Reviewer #4: The authors are still failing to take the problem of heterogeneity seriously. The latest response was "well, yes, the results are all over the place, but most results are pretty good". First off, it is not at all clear to this reviewer that a negative predictive value of 85%, the threshold used by the authors, is acceptable. Would most patients be happy to be told that they only have a 15% risk of high-grade cancer, so not to worry about it? That is approximately the risk for a PSA of 10. How many urologists would argue against giving a biopsy to a man with a PSA of 10?

Second, the point about heterogeneity is absolutely not to find out where the majority of results lie. For instance, if this was a meta-analysis of complication risk after abortion, and there was gross heterogeneity, with most clinics having very low rates and a few having very high rates, I'm pretty sure the conclusion wouldn't be only that "abortion is a safe procedure". More likely, there would be something about there being some unsafe clinics, and then an analysis to determine the characteristics of those clinics (e.g. solo practice) or a call for further research to determine why some clinics have poor outcomes. The current paper reads as if the authors came up with the conclusion first and then went through the motions of reporting forest plots and I2 statistics and so on. The conclusion that "Multi-parametric MRI of the prostate is an accurate test for ruling-out clinically significant prostate cancer" is simply not an appropriate reflection of the data presented by the authors.

Thank you to the reviewer for these further helpful comments, particularly regarding heterogeneity. We do want to make clear that the purpose of this paper has never been to prove that MRI is good. The purpose instead has been to provide a pooled statistic which can be used by clinicians in this counselling discussion with patients which you so helpfully refer to above in your first paragraph above. Indeed, in order to do this, we have intentionally included 'real-world data' here which is necessarily more varied. Furthermore, we provide no steer as to whether 9.2% (the key percentage chance that we "have missed clinically significant disease" according to most commonly accepted definitions) means "not to worry about it". Some surgeons will argue to proceed to biopsy even with a 'negative' MRI on this basis – as do several of my colleagues. Others will argue that it is an acceptable risk and biopsy is not needed. The key is that we help the patient make an informed decision based on intelligent discussion using real-world data.

We have addressed each of your comments and made some further detailed changes to the manuscript.

Formatted: Font: Bold

I would recommend the following to the authors

1. Include reference to the heterogeneity in the abstract.

We thank the reviewer for this comment and this has ~~already~~ been ~~done~~ ammended in the conclusion part of the abstract: *"However, we observed heterogeneity in the NPV estimates and local institutional data should form the basis of decision-making if available."*

Commented [AL1]: I know we'd already done it but we don't need to point that out to AV. Hopefully he'll think we've now spotted the wisdom of his words and made the change requested. Don't want to give him any reason to dig his heals in.

2. Change "MRI is an accurate test" to something like "MRI can be an accurate test when implemented in at least some centers" (I'm not joysticking here, the authors should choose

Commented [AL2]: Never heard this term used before…any ideas?

their own words to make reference to the results, which is that it clearly works in some settings, but has insufficient properties in others).

We thank the reviewer for their comment and have adjusted both the conclusion in the abstract and text accordingly.

Abstract: ***"mpMRI of the prostate is generally an accurate test for ruling-out clinically significant prostate cancer. However, we observed heterogeneity in the NPV estimates and local institutional data should form the basis of decision-making if available."***

Text: ***"Multi-parametric MRI of the prostate can be an accurate test for ruling-out clinically significant prostate cancer when implemented in at least some centers"***

~~However, to~~We do wish to emphasise~~clarify~~ that~~,~~ we have always stated that the pooled statistic is no substitute for local level data and this was stated in the Discussion: ***"This highlights the importance of each institution auditing their own performance so that patients can be counselled reliably and to improve clinical practice."*** **However in recognition of the reviewer's repeated highlighting of this point, ~~To~~ we have further reiterate the~~is~~ point, ~~we have included~~ including~~this~~ as part of the Take Home Message for readers:** ***"Given the institutional variation in results, it is of upmost importance to base decision-making on local data if available."*** **We thank the reviewer for prompting what is, no doubt, an essential take home message for the review.**

Formatted: Font: Not Italic

2. Explain in the discussion how "heterogeneity should be considered" when implementing the results of the study.

We thank the reviewer for this comment and contend that this has been addressed in the second paragraph of the Discussion section in some detail ***"Furthermore, institutional and individual differences in technique, skill and experience also have an impact on the diagnostic performance of mpMRI. Sonn et al highlighted the difference in the interpretation of mpMRI with individual radiologists being a significant factor on multivariable analysis for cancer detection[55]. Furthermore, amongst different radiologists, the prevalence of clinically significant cancer varied from 13% to 60% in men with PIRADS 1-2.".*** **We have also acknowledged the heterogeneity in the last paragraph of our Discussion:** ***"First, there is substantial variation between each of the studies, especially regarding population, protocols, technique and experience of clinicians, which we were unable to explain through our subgroup analyses."***

4. Add something to the conclusions in the text and in the abstract about the research that should be done next.

We thanks the reviewer for this helpful comment and have added the following into the conclusion: ***"Further research needs to be conducted to identify the reasons behind the heterogeneity observed ~~so that they can be addressed and~~in order for consistent quality~~ is consistent~~ between centres."*** **Unfortunately, due to word constraints we have not been able to make such a comment in the abstract.**

With respect to publication bias, asymmetry of the funnel plot etc., as I said in my first review, the funnel plot and associated test is not valid under the conditions, observed here, that there is heterogeneity and the direction of results is likely correlated with the size of the study. If high volume centers do good MRIs and have a lot of patients and do big studies, and low volume centers do bad MRIs and do small studies, you would see the funnel plot in supplementary figure 1.

Commented [AL3]: Does he mean that we would see this in the funnel plot if it were true? Or is he saying that this is true and that's what we see? What do you understand by his meaning?

We thank the reviewer for their comment and have removed the statement about the trim and fill method and instead stated that publication bias is likely to be present.

Negative predictive value of multi-parametric MRI in detection of clinically significant prostate cancer in the PI-RADS era: a systematic review and meta-analysis

Niranjan J Sathianathen[1,2], Altan Omer[3], Eli Harriss[4], Lucy Davies[3], Veeru Kasivisvanathan[5], Shonit Punwani[5], Caroline M Moore[5], Christof Kastner[6], Tristan Barrett[6], Roderick Van Den Bergh[7], Ben A Eddy[8], Fergus Gleeson[3], Ruth Macpherson[3], Richard J Bryant[3], James WF Catto[9], Declan G Murphy[1], Freddie C Hamdy[3], Hashim U. Ahmed[10], Alastair D Lamb[3]

[1]University of Melbourne, Peter MacCallum Cancer Centre, Melbourne, Australia

[2] University of Minnesota, Dept. of Urology, Minneapolis, United States of America

[3]University of Oxford, Nuffield Dept. of Surgical Sciences, Oxford, United Kingdom

[4]Bodleian Health Care Libraries, University of Oxford, Oxford, United Kingdom

[5]University College London Hospital, Dept. of Urology, London, United Kingdom

[6]Cambridge University Hospitals NHS Foundation Trust, CamPARI Clinic, Cambridge, United Kingdom

[7]Utrecht Medical Centre, Dept. of Urology, Utrecht, The Netherlands

[8]Department of Urology, Canterbury Hospital, Canterbury, Kent, United Kingdom

[9]University of Sheffield, Sheffield, United Kingdom

[10]Division of Surgery, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom

Funding: none

Conflicts of interest: none relevant

Abstract word count: 300

Wound count: 3,205

Figure: 4

Tables: 1

Key words: Prostate cancer, MP-MRI, negative predictive value, diagnosis, biopsy, PSA

Context: Pre-biopsy multi-parametric MRI (mpMRI) is increasingly used in prostate cancer diagnosis. The reported negative predictive value (NPV) of mpMRI is used by some clinicians to aid decision making about whether or not to proceed to biopsy.

Objective: We aim to perform a contemporary systematic review that reflects the latest literature on optimal mpMRI techniques and scoring systems to update the NPV of mpMRI for clinically significant prostate cancer (csPCa).

Evidence acquisition: We conducted a systematic literature search and included studies from 2016-4th September 2019 which assessed the NPV of mpMRI for csPCa, using biopsy or clinical follow-up as the reference standard. To ensure that studies included in this analysis reflect contemporary practice, we only included studies in which mpMRIs were interpreted according to the PIRADS or similar Likert grading system. We define a negative mpMRI as either (1) PIRADS/Likert 1-2 or (2) PIRADS/Likert 1-3. csPCa was defined as either (1) Gleason grade group ≥2 or (2) Gleason grade group ≥3. We calculated NPV separately for each combination of negative mpMRI and csPCa.

Evidence synthesis: A total of 42 studies with 7,321 patients met our inclusion criteria and were included for analysis. Using definition (1) for a negative mpMRI and csPCa, the NPV for biopsy-naïve men pooled NPV was 90.8% [95%CI 88.1-93.1%]. When defining csPCa using definition (2), the NPV for csPCa 97.1% [95%CI 94.9-98.7%]. Calculation of the pooled NPV using definition (2) for negative mpMRI and definition (1) for csPCa: 86.8% [95%CI 80.1-92.4%]. Using definition (2) for both negative mpMRI and csPCa, the pooled NPV from two studies was 96.1% [95%CI 93.4-98.2%].

