML, Russo A, Bianchi R, et al. Reliability of Frozen Section Examination in a Large Cohort of Testicular Masses: What Did We Learn? Clinical Genitourinary Cancer. 2017 Aug 1;15(4):e689–96.

48. Elert A, Olbert P, Hegele A, Barth P, Hofmann R, Heidenreich A. Accuracy of Frozen Section Examination of Testicular Tumors of Uncertain Origin. Vol. 41, European Urology. 2002.

49. Nason GJ, Aditya I, Leao R, Anson-Cartwright L, Jewett MAS, O'Malley M, et al. Partial orchiectomy: The Princess Margaret cancer centre experience. Urologic Oncology: Seminars and Original Investigations. 2020 Jun 1;38(6):605.e19-605.e24.

50. Shamash J, Ansell W, Alifrangis C, Thomas B, Wilson P, Stoneham S, et al. The impact of a supranetwork multidisciplinary team (SMDT) on decision-making in testicular cancers: a 10-year overview of the Anglian Germ Cell Cancer Collaborative Group (AGCCCG). British Journal of Cancer. 2021 Jan 19;124(2):368–74.

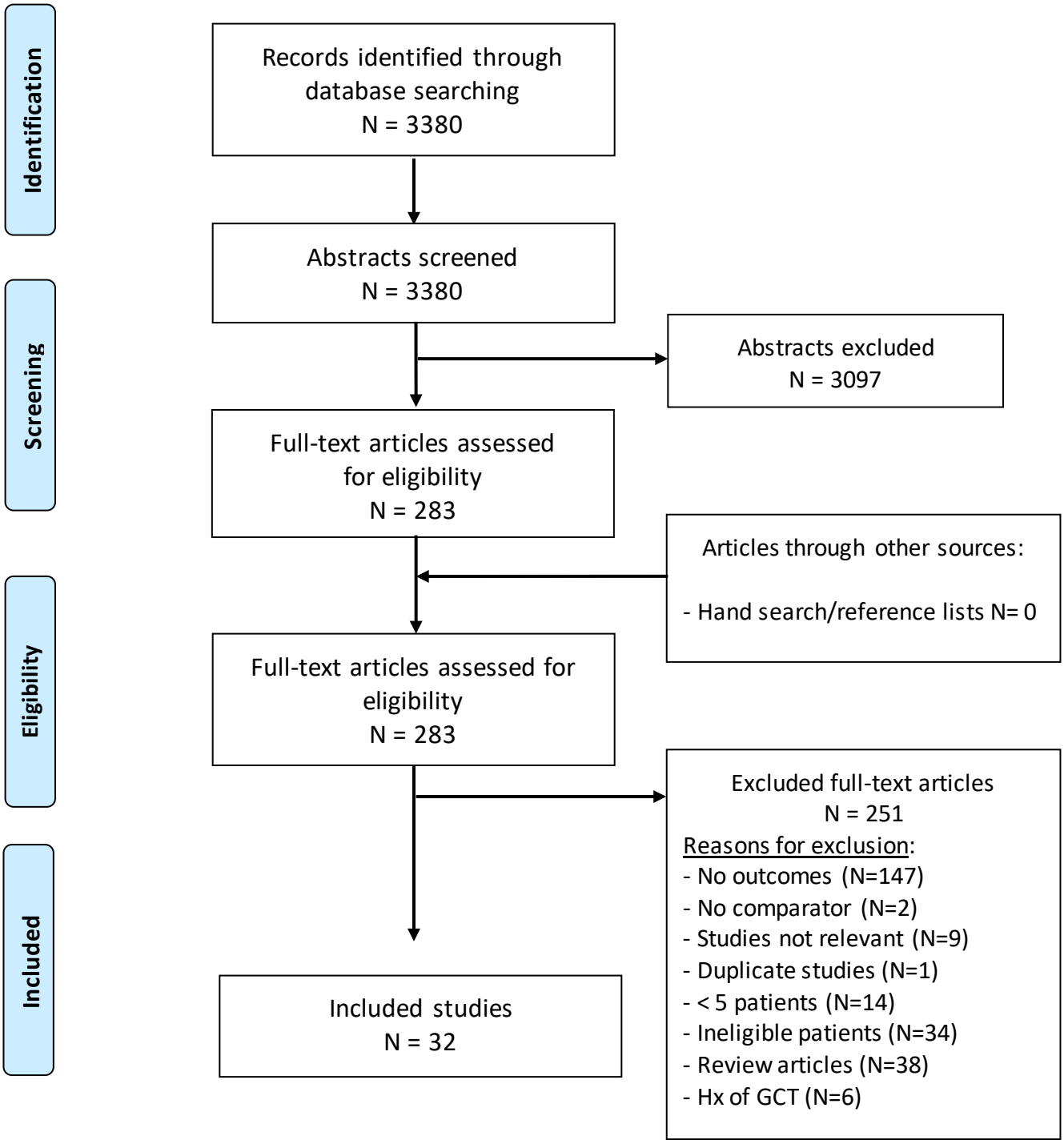

Fig 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of studies identified, included, and excluded.

GCT: germ cell tumour, N: number.