Conclusion: ~~Multi-parametric~~mp ~~-~~MRI of the prostate is generally an accurate test for ruling-out clinically significant prostate cancer. However, we observed heterogeneity in the NPV estimates ~~that must be considered in interpreting the applicability of these data~~and local institutional data should form the basis of decision-making if available.~~.~~

Patient summary: These NPV figures should assist decision making for clinicians considering not proceeding to biopsy in men with an elevated age-specific PSA and an mpMRI reported as negative (or equivocal) on PIRADS/Likert scoring. Some 7-10% of men, depending on setting, will miss a diagnosis of clinically significant cancer if they do not proceed to biopsy. Given the institutional variation in results, it is of upmost importance to base decision-making on local data if available.

Introduction

Multiparametric magnetic resonance imaging (mpMRI) is increasingly employed as a diagnostic tool following multiple studies demonstrating the accuracy in prostate cancer detection in men with an elevated prostate specific antigen (PSA) level or other clinical suspicion such as an abnormal digital rectal examination (DRE) [1]. Prior to the mpMRI era, serum PSA testing and DRE alone were used to inform decision-making about proceeding to prostate biopsy using transrectal ultrasound (TRUS), an imaging modality that is targeted to the prostate alone, rather than suspicious lesions, and does not provide adequate visualisation of intra-prostatic lesions. mpMRI brings another tool to this decision-making process improving both the performance of biopsy by permitting targeting of visualised lesion and raises the possibility of not proceeding to biopsy in the context of a non-suspicious mpMRI . The PROMIS study was pivotal in establishing the clinical validity of mpMRI as it demonstrated that use of mpMRI prior to biopsy resulted in 18% more cases of clinically significant prostate cancer (csPCa) being detected, compared to TRUS biopsy alone [2]. This study used a transperineal mapping biopsy (saturation sampling of the prostate at 5mm increments) as a reference point. Furthermore, it also showed that 27% of men could potentially avoid biopsy in cases where mpMRI did not display a suspicious lesion and thus prevent the diagnosis of low risk prostate cancer in 5%. This was a practice-changing study that provided justification for the use of mpMRI as a triage tool for biopsy, especially in the biopsy-naive setting. Subsequent studies, summarised in a recent Cochrane review on the subject, have demonstrated the clinical utility of mpMRI pre-biopsy to triage men to biopsy or as an adjunct to systematic biopsies using MRI-targeted sampling in addition[3]. Along these lines, a recent systematic review demonstrated that sampling of MRI suspicious areas improved the diagnosis of significant cancer and reduced the diagnosis of indolent disease[4]. Whereas the review by Kasivisvanathan et al measures the accuracy of a positive mpMRI, our paper is the complementary analysis and assesses the reliability of a negative mpMRI - the key question if men are to be discharged without biopsy.

For mpMRI to be used in the decision whether or not to biopsy, it is important that both clinicians and patients are aware of the potential risk of missing clinically significant disease, that is, the negative predictive value (NPV), as well as the potential for over-diagnosis. In addition, the burden of active surveillance for men and the healthcare system should not be underestimated. Although patients are often counselled using data from large, prominent trials and centres, such high levels of diagnostic performance may not reflect true practice in the wider urological community. There are a multitude of factors that impact the sensitivity and specificity of mpMRI and the marked variation between centres is highlighted by the disparity in diagnostic accuracy that has been reported in the past [3]. Hence, in the absence of the ideal scenario whereby data from each centre is available, an alternative is to counsel patients using pooled data from the entirety of the literature using a systematic approach. Such an approach by Moldovan and colleagues in 2016 identified 48 published studies that had reported the NPV of mpMRI but only eight could form the basis of a subsequent meta-analysis. Indeed, only one actually informed the NPV for clinically significant cancer [5]. Additionally, the majority of the included studies were from the early mpMRI era, published before PI-RADS v1; there have been considerable advances in image-acquisition protocols and reporting standardisation since which may impact on diagnostic performance. We therefore performed an updated systematic review and meta-analysis to include recently published studies in order to determine the NPV of mpMRI for csPCA and therefore whether it is safe to omit biopsy in men with a negative mpMRI in the primary diagnostic setting.

Evidence acquisition

The review was conducted in accordance to PRISMA guidelines (Supplementary Table 1). A protocol was registered in PROSPERO [CRD42018111619] on 18th October 2018. We searched six databases for relevant studies published between 2016 to the search date (4th September 2019): Ovid Medline; Ovid EMBASE; The Cochrane Library; Scopus; the Web of Science Core Collection; and the Centre for Reviews and Dissemination HTA database. The search strategies used text words, phrases, and relevant index terms to combine the concepts for “prostate cancer” or “prostate neoplasms”, “magnetic resonance imaging”, biopsy or pre-biopsy, and a search filter for diagnostic studies, using

an adapted version of the search strategy used by Moldovan et al[5]. We did not make any restrictions on study design nor the language of publication. The full search strategies are shown in Supplementary Table 2.

We included patients with no previous diagnosis of prostate cancer, i.e. primary biopsy or previously negative biopsy. We had initially specified that all indications for prostate biopsy would be included but recognised the variability and likely bias by including patients on active surveillance which would compromise the applicability of our results to patients not diagnosed with prostate cancer and decided to exclude this population. We will investigate this population in a separate study.

Studies evaluating the prevalence of clinically significant prostate cancers defined as the presence of any Gleason Grade Group 2 (Gleason Score 3+4=7) cancer or above in negative mpMRI scans were included. We defined an mpMRI as being negative when PIRADS or equivalent scoring system were 1 or 2. We included studies that verified mpMRI results using any form of biopsy and/or clinical follow-up. Biopsy approaches could include systematic biopsy methods, including transrectal (TR) or transperineal (TP) or mapping biopsy, or clinical follow-up with subsequent diagnosis using any of the above. When clinical follow-up was used to verify negative mpMRI results, we defined true negatives as participants with a negative mpMRI that did not present with clinically significant prostate cancer during the period of follow up of the study, accepting that this may vary. We included all mpMRI phases and techniques used to detect cancer (e.g. T2W, DWI, DCE or spectroscopy) and all mpMRI systems.

The results of the search were screened initially by title and abstract for relevance. Articles that were determined to be of interest were reviewed based on full text to determine relevance and whether they satisfied the inclusion/exclusion criteria. If more than one report of the same trial was found, only the most up-to-date publication was included in the analysis. This process was performed by two independent authors (NS and AO) with a third author (AL) consulted to resolve any disagreements.

We collected data on cohort characteristics, mpMRI protocol, biopsy protocol and outcomes of interest. The primary outcome of interest was the presence of clinically significant prostate cancer defined as the presence of any Gleason 3+4 (Gleason Grade Group 2) cancer or above in negative MP-MRI scans defined as PIRADS 1 or 2.

We intended to perform subgroup analyses based on the following factors if the data was available:

- Strength of MRI magnet: 3 Tesla versus 1.5 Tesla
- Clinical setting: university hospitals versus district or non-academic institutions

We performed sensitivity analysis on the primary outcome based on the definition of a negative mpMRI where the PIRADS score threshold was changed to incorporate a score 3 as negative or non-suspicious. We also performed sensitivity analysis on the definition of 'clinically significant prostate cancer' by evaluating any Gleason 4+3 or higher as a histological threshold for significance. Finally, we performed post-hoc testing to characterise the prevalence of Gleason Grade Group 1 (Gleason score 3+3=6) cancer in men with a negative mpMRI according to our primary definition.

Risk of bias

Studies were assessed for quality using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool[6] by two independent authors, with consultation with a senior author to resolve any disagreements. Publication bias was assessed visually by examining funnel plots for asymmetry and statistically using Egger's test.

Statistical analysis

Proportions were extracted from the publications and transformed with the Freeman-Tukey double inverse sine transformation; the Wilson Score method was used to calculate confidence intervals. We pooled the effect estimates using a DerSimonian and Laird inverse of variance random effects model and the results are presented in forest plots. Heterogeneity was determined using the I^2 statistic which was interpreted in the following way: 0-40% low, 30-60% moderate, 50-90% substantial and

75-100% considerable heterogeneity based on GRADE guidance. We conducted subgroup analyses and tested for interaction using the Borenstein method. We also conducted a sensitivity analysis based on risk of bias (excluding studies with high or unclear risk). We used funnel plots to examine the extent of publication bias. This review was reported in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (Supplementary Table 1). Extracted data was collated in Excel (Microsoft Corporation, Redmond, CA, USA) and analysis performed using R Statistical Software (Foundation for Statistical Computing, Vienna, Austria).

Evidence synthesis

Our initial search returned 3,763 records of which 487 were included for full-text review. Of these, 42 met the eligibility criteria and were included for analysis (Figure 1). We include the key characteristics for each trial (Figure 1, Table 1[2, 7-52]), and the risk of bias assessment for each (Supplementary Table 3). PIRADS was used to interpret mpMRIs in 74% (n=31). Histopathological verification was performed using transperineal mapping biopsy in 12% (n=5), the majority utilising TRUS 12 core biopsy (62%, n=26).

Primary definition: MRI negative = PIRADS 1-2, csPCa = Gleason Grade Group >2 (Gleason score ≥3+4=7)

A total of 42 studies, including 7,321 men, reported sufficient data which could be pooled and analysed for the NPV of mpMRI using our primary definition of negative MRI (PIRADS 1-2) and csPCa (Gleason score ≥3+4=7, GGG 2). The summary NPV for all patients with no previous diagnosis of prostate cancer was 91.3% [95%CI 89.3-93.2%, I^2=86%] (Figure 2).

Of the above studies, 36 published findings for patients who were undergoing primary biopsy and for whom the pooled NPV of mpMRI was 90.8% [95%CI 88.1-93.1%, I^2=88%]. For the men in the 16 studies who had undergone a previous negative prostate biopsy, the pooled NPV was 92.7% [95%CI

89.8-95.2%, I^2=73%]. Statistically, there was no robust evidence of a difference in NPV between these populations (p=0.4). There was significant asymmetry observed in the funnel plot (p=0.01) (Supplementary Figure 1). ~~This was primarily due to studies missing on the right side of the funnel, i.e. high NPV.~~

Sensitivity analysis

We performed a sensitivity analysis only including the ten studies that were considered to be 'low' risk of bias (Supplementary Table 3). The pooled NPV based on the primary definition used above for all men undergoing primary biopsy or those with a previous negative biopsy was 93.4% [95%CI 90.8-95.7%, I^2=73%] (Supplementary Figure 2). Of these low risk of bias studies, nine provided data for men undergoing primary biopsy only and the summary NPV for this group was 92.7% [95%CI 88.7-95.7%, I^2=80%]. Whereas, five studies reported on men with a previous negative biopsy for whom the summary NPV was 94.5% [95%CI 90.1-97.8%, I^2=66%). There was no difference between the NPV of these groups (p=0.5).

Subgroup analysis

Magnet strength: 1.5T vs 3T

Of the studies included in the primary analysis, 22 studies of biopsy naïve men only used a single magnet strength. Interestingly, there was no difference (p=0.45) in the NPV between the studies using 1.5T [90.2%, 95%CI 81.3-96.5%, studies=6, I^2=93%] or 3T [93.2%, 95%CI 90.4-95.5%, studies=16, I^2=77%]. There was no difference in the previous negative biopsy subpopulation, (86.0% [95%CI 67.9-97.7%, studies=2, I^2=76%] vs 93.9% [95%CI 87.0-98.6%, studies=6, I^2=72%], p=0.4).

Clinical setting: university hospitals versus district or non-academic institutions

There was insufficient data from non-academic institutions to perform this subgroup analysis.

Secondary definitions

MRI negative = PIRADS 1-3, csPCa = Gleason score ≥3+4=7 (GGG 2)

We found eleven studies that were able to be pooled for a NPV of 87.0% [95%CI 81.8-91.5%, I^2=93%] for when a 'negative' MRI includes PIRADS 1-3 and csPCa as any Gleason score ≥3+4=7 (Figure 3A). There was no difference in NPV between the biopsy naïve [86.8%, 95%CI 80.1-92.4%, studies=9, I^2=95%] and previous negative biopsy [87.4%, 95%CI 78.9-94.1%, studies=4, I^2=83%] subpopulations (p=0.9).

MRI negative = PIRADS 1-2, csPCa = Gleason score ≥4+3=7 (GGG 3)

Seventeen studies reported sufficient data to calculate NPV based on the above secondary definition of negative MRI and csPCa which was 97.4% [95%CI 95.8-98.7%, I^2=75%] (Figure 3B). Of these studies, 13 included patients who were undergoing their first biopsy and had a summary NPV in the subgroup of 97.1% [95%CI 94.9-98.7%, I^2=79%]. Eight studies included men with previous negative biopsies and had a pooled NPV of 97.9% [95%CI 95.2-99.7%, I^2=66%]. There was no difference in NPV between these subgroups (p=0.6).

MRI negative = PIRADS 1-3, csPCa = Gleason score ≥4+3=7 (GGG 3)

We calculated a pooled NPV of 94.7% [95%CI 92.0-96.9%, I^2=78%] from seven studies for the above secondary definition (Figure 3C). There was no difference (p=0.24) in NPV between the biopsy naïve [96.1%, 95%CI 93.4-98.2%, studies=4, I^2=70%] and previous negative biopsy [93.3%, 95%CI 87.9-97.2%, studies=5, I^2=82%].

Prevalence of Gleason score 3+3=6 (GGG 1) in negative mpMRI (PIRADS 1-2)

Thirty two studies reported sufficient data to estimate the prevalence of Gleason score 3+3=6 (GGG 1) in men with an mpMRI score of 1 or 2 (Figure 2, Supplementary Figure 3). Including both men undergoing primary biopsy and those who had a history of a negative prostate biopsy, we estimated the prevalence of GGG 1 prostate cancer in negative mpMRI as 17% [95%CI 14-20%]. There was

significant heterogeneity in the pooled studies [I^2=91%]. There was no difference in the prevalence of GGG 1 in negative mpMRI between men undergoing primary biopsy and those who had a previous negative biopsy (p=0.5).

Discussion

In summary, based on our primary definition of negative mpMRI being PIRADS 1-2 and csPCa being Gleason score ≥3+4=7, men who have a negative or non-suspicious mpMRI have an approximate 1 in 10 probability of harbouring csPCa. This is consistent with a recent Cochrane review which estimated from twelve studies that the NPV of MRI for ≥3+4=7 ranges between 86-97%[3]. Our systematic review was unable to determine summary statistics for the exact burden in terms of length of cancer or amount of pattern 4 involvement. For cspCa defined as any Gleason 4+3 or higher, men have a 1 in 20 chance of harbouring such disease despite a negative mpMRI. The findings from our study represent an average NPV taking into account the multiple factors impacting this, and these results are therefore likely to broadly represent clinical practice across many similar institutions. These estimates can be used by clinicians who do not have access to their own data to counsel patients rather than relying on performance data from single institutional studies from centres of excellence, which may be unattainable in the district or non-academic setting. When contexualised with the findings from previous systematic reviews assessing the PPV[4], it ~~strongly~~ suggests that mpMRI should be routinely incorporated in the diagnostic pathway of prostate cancer as it can both reduce the number of biopsies in men with clinically insignificant disease and also improves diagnosis in those with significant cancer. This concept has been supported by the 2019 European Association of Urology and NICE guidelines[53, 54].

It should however be noted that there was considerable heterogeneity between the NPV between studies which is likely a reflection of difference in population characteristics such as cancer prevalence which inherently influences the NPV. Examining the range of point estimates of the individual studies, the lowest NPV for our primary definition was 63% based on a small study of 24

patients but this was one of two outliers. ~~This study was one of two outliers as the majority had an NPV above 85%.~~ Furthermore, institutional and individual differences in technique, skill and experience also have an impact on the diagnostic performance of mpMRI. Sonn et al highlighted the difference in the interpretation of mpMRI with individual radiologists being a significant factor on multivariable analysis for cancer detection[55]. Furthermore, amongst different radiologists, the prevalence of clinically significant cancer varied from 13% to 60% in men with PIRADS 1-2. On the contrary, a separate institution demonstrated equivalence amongst both urologists and radiologists in the detection of significant disease[56]. This highlights the importance of each institution auditing their own performance so that patients can be counselled reliably and to improve clinical practice.

Although demonstrating that mpMRI has a high NPV, the risk of missing significant cancer by omitting biopsy is not insignificant and each individual patient and their physician will need to make an informed choice about whether such a risk for the definition of csPCa we have used warrants a systematic biopsy, and if so, what type of biopsy. Nonetheless, there is a need for research to evaluate risk stratification tools that might further improve the NPV. One such variable that is emerging as an important stratification tool is PSA density[14, 57-59]. Venderink et al showed in a study of 1057 men, that for PI-RADS 3 lesions (156/1057), where the PSA density is <0.15 ng/ml/cc, the false negative rate was only 6% (NPV 94%) for csPCa determined as Gleason Grade Group ≥2[60]. Bryant and colleagues suggested that a PSA density above 0.15 ng/ml/cc only acts a predictor of an increased likelihood of underlying csPCa in men with negative mpMRI when csPCa is defined as Gleason Grade Group 3-5 disease and/or tumour length ≥6 mm[28, 57, 58]. A prior negative biopsy may also a reassuring feature although we did not find a statistically significant difference in the summary NPVs between the primary biopsy (89%) and prior negative biopsy groups (93%) [61]. The concept of omitting biopsy in patients with a negative mpMRI while taking into account clinical risk has been supported by the PIRADS Steering committee and the ESUR Prostate MRI Working Group[62, 63]. The incorporation of mpMRI findings into current risk calculators, such as the Rotterdam European Randomized Study of Screening for Prostate Cancer risk calculators, has been

shown to improve accuracy of risk stratification (AUC 0.84 vs 0.76), and facilitates its use in daily practice[64].

This meta-analysis does have some limitations. First, there is substantial variation between each of the studies, especially regarding population, protocols, technique and experience of clinicians, which we were unable to explain through our subgroup analyses. Second, it is important to note that the reference test used to determine 'true negative' rates was, in many studies, 12-core transrectal ultrasound-guided biopsy which has been shown to miss significant disease and therefore our results may overestimate NPV. Nevertheless, given that this method of primary sampling is still employed by the vast majority of urology units world-wide, these results represent an important a clinically meaningful endpoint for patients.

Conclusion

Multi-parametric MRI of the prostate ~~is~~ can be an accurate test for ruling-out clinically significant prostate cancer when implemented in at least some centers. Our systematic review demonstrated that approximately 1 in 10 men who have a non-suspicious MRI can still harbour clinically significant prostate cancer, although the actual probability depends on the clinical setting, threshold used for defining a non-suspicious mpMRI, definition of significant prostate cancer and the diagnostic performance of each local institution. For our primary definition, the NPV ranged from 63-100% for individual studies~~, with the majority exceeding 85% and being highly accurate in ruling-out significant disease~~.~~-~~ Improved risk stratification using clinical factors, and possibly in future other biomarkers, may provide further improvements to reassure men and physicians contemplating avoiding a biopsy. Further research needs to be conducted to identify the reasons behind the heterogeneity observed so that they can be addressed and quality is consistent between centres.