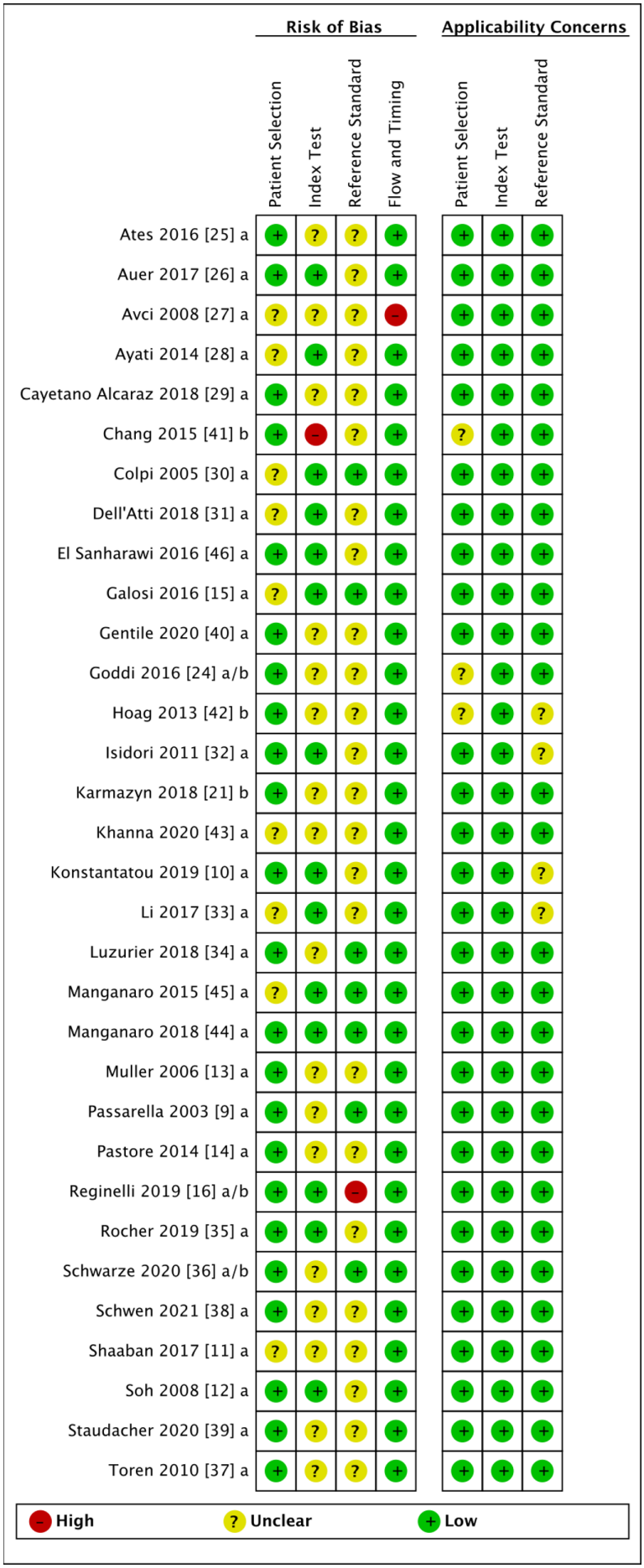

Figure 2. Summary of risk of bias assessment and applicability concerns (22)**.**

a – post pubertal studies, b – pre pubertal studies, a/b – mixed population (pre and post pubertal).