References

[1] Kim SP, Karnes RJ, Mwangi R, Van Houten H, Gross CP, Gershman B, et al. Contemporary Trends in Magnetic Resonance Imaging at the Time of Prostate Biopsy: Results from a Large Private Insurance Database. European urology focus. 2019.
[2] Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. The Lancet. 2017;389:815-22.
[3] Drost FJH, Osses DF, Nieboer D, Steyerberg EW, Bangma CH, Roobol MJ, et al. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. Cochrane Database of Systematic Reviews. 2019.
[4] Kasivisvanathan V, Stabile A, Neves JB, Giganti F, Valerio M, Shanmugabavan Y, et al. Magnetic Resonance Imaging-targeted Biopsy Versus Systematic Biopsy in the Detection of Prostate Cancer: A Systematic Review and Meta-analysis. European urology. 2019;76:284-303.
[5] Moldovan PC, Van den Broeck T, Sylvester R, Marconi L, Bellmunt J, van den Bergh RCN, et al. What Is the Negative Predictive Value of Multiparametric Magnetic Resonance Imaging in Excluding Prostate Cancer at Biopsy? A Systematic Review and Meta-analysis from the European Association of Urology Prostate Cancer Guidelines Panel. European urology. 2017;72:250-66.
[6] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011;155:529-36.
[7] Filson CP, Natarajan S, Margolis DJ, Huang J, Lieu P, Dorey FJ, et al. Prostate cancer detection with magnetic resonance-ultrasound fusion biopsy: The role of systematic and targeted biopsies. Cancer. 2016;122:884-92.
[8] Hansen N, Patruno G, Wadhwa K, Gaziev G, Miano R, Barrett T, et al. Magnetic Resonance and Ultrasound Image Fusion Supported Transperineal Prostate Biopsy Using the Ginsburg Protocol: Technique, Learning Points, and Biopsy Results. European urology. 2016;70:332-40.
[9] Kam J, Yuminaga Y, Khoury S, Macneil F, Ouyang R, Ruthven S, et al. Can multi-parametric MRI (mpMRI) decrease the number of unnecessary biopsies performed for prostate cancer detection? Early experience in a regional centre2016.
[10] Mahon J, P Kaufman Jr R, O'Malley R, Mian B, Fisher H, Essa A. MP53-02 NEGATIVE PREDICTIVE VALUE OF A NEGATIVE MULTI-PARAMETRIC MRI OF PROSTATE2016.
[11] Renard-Penna R, Roupret M, Comperat E, Rozet F, Granger B, Barkatz J, et al. Relationship between non-suspicious MRI and insignificant prostate cancer: results from a monocentric study. World journal of urology. 2016;34:673-8.
[12] Schouten MG, van der Leest M, Pokorny M, Hoogenboom M, Barentsz JO, Thompson LC, et al. Why and Where do We Miss Significant Prostate Cancer with Multi-parametric Magnetic Resonance Imaging followed by Magnetic Resonance-guided and Transrectal Ultrasound-guided Biopsy in Biopsy-naive Men? European urology. 2017;71:896-903.
[13] Thompson JE, van Leeuwen PJ, Moses D, Shnier R, Brenner P, Delprado W, et al. The Diagnostic Performance of Multiparametric Magnetic Resonance Imaging to Detect Significant Prostate Cancer. The Journal of urology. 2016;195:1428-35.
[14] Washino S, Okochi T, Saito K, Konishi T, Hirai M, Kobayashi Y, et al. Combination of prostate imaging reporting and data system (PI-RADS) score and prostate-specific antigen (PSA) density predicts biopsy outcome in prostate biopsy naive patients. BJU international. 2017;119:225-33.
[15] Adamcová V, Záleský M, Stejskal J, Minárik I, Koldová M, Pavličko A, et al. MRI as a triage test in prostate cancer diagnostic algorithm: Prospective study. European Urology Supplements. 2017;16:e2876-e7.
[16] Boesen L, Norgaard N, Logager V, Thomsen HS. Clinical Outcome Following Low Suspicion Multiparametric Prostate Magnetic Resonance Imaging or Benign Magnetic Resonance Imaging Guided Biopsy to Detect Prostate Cancer. The Journal of urology. 2017;198:310-5.
[17] Gaunay G, Patel V, Shah P, Moreira D, Hall SJ, Vira MA, et al. Role of multi-parametric MRI of the prostate for screening and staging: Experience with over 1500 cases. Asian Journal of Urology. 2017;4:68-74.

[18] Hoffmann MA, Taymoorian K, Ruf C, Gerhards A, Leyendecker K, Stein T, et al. Diagnostic Performance of Multiparametric Magnetic Resonance Imaging and Fusion Targeted Biopsy to Detect Significant Prostate Cancer. Anticancer research. 2017;37:6871-7.
[19] Hansen NL, Kesch C, Barrett T, Koo B, Radtke JP, Bonekamp D, et al. Multicentre evaluation of targeted and systematic biopsies using magnetic resonance and ultrasound image-fusion guided transperineal prostate biopsy in patients with a previous negative biopsy. BJU international. 2017;120:631-8.
[20] Jambor I, Bostrom PJ, Taimen P, Syvanen K, Kahkonen E, Kallajoki M, et al. Novel biparametric MRI and targeted biopsy improves risk stratification in men with a clinical suspicion of prostate cancer (IMPROD Trial). Journal of magnetic resonance imaging : JMRI. 2017;46:1089-95.
[21] Karman A, Shymanets S, Kharuzhyk S, Minich A, Rolevich AI, Krasny S. Diagnostic accuracy of transrectal ultrasound and multiparametric MRI after primary negative prostate biopsy2017.
[22] Lu AJ, Syed JS, Nguyen KA, Nawaf CB, Rosoff J, Spektor M, et al. Negative Multiparametric Magnetic Resonance Imaging of the Prostate Predicts Absence of Clinically Significant Prostate Cancer on 12-Core Template Prostate Biopsy. Urology. 2017;105:118-22.
[23] Porpiglia F, Manfredi M, Mele F, Cossu M, Bollito E, Veltri A, et al. Diagnostic Pathway with Multiparametric Magnetic Resonance Imaging Versus Standard Pathway: Results from a Randomized Prospective Study in Biopsy-naive Patients with Suspected Prostate Cancer. European urology. 2017;72:282-8.
[24] Rhudd A, Moghul M, McDonald J. The diagnostic outcomes of men with a negative multi-parametric magnetic resonance imaging (mpMRI) for the diagnosis of prostate cancer, in low-risk biopsy naive men. J Endourol. 2018;31:A229-A30.
[25] Simmons LAM, Kanthabalan A, Arya M, Briggs T, Barratt D, Charman SC, et al. The PICTURE study: diagnostic accuracy of multiparametric MRI in men requiring a repeat prostate biopsy. British journal of cancer. 2017;116:1159-65.
[26] Wang RS, Kim EH, Vetter JM, Fowler KJ, Shetty AS, Mintz AJ, et al. Determination of the Role of Negative Magnetic Resonance Imaging of the Prostate in Clinical Practice: Is Biopsy Still Necessary? Urology. 2017;102:190-7.
[27] Boesen LP, Nørgaard N, Løgager V, Balslev I, Thestrup KC, Winther M, et al. Simple biparametric MRI in detection and ruling out significant prostate cancer in biopsy-naïve men (BIDOC study). European Urology Supplements. 2018;17:e891.
[28] Bryant RJ, Hobbs CP, Eyre KS, Davies LC, Sullivan ME, Shields W, et al. Comparison of Prostate Biopsy with or without Prebiopsy Multiparametric Magnetic Resonance Imaging for Prostate Cancer Detection: An Observational Cohort Study. The Journal of urology. 2019;201:510-9.
[29] Hansen NL, Barrett T, Kesch C, Pepdjonovic L, Bonekamp D, O'Sullivan R, et al. Multicentre evaluation of magnetic resonance imaging supported transperineal prostate biopsy in biopsy-naive men with suspicion of prostate cancer. BJU international. 2018;122:40-9.
[30] Hwang SI, Lee HJ, Lee SE, Hong SK, Byun SS, Lee SC, et al. Value of MR-US fusion in guidance of repeated prostate biopsy in men with PSA<10ng/mL. Clinical imaging. 2019;53:1-5.
[31] Karman AV, Shimanets SV, Minich AA, Rolevich AI, Karman TA, Krasny SA. Repeated 24-core systematic and targeted biopsy in detection of clinically significant prostate cancer. European Urology Supplements. 2018;17:e2189.
[32] Kim JG, Kim CK, Kim JH. Prebiopsy Multiparametric MRI With Cancer-Negative Findings in Men With Suspected Prostate Cancer: Evaluation Using Prostate Imaging Reporting and Data System Version 2. AJR American journal of roentgenology. 2018;211:121-6.
[33] Lobo N, Petrides N, Stanowski M, Morrison I, Thomas M, Kommu S, et al. Can we rely on a negative multiparametric MRI to exclude significant prostate cancer at biopsy? Results from a regional cancer centre. European Urology Supplements. 2019;18:e1876.
[34] Mannaerts CK, Gayet M, Verbeek JF, Engelbrecht MRW, Savci-Heijink CD, Jager GJ, et al. Prostate Cancer Risk Assessment in Biopsy-naive Patients: The Rotterdam Prostate Cancer Risk Calculator in Multiparametric Magnetic Resonance Imaging-Transrectal Ultrasound (TRUS) Fusion Biopsy and Systematic TRUS Biopsy. Eur Urol Oncol. 2018;1:109-17.
[35] Meng Y, Vetter JM, Parker AA, Arett CT, Andriole GL, Shetty AS, et al. Improved Detection of Clinically Significant Prostate Cancer With Software-assisted Systematic Biopsy Using MR/US Fusion in Patients With Negative Prostate MRI. Urology. 2018;120:162-6.

[36] Morote J, Celma A, Roche S, de Torres IM, Mast R, Semedey ME, et al. Who Benefits from Multiparametric Magnetic Resonance Imaging After Suspicion of Prostate Cancer? European Urology Oncology.
[37] Mortezavi A, Marzendorfer O, Donati OF, Rizzi G, Rupp NJ, Wettstein MS, et al. Diagnostic Accuracy of Multiparametric Magnetic Resonance Imaging and Fusion Guided Targeted Biopsy Evaluated by Transperineal Template Saturation Prostate Biopsy for the Detection and Characterization of Prostate Cancer. The Journal of urology. 2018;200:309-18.
[38] Oishi M, Ashrafi A, Cacciamani GE, Shin T, Ohe C, Ghodoussipour S, et al. Which patients should consider and which patients could safely avoid prostate biopsy in the setting of negative mpMRI? European Urology Supplements. 2018;17:e904.
[39] Otti VC, Miller C, Powell RJ, Thomas RM, McGrath JS. The diagnostic accuracy of multiparametric magnetic resonance imaging before biopsy in the detection of prostate cancer. BJU international. 2019;123:82-90.
[40] Pal RP, Ahmad R, Trecartan S, Voss J, Ahmed S, Bazo A, et al. A Single Center Evaluation of the Diagnostic Accuracy of Multiparametric Magnetic Resonance Imaging against Transperineal Prostate Mapping Biopsy: An Analysis of Men with Benign Histology and Insignificant Cancer following Transrectal Ultrasound Biopsy. The Journal of urology. 2018;200:302-8.
[41] Panebianco V, Barchetti G, Simone G, Del Monte M, Ciardi A, Grompone MD, et al. Negative multiparametric magnetic resonance imaging for prostate cancer: what's next? European urology. 2018;74:48-54.
[42] Simmons LAM, Kanthabalan A, Arya M, Briggs T, Barratt D, Charman SC, et al. Accuracy of Transperineal Targeted Prostate Biopsies, Visual Estimation and Image Fusion in Men Needing Repeat Biopsy in the PICTURE Trial. The Journal of urology. 2018;200:1227-34.
[43] Zhou Y, Zhou Z, Li Q, Xu Y, Sun H, Xiao Y, et al. Diagnostic accuracy of magnetic resonance-guided prostate biopsy and template-guided transperineal saturation biopsy. Medicine. 2018;97:e12495.
[44] Barrett T, Slough R, Sushentsev N, Shaida N, Koo BC, Caglic I, et al. Three-year experience of a dedicated prostate mpMRI pre-biopsy programme and effect on timed cancer diagnostic pathways. Clin Radiol. 2019.
[45] Elkhoury FF, Felker ER, Kwan L, Sisk AE, Delfin M, Natarajan S, et al. Comparison of Targeted vs Systematic Prostate Biopsy in Men Who Are Biopsy Naive: The Prospective Assessment of Image Registration in the Diagnosis of Prostate Cancer (PAIREDCAP) Study. JAMA Surg. 2019.
[46] Kruger-Stokke B, Selnaes KM, Langorgen S, Sjobakk TAE, Bathen TF, Bertilsson H. Multiparametric prostate MRI in biopsy naive men; then what? SCANDINAVIAN JOURNAL OF UROLOGY: TAYLOR & FRANCIS LTD 2-4 PARK SQUARE, MILTON PARK, ABINGDON OR14 4RN, OXON …; 2019. p. 9-11.
[47] Regis L, Celma A, Planas J, Lopez R, Roche S, Lorente D, et al. The role of negative magnetic resonance imaging: can we safely avoid biopsy in P.I.-R.A.D.S. 2 as in P.I.-R.A.D.S. 1? Scand J Urol. 2019;53:21-5.
[48] Rouviere O, Puech P, Renard-Penna R, Claudon M, Roy C, Mege-Lechevallier F, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. Lancet Oncol. 2019;20:100-9.
[49] Rozas GdQ, Saad LS, Melo HJdFe, Gabrielle HAA, Szejnfeld J. Impact of PI-RADS v2 on indication of prostate biopsy. International braz j urol. 2019;45:486-94.
[50] van der Leest M, Cornel E, Israel B, Hendriks R, Padhani AR, Hoogenboom M, et al. Head-to-head Comparison of Transrectal Ultrasound-guided Prostate Biopsy Versus Multiparametric Prostate Resonance Imaging with Subsequent Magnetic Resonance-guided Biopsy in Biopsy-naive Men with Elevated Prostate-specific Antigen: A Large Prospective Multicenter Clinical Study. European urology. 2019;75:570-8.
[51] Zalesky M, Stejskal J, Adamcova V, Hrbacek J, Minarik I, Pavlicko A, et al. Use of Prostate Specific Antigen Density Combined with Multiparametric Magnetic Resonance Imaging Improves Triage for Prostate Biopsy. Urol Int. 2019;103:33-40.
[52] Zhang Y, Zeng N, Zhang F, Huang Y, Tian Y. How to make clinical decisions to avoid unnecessary prostate screening in biopsy-naive men with PI-RADs v2 score </= 3? Int J Clin Oncol. 2019.

[53] NICE Guidance – Prostate cancer: diagnosis and management. BJU international. 2019;124:9-26.
[54] Mottet N, Bellmunt J, Briers E, Bolla M, Cornford P, De Santis M, et al. EAU–ESTRO–SIOG Guidelines on prostate cancer. Europ Association Urol. 2018.
[55] Sonn GA, Fan RE, Ghanouni P, Wang NN, Brooks JD, Loening AM, et al. Prostate Magnetic Resonance Imaging Interpretation Varies Substantially Across Radiologists. European urology focus. 2019;5:592-9.
[56] Sathianathen NJ, Warlick CA, Soubra A, Balaji P, Metzger GJ, Spilseth B, et al. Difference in MRI-guided biopsy cancer detection rates between individual clinicians. Urol Oncol. 2019;37:299.e1-.e6.
[57] Rais-Bahrami S, Siddiqui MM, Vourganti S, Turkbey B, Rastinehad AR, Stamatakis L, et al. Diagnostic value of biparametric magnetic resonance imaging (MRI) as an adjunct to prostate-specific antigen (PSA)-based detection of prostate cancer in men without prior biopsies. BJU international. 2015;115:381-8.
[58] Hansen NL, Barrett T, Koo B, Doble A, Gnanapragasam V, Warren A, et al. The influence of prostate-specific antigen density on positive and negative predictive values of multiparametric magnetic resonance imaging to detect Gleason score 7-10 prostate cancer in a repeat biopsy setting. BJU international. 2017;119:724-30.
[59] Sim KC, Sung DJ, Kang KW, Yang KS, Han NY, Park BJ, et al. Magnetic Resonance Imaging-Based Prostate-Specific Antigen Density for Prediction of Gleason Score Upgrade in Patients With Low-Risk Prostate Cancer on Initial Biopsy. J Comput Assist Tomogr. 2017;41:731-6.
[60] Venderink W, van Luijtelaar A, Bomers JG, van der Leest M, Hulsbergen-van de Kaa C, Barentsz JO, et al. Results of Targeted Biopsy in Men with Magnetic Resonance Imaging Lesions Classified Equivocal, Likely or Highly Likely to Be Clinically Significant Prostate Cancer. European urology. 2017.
[61] Oishi M, Shin T, Ohe C, Nassiri N, Palmer SL, Aron M, et al. Which Patients with Negative Magnetic Resonance Imaging Can Safely Avoid Biopsy for Prostate Cancer? The Journal of urology. 2019;201:268-77.
[62] Padhani AR, Barentsz J, Villeirs G, Rosenkrantz AB, Margolis DJ, Turkbey B, et al. PI-RADS Steering Committee: The PI-RADS Multiparametric MRI and MRI-directed Biopsy Pathway. Radiology.0:182946.
[63] Richenberg J, Logager V, Panebianco V, Rouviere O, Villeirs G, Schoots IG. The primacy of multiparametric MRI in men with suspected prostate cancer. Eur Radiol. 2019.
[64] Alberts AR, Roobol MJ, Verbeek JFM, Schoots IG, Chiu PK, Osses DF, et al. Prediction of High-grade Prostate Cancer Following Multiparametric Magnetic Resonance Imaging: Improving the Rotterdam European Randomized Study of Screening for Prostate Cancer Risk Calculators. European urology. 2019;75:310-8.