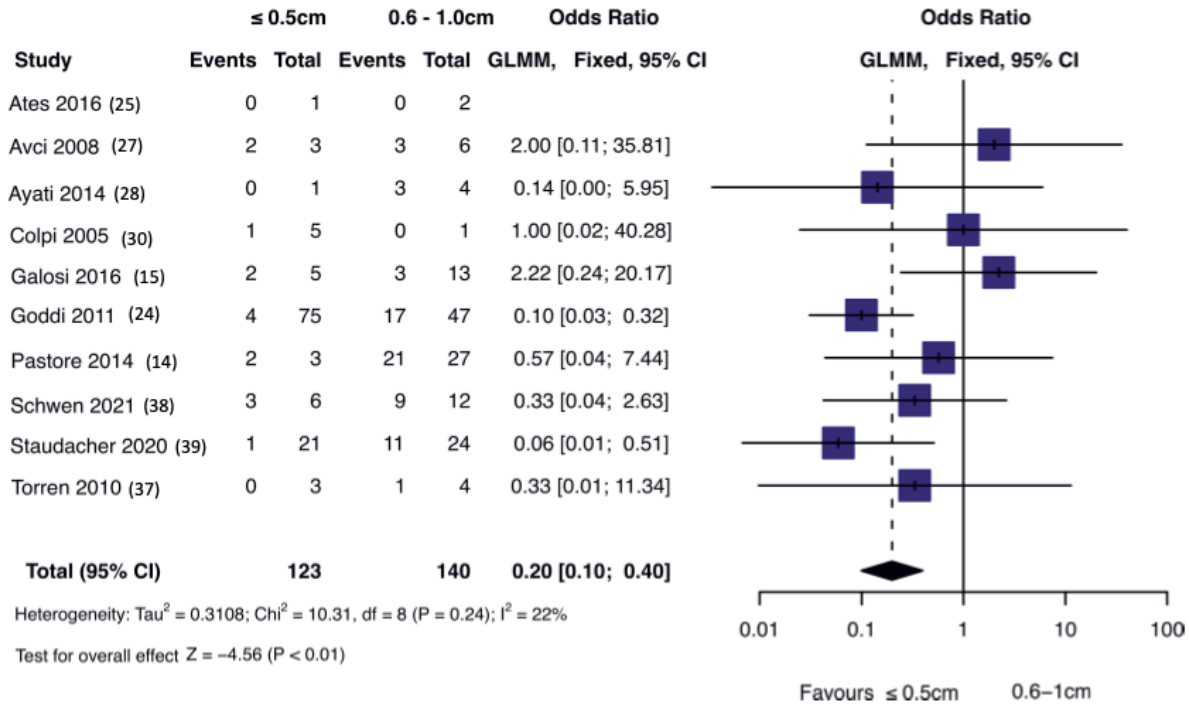

Figure 3: Forest plot showing results of OR of malignancy for individual studies and meta-analysis of combined studies for masses ≤0.5cm vs 0.6cm - ≤1.0cm on SUS.

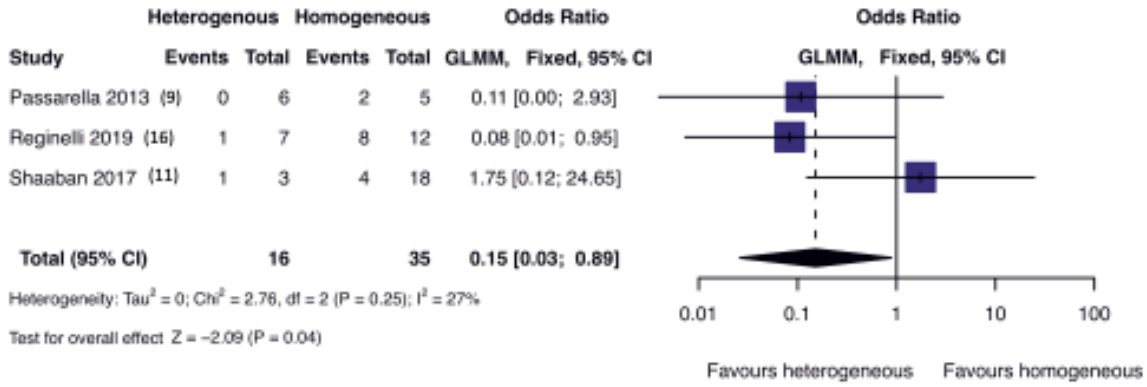

Figure 4: Forest plot showing results of OR of malignancy of individual studies and meta-analysis of combined studies for heterogenous masses vs homogeneous masses on SUS.

Author	Country	Study type	N patients (N masses)	Age: Mean (SD), *Median [Range]	Target condition	Index test	Reference test			Benign	TGCT	Lymphoma / metastasis	
							TSS	RO	FU				
SUS post pubertal males													
Konstantatou 2019 (10)	UK	Retrospective cohort	86 (86)	36 [16-81]	'Indeterminate focal intratesticular mass'	SUS	0	52	34	31	55	0	
Shaaban 2017 (11)	Egypt	Prospective cohort	21 (23)	30 [18-54]	'Focal testicular masses'	SUS	0	7	14	16	5	0	
Soh 2008 (12)	UK	Prospective cohort	5 (5)	42.2 [28-56]	'Focal indeterminate mass on SUS'	SUS	0	3	2	2	3	0	

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

Muller 2006 [(13)	Austria	Prospective cohort	20 (20)	36.4 [26-58]	≤0.5cm	SUS	16	4	0	16	4	0
Pastore 2014 (14)	Italy	Prospective cohort	30(30)	37 [22-50]	≤1.0cm	SUS	0	30	0	23	7	0
Galosi 2016 (15)	Italy	Prospective cohort	28 (28)	38 [18-68]	≤1.5cm	SUS	17	11	0	22	6	0
Passarella 2003 (9)	USA	Retrospective Cohort	11 (11)	43.18 [27-63]	Lesion unable to be defined by size, history, or SUS feature with normal	SUS	4	7	0	9	2	0

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

					tumour markers							
Ates 2016 (25)	Turkey	Retrospective cohort	15 (15)	24.22 [20-36]	"Small testicular masses"	SUS	1	14	0	14	1	0
Auer 2017 (26)	Austria	Retrospective cohort	55 (55)	*39.5 (+/-14.9)	'Focal testicular lesions indeterminate on gray scale'	SUS	0	24	31	43	12	0
Avci 2008 (27)	Turkey	Retrospective cohort	9 (9)	20. 24 [19-33]	Non palpable masses	SUS	0	9	0	4	5	0

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

Ayati 2014 (28)	Iran	Retrospective cohort	10 (10)	32.2 [21-54]	Non palpable < 2.0cm	SUS	4	6	0	4	6	0
Cayetano Alcaraz 2018 (29)	Mexico	Retrospective cohort	23 (23)	*37 (+/-13)	'Equivocal ,malignant testicular masses'	SUS	13	10	0	17	4	2
Colpi 2005 (30)	Italy	Retrospective cohort	6 (6)	33.1 [34-42]	Incidental SUS detected masses	SUS	1	1	4	5	1	0
Dell'Atti 2018 (31)	Italy	Retrospective cohort	77 (77)	36.5 [22-74]	Non palpable testicular masses	SUS	37	40	0	28	47	2