Table 1. Characteristics of included studies

Study	Year	Setting	Magnet Brand	Magnet Strength	bp/mpMRI	Academic vs Community	MR Scoring System	Type biopsy	Age	PSA	Total Patients	Total Negative MRI (1-2) (%)
Filson[7]	2016	Biopsy naïve	Siemens	3	mpMRI	Academic	PIRADS	TRUS 12 core	Median 64 (IQR 59-69)	Median 5.8 (IQR 4.4-8.1)	328	56 (17)
		Previous negative							Median 65.7 (IQR 59.3-70.2)	Median 7.6 (IQR 5.0-11.5)	324	59 (18)
Hansen[8]	2016	Biopsy naïve	GE	Both	mpMRI	Academic	Likert	TP 24 core	Median 64 (IQR 58-70)	Median 6.2 IQR (4.8-8.6)	107	22 (21)
		Previous negative							Median 65 IQR 69-69	Median 7.8 IQR 6.0-12	295	91 (31)
Kam[9]	2016	Biopsy naïve	NR	NR	mpMRI	NR	PIRADS	TRUS 12 core	Mean 63	Mean 9.17	187	94 (50)
Mahon[10]	2016	Biopsy naïve	NR	NR	mpMRI	NR	PIRADS	TRUS 12 core	NR	NR	395	124 (31)
		Previous negative							NR	NR	395	45 (11)
Renard-Penna[11]	2016	Biopsy naïve	NR	1.5	mpMRI	Academic	Likert	TRUS 12 core	NR	NR	NR	78 (NA)
Schouten[12]	2016	Biopsy naïve	Siemens	3	mpMRI	NR	PIRADS	TRUS 12 core	NR	NR	223	81 (36)
Thompson[13]	2016	Biopsy naïve	NR	Both	mpMRI	Academic	PIRADS	TPM	NR	NR	344	79 (23)
Washino[14]	2016	Biopsy naïve	NR	Both	mpMRI	NR	PIRADS	TRUS 14 core	NR	NR	288	131 (46)
Adamcova[15]	2017	Biopsy naïve	NR	NR	mpMRI	NR	PIRADS	TRUS 12 core	Mean 62	Mean 6.33	206	47 (23)
		Previous negative							Mean 63.9	Mean 10.8	372	27 (7)
Ahmed[2]	2017	Biopsy naïve	NR	1.5	mpMRI	Multi	Likert	TPM	Mean 63 (SD 8)	Mean 7.1 (SD 2.9	576	158 (27)
Boesen[16]	2017	Previous negative		3	mpMRI	Academic	PIRADS	TRUS 10 core or follow-up	NR	NR	289	194 (67)
Gaunay[17]	2017	Biopsy naïve	NR	3	mpMRI	NR	PIRADS	TRUS 12 core	NR	NR	231	45 (20)
		Previous negative							NR	NR	282	68 (24)
Hansen[19]	2017	Previous negative	Siemens/ Ge	Both	mpMRI	Multi	PIRADS	TTPB 24 core	Median 66 (IQR 60-71)	Median 9 (IQR 7-13)	487	144 (30)
Hoffman[18]	2017	Previous negative	GE	3	mpMRI	Academic	PIRADS	TTPB 24 core	Mean 66.5 (range 49-80)	NR	99	6 (6)
Jambor[20]	2017	Biopsy naïve	NR	3	mpMRI	NR	Likert	TRUS 12 core	NR	NR	166	38 (6)
Karman[21]	2017	Previous negative	NR	1.5	mpMRI	Academic	PIRADS	TRUS 24 core	Median 66	Median 7.1	252	161 (63)
Lu[22]	2017	Biopsy naïve	Siemens	3	mpMRI	Academic	Mixed	TRUS 12 core	NR	NR	670	38 (6)

		Previous negative							NR	NR	670	33 (5)
Porpiglia[23]	2017	Biopsy naïve	NR	1.5	mpMRI	Academic	PIRADS	TRUS 12 core	Median 64 (IQR 58-70)	Median 5.9 (IQR 4.8-7.5)	107	26 (24)
Rhudd[24]	2017	Biopsy naïve	NR	NR	mpMRI	NR	PIRADS	NR	NR	NR	181	24 (13)
Simmons[25]	2017	Previous negative	NR	3	mpMRI	Academic	Likert	TTPM	Mean 63 (SD 7)	NR	249	35 (14)
Wang[26]	2017	Biopsy naïve	Siemens	3	mpMRI	Academic	PIRADS	TRUS 12 core	Mean 61.9 (SD 7)	Mean 5.9 (SD 6.3)	1072	39 (4)
		Previous negative							Mean 64 (SD 7)	Mean 9.4 (SD 6.9)	1072	30 (3)
Boesen[27]	2018	Biopsy naïve	NR	NR	BP	NR	NR	NR	NR	NR	1020	305 (30)
Bryant[28]	2018	Biopsy naïve	NR	Both	mpMRI	Academic	PIRADS	TRUS 12 core	NR	NR	792	278 (35)
Hansen[29]	2018	Biopsy naïve	GE	NR	mpMRI	Academic	PIRADS	TP 24 core	NR	NR	807	236 (29)
Hwang[30]	2018	Previous negative	NR	3	mpMRI		PIRADS	TRUS 12 core	NR	NR	39	15 (39)
Karman[31]	2018	Previous negative	NR	NR	mpMRI	NR	PIRADS	TRUS 24 core	NR	NR	714	318 (45)
Kim[32]	2018	Biopsy naïve	Philips	3	mpMRI	NR	Mixed	TRUS	Mean 63 (Range 30–86)	Mean 9.67 (Range 0.83–997.1)	NR	324 (NA)
		Previous negative							Mean 62 (Range 40–78)	Mean 7.89 (Range 0.6–66.35)	NR	161 (NA)
Lobo[33]	2018	Biopsy naïve	NR	Both	mpMRI	Academic	PIRADS	TRUS 12 core	NR	NR	900	404 (45)
Mannaerts[34]	2018	Biopsy naïve	NR	Both	mpMRI	Academic	PIRADS	TRUS	Median 64 (IQR 59–68)	Median 6.4 (IQR 5.1–9.1)	200	96 (48)
Meng[35]	2018	Biopsy naïve	Siemens	NR	mpMRI	NR	PIRADS	TRUS 12 core	Mean 61.87 (SD 7.9)	NR	NR	52 (NA)
Morote[36]	2018	Biopsy naïve	Siemens	3	mpMRI	NR	PIRADS	TRUS 12 core	Median 68 (IQR 61–73)	Median 5.8 (IQR 4.3–8.7)	549	141 (26)
		Previous negative									219	45 (21)
Mortezavi[37]	2018	Biopsy naïve	Mixed	3	mpMRI	Academic	PIRADS	TPM 24 core	Median 63 (IQR 57-68)	Median 5.8 (IQR 4.4-8.9)	163	49 (30)
		Previous negative							Median 64 (IQR 60-69)	Median 8.6 (IQR 5.7-13)	86	36 (412)
Oishi[38]	2018	Biopsy naïve	NR	3	mpMRI	Academic	PIRADS	TRUS 12 core	NR	NR	1149	48 (2)
		Previous negative							NR	NR	1149	60 (5)
Otti[39]	2018	Biopsy naïve		1.5	mpMRI	Academic	PIRADS	Mixed	NR	NR	NR	348 (NA)
Pal[40]	2018	Biopsy naïve	NR	1.5	mpMRI	Academic	PIRADS	TPM	Median 64 (IQR 41-77)	NR	NR	208 (NA)
Panebianco[41]	2018	Biopsy naïve	NR	3	mpMRI	Academic	PIRADS	SB or Saturation	Median 66 (IQR 62-69)	Median 5.9 (IQR 3.9-7.6)	659	NR (NA)

		Previous negative							Median 68 (IQR 60-72)	Median 5.6 (IQR 3.2-7.8)	596	NR (NA)
Simmons[42]	2018	Previous negative	NR	3	mpMRI	Academic	Likert	TPM	NR	NR	249	49 (20)
Zhou[43]	2018	Biopsy naïve	NR	3	mpMRI	NR	PIRADS	TP saturation	NR	NR	219	66 (30)
Barrett[44]	2019	Biopsy naïve	GE	Both	mpMRI	Academic	PIRADS	TPM 24 core or TRUS 12 core	Median 63 (IQR 58-67)	Median 6 (IQR 4.6-8.9)	833	484 (58)
Elkhoury[45]	2019	Biopsy naïve	Siemens	3	mpMRI	Academic	PIRADS	TRUS 12 core	Mean 63.6 (SD 5.9)	Median 5.2 (IQR 4.1-6.6)	300	52 (17)
Kruger-Stokke[46]	2019	Biopsy naïve	NR	NR	mpMRI	NR	PIRADS	TRUS	NR	NR	201	101 (50)
Regis[47]	2019	Biopsy naïve	Siemens	3	mpMRI	NR	PIRADS	TRUS 12 core	Median 66 (IQR 51–81)	Median 6.1 (IQR 1.8–67.0)	NR	93
		Previous negative									NR	129
Rouviere[48]	2019	Biopsy naïve	NR	Both	mpMRI	Multi	PIRADS	TRUS 12 core	Median 64 (IQR 59-68)	Median 6.5 (IQR 5.6-9.6)	266	45 (17)
Rozas[49]	2019	Biopsy naïve	Siemens	3	mpMRI	NR	PIRADS	TRUS 18 core	NR	NR	342	201 (59)
van der Leest[50]	2019	Biopsy naïve	Siemens	3	mpMRI	Multi	PIRADS	TRUS 12 core	Median 65 (IQR 59–68)	Median 6.4 (IQR 5.0–8.6)	626	309 (50)
Zalesky[51]	2019	Biopsy naïve	GE	1.5	mpMRI	Academic	PIRADS	TRUS 12 core	Median 62 (IQR 56-67)	Median 5.4 (IQR 4.1-7.3)	223	53 (24)
		Previous negative							Median 65 IQR (60-69)	Median 8.3 (IQR 6.1-13.5)	174	25 (14)
Zhang[52]	2019	Biopsy naïve	NR	3	mpMRI	NR	PIRADS	TRUS 24 core	NR	NR	NR	273