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

Isidori 2011 (32)	Italy	Prospective cohort	115 (122)	34 [28-40]	≤ 1.5cm	SUS	47	43	25 (32)	70 (77)	44	1
Li 2017 (33)	USA	Retrospective cohort	101(101)	42 [18-91]	≤1.0cm	SUS	3	22	76	86	14	1
Luzurier 2018 (34)	France	Prospective cohort	40 (40)	35.5 [20-58]	Non palpable masses with normal tumour markers	SUS	15	25	0	16	24	0
Rocher 2019 (35)	France	Prospective cohort	86 (89)	*37.9 (+/-13.2)	Small incidental	SUS	0	81 (82)	5 (7)	38 (41)	47 (48)	1

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

					testicular masses							
Toren 2010 (37)	Canada	Retrospective cohort	46 (56)	35 [21-71]	≤1.0cm	SUS	7	1	38	45	1	0
Schwen 2021 (38)	USA	Retrospective cohort	208 (208)	32 [26-42]	Nonpalpable small testicular mass	SUS	10	98	0	22	186	0
Staudacher 2020 (39)	Austria	Retrospective cohort	89 (99)	*38.4 (+/-16.2)	≤2.0cm negative	SUS	67	32	0	(67)	(32)	0

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

					tumour markers							
Gentile 2020 (40)	Italy	Retrospective cohort	147 (147)	*35 (+/-11)	≤2.0cm normal tumour markers	SUS	147	0	0	126	21	0
SUS pre and post pubertal males												
Reginelli 2019 (16)	Italy	Retrospective cohort	19 (19)	42.2 [10-64]	≤2.0cm	SUS	2	11	6	10	9	0
Goddi 2011 (24)	Italy	Prospective cohort	88 (testicles) (144)	34 [2months to 89 years]		SUS	0	33	11 1	112	32	0

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

					'Focal testicular masses'							
Schwarze 2020 (36)	German y	Retrospective cohort	49 (49)	46 [7-80]	'Unknown testicular masses'	SUS	48	0	1	13	31	5
Scrotal MRI post pubertal males												
Khanna 2021 (43)	USA	Retrospective cohort	20 (20)	STs: *30.9 (+/-14.2) MTN:	Indeterminate testicular mass	MRI	0	20	0	11	7	2

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

				*35.2 (+/-15.3)								
Manganaro 2018 (44)	Italy	Prospective cohort	47 (47)	36 [27-41]	≤1.5cm	MRI	24	23	0	25	21	1
Manganaro 2015 (45)	Italy	Prospective cohort	44 (44)	*34.45 (+/-8.97)	Non palpable testicular lesion	MRI	23	21	0	21	23	0
El Sanharawi 2016 (46)	France	Retrospective cohort	31 (31)	*37.3 (+/-7.5)	non-palpable, incidental testicular tumours	MRI	0	31	0	12	19	0
SUS pre pubertal males												

Table 1: Baseline characteristics for post pubertal SUS (n=22), mixed pre and post pubertal SUS (n=3), post pubertal scrotal MRI studies (n=4), pre pubertal SUS (n = 3).