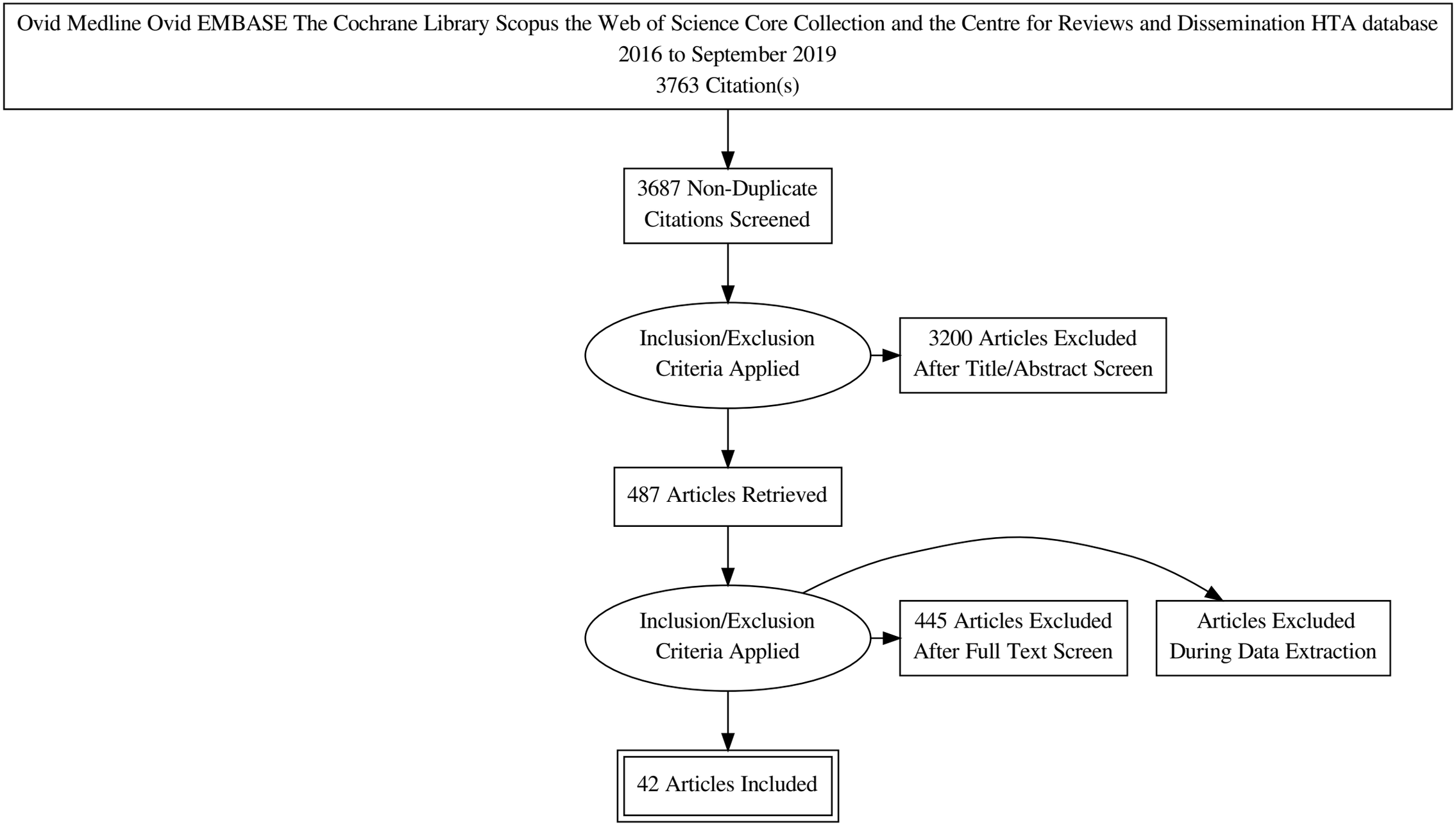
Figure Legends:

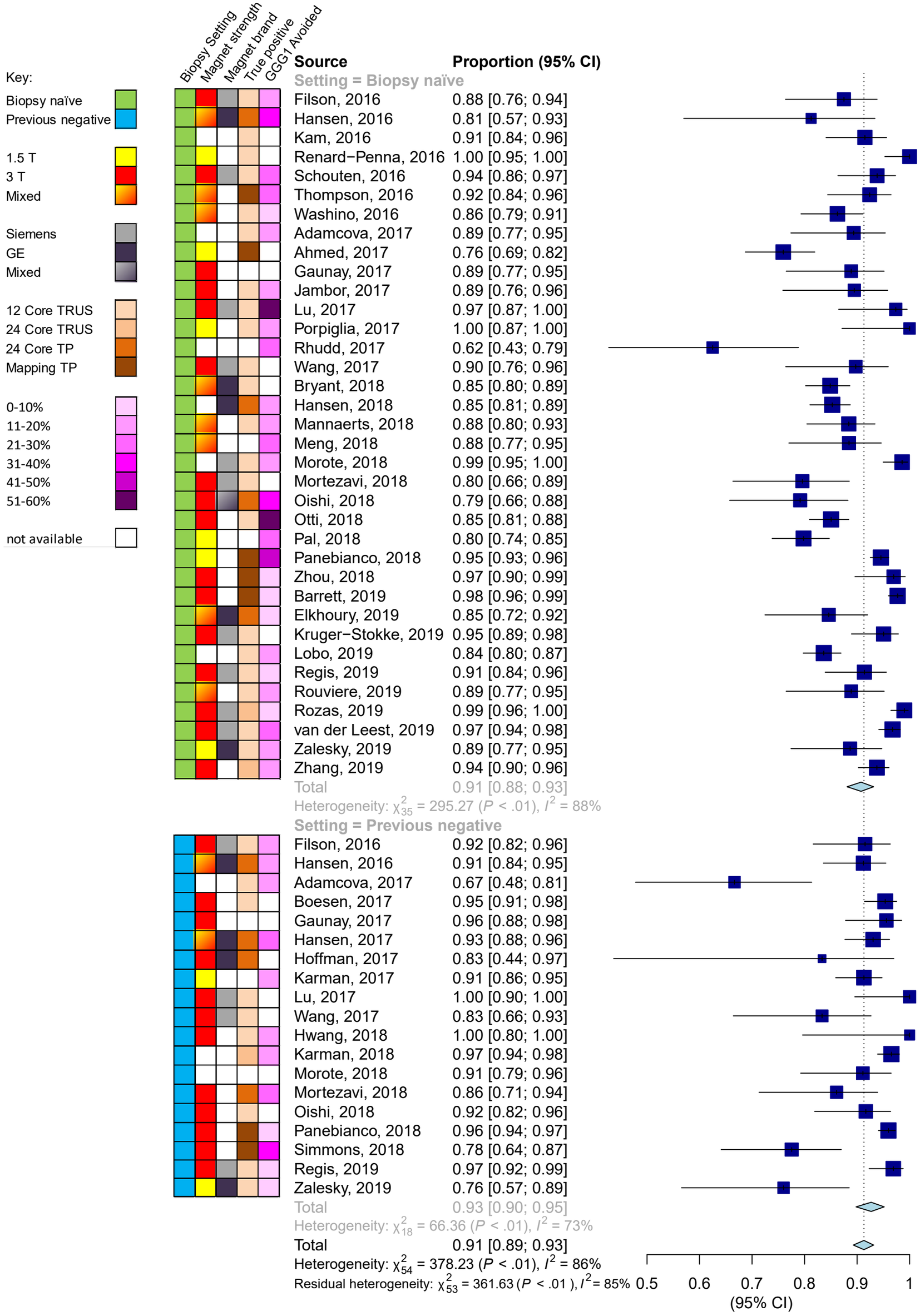
Figure 1. Study selection flow chart

Figure 2. Forest plot of NPV using primary definition (MRI negative = PIRADS 1-2, csPCa = Gleason score ≥3+4=7 (GGG 2))