TSS – Testis sparing surgery, RO – Radical orchidectomy, FU – Follow up (radiological)

Chang 2014 (41)	South Korea	Retrospective cohort	18 (19)	13.5 months (1 – 64)	Prepubertal testicular masses	SUS	0	18 (19)	0	16	3	0
Karmazyn 2018 (21)	USA	Retrospective cohort	52 (52)	264 days (1 day to 18 years)	SUS diagnosed testicular masses	SUS	11	41	0	27	25	0
Hoag 2012 (42)	Canada	Retrospective case series	7 (7)	6.6 months (0-15)	Intratesticular cysts	SUS	5	1	1	7	0	0

Comparison	Statistical heterogeneity across studies	N Studies (N masses)	Odds ratio (95%CI)	P value
(≤0.5cm) vs. (>0.5cm - ≤1.0cm)	Chi^2 = 10.31, df = 8 (P = 0.24); I^2 = 22%	10(163)	0.20[0.10; 0.40]	<0.01
(≤0.5cm) vs (>1.0cm - ≤1.5cm)	Chi^2 = 5.45, df = 3 (P = 0.14); I^2 = 45%	5(77)	0.29[0.07; 1.16]	0.06
(≤0.5cm) vs. (>1.5cm)	Chi^2 = 0.00, df = 1 (P = 1.00); I^2 = 0%	3(48)	0.03[0; 0.25]	<0.01
(>1.0cm - ≤1.5cm) vs. (>0.5cm-≤1.0cm)	Chi^2 = 0.26, df = 4 (P = 0.99); I^2 = 0%	5(94)	1.77[0.70; 4.50]	0.23
(>0.5-1cm) vs (>1.5cm)	Chi^2 = 0.00, df = 1 (P = 1.00); I^2 = 0%	3(65)	0.30[0.09; 0.96]	0.04
(>1.0-≤1.5cm) vs (>1.5cm)	Chi^2 = 0.00, df = 2 (P = 1.00); I^2 = 0%	3(61)	0.40(0.10; 1.17)	0.12

Table 2: Summary of results from logistic regression meta-analysis of comparison of each reported SUS characteristic.

Heterogenous vs homogenous	$Chi^2 = 2.76$, df = 2 (P = 0.25); $I^2 = 27\%$	3(51)	0.15[0.03; 0.89]	0.04
Hyperechoic vs hypoechoic	$Chi^2 = 0.00$, df = 6 (P = 1.00); $I^2 = 0\%$	7(385)	0.26[0.11; 0.58]	< 0.01
Hyperechoic vs isoechoic	$Chi^2 = 0.00$, df = 5 (P = 1.00); $I^2 = 0\%$	6(82)	0.25[0.07; 0.83]	0.02
Hypoechoic vs isoechoic	$Chi^2 = 2.74$, df = 8 (P = 0.95); $I^2 = 0\%$	9(379)	0.77[0.42; 1.43]	0.41
No microlithiasis vs microlithiasis	$Chi^2 = 25.53$, df = 5 (P < 0.01); $I^2 = 80\%$	6(112)	0.64[0.11; 3.69]	0.62
Regular vs irregular margins	$Chi^2 = 20.86$, df = 3 (P < 0.01); $I^2 = 86\%$	4(167)	0.34[0.06;1.93]	0.22
Colour flow doppler				
Peripheral vs central doppler flow	$Chi^2 = 0.05$, df = 3 (P = 1.00); $I^2 = 0\%$	4(197)	0.09[0.04; 0.20]	< 0.01
Peripheral vs mixed doppler flow	$Chi^2 = 0.00$, df = 2 (P = 1.00); $I^2 = 0\%$	3(75)	0.17[0.04; 0.70]	0.01

Table 2: Summary of results from logistic regression meta-analysis of comparison of each reported SUS characteristic.

Mixed vs central doppler flow	$Chi^2 = 0.54$, df = 2 (P = 0.77); $I^2 = 0\%$	3(83)	0.70[0.17; 2.96]	0.63
Contrast enhanced SUS				
Enhancement: normal vs increased	$Chi^2 = 0.00$, df = 1 (P = 1.00); $I^2 = 0\%$	2(87)	0.14[0.05; 0.44]	< 0.01
Enhancement increased vs delayed	$Chi^2 = 0.00$, df = 1 (P = 1.00); $I^2 = 0\%$	2(74)	0.15[0.02; 1.38]	0.09