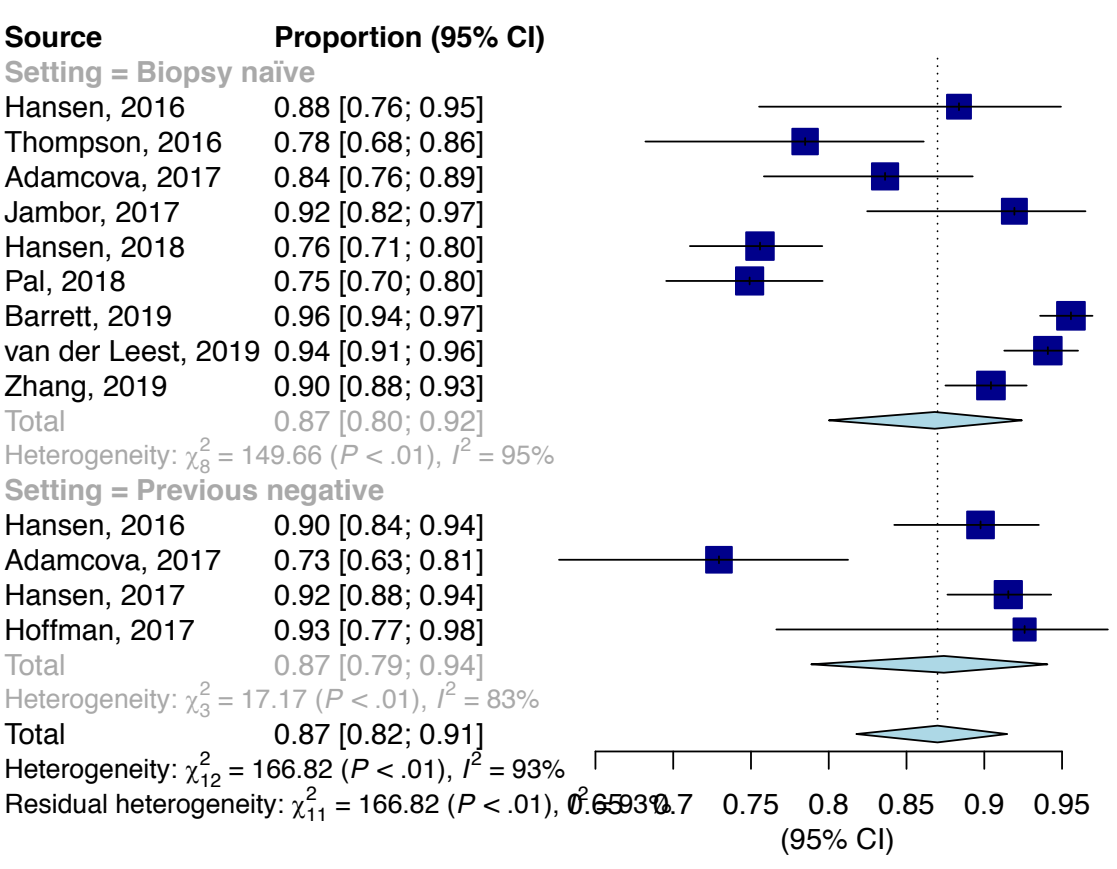
Figure 3. Secondary definition forest plots of NPV: (A) MRI negative = PIRADS 1-3, csPCa = Gleason Grade Group ≥2 (Gleason Score 3+4=7), (B) MRI negative = PIRADS 1-2, csPCa = Gleason Grade Group ≥3 (Gleason Score ≥4+3=7), (C) MRI negative = PIRADS 1-3, csPCa = (Gleason Grade Group ≥3 (Gleason Score ≥4+3=7)

Figure 4. Proportion of Gleason Grade Group 1 (Gleason Score 3+3=6) negative MRI (PI-RADS 1 and 2). This represents avoided diagnoses of non-clinically significant prostate cancer.

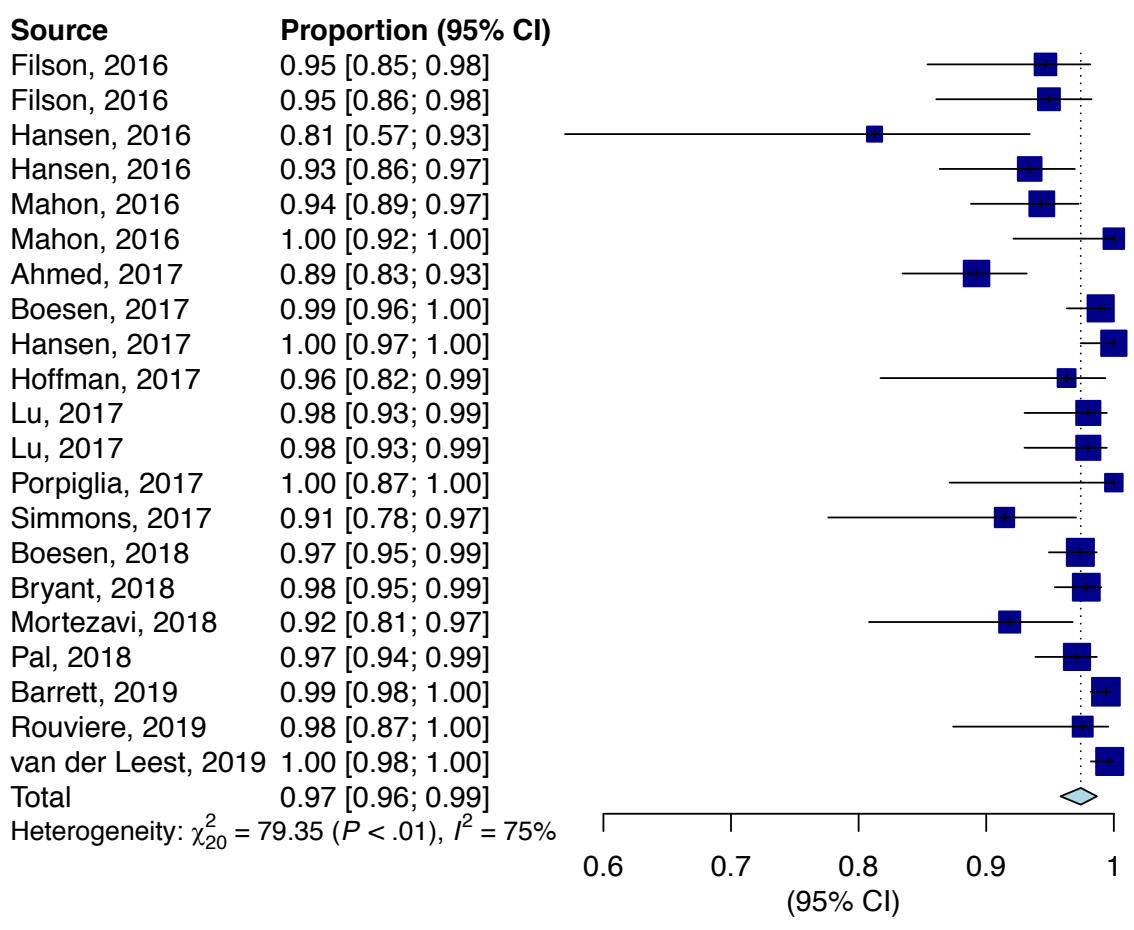

Ovid Medline Ovid EMBASE The Cochrane Library Scopus the Web of Science Core Collection and the Centre for Reviews and Dissemination HTA database
2016 to September 2019
3763 Citation(s)
3687 Non-Duplicate Citations Screened
Inclusion/Exclusion Criteria Applied
3200 Articles Excluded After Title/Abstract Screen
487 Articles Retrieved
Inclusion/Exclusion Criteria Applied
445 Articles Excluded After Full Text Screen
Articles Excluded During Data Extraction
42 Articles Included


Biopsy Setting
Magnet strength
Magnet brand
True positive
GGG1 Avoided
Key:
Biopsy naïve
Previous negative
1.5 T
3 T
Mixed
Siemens
GE
Mixed
12 Core TRUS
24 Core TRUS
24 Core TP
Mapping TP
0-10%
11-20%
21-30%
31-40%
41-50%
51-60%
not available
Source
Proportion (95% CI)
Setting = Biopsy naïve
Filson, 2016 0.88 [0.76; 0.94]
Hansen, 2016 0.81 [0.57; 0.93]
Kam, 2016 0.91 [0.84; 0.96]
Renard−Penna, 2016 1.00 [0.95; 1.00]
Schouten, 2016 0.94 [0.86; 0.97]
Thompson, 2016 0.92 [0.84; 0.96]
Washino, 2016 0.86 [0.79; 0.91]
Adamcova, 2017 0.89 [0.77; 0.95]
Ahmed, 2017 0.76 [0.69; 0.82]
Gaunay, 2017 0.89 [0.77; 0.95]
Jambor, 2017 0.89 [0.76; 0.96]
Lu, 2017 0.97 [0.87; 1.00]
Porpiglia, 2017 1.00 [0.87; 1.00]
Rhudd, 2017 0.62 [0.43; 0.79]
Wang, 2017 0.90 [0.76; 0.96]
Bryant, 2018 0.85 [0.80; 0.89]
Hansen, 2018 0.85 [0.81; 0.89]
Mannaerts, 2018 0.88 [0.80; 0.93]
Meng, 2018 0.88 [0.77; 0.95]
Morote, 2018 0.99 [0.95; 1.00]
Mortezavi, 2018 0.80 [0.66; 0.89]
Oishi, 2018 0.79 [0.66; 0.88]
Otti, 2018 0.85 [0.81; 0.88]
Pal, 2018 0.80 [0.74; 0.85]
Panebianco, 2018 0.95 [0.93; 0.96]
Zhou, 2018 0.97 [0.90; 0.99]
Barrett, 2019 0.98 [0.96; 0.99]
Elkhoury, 2019 0.85 [0.72; 0.92]
Kruger−Stokke, 2019 0.95 [0.89; 0.98]
Lobo, 2019 0.84 [0.80; 0.87]
Regis, 2019 0.91 [0.84; 0.96]
Rouviere, 2019 0.89 [0.77; 0.95]
Rozas, 2019 0.99 [0.96; 1.00]
van der Leest, 2019 0.97 [0.94; 0.98]
Zalesky, 2019 0.89 [0.77; 0.95]
Zhang, 2019 0.94 [0.90; 0.96]
Total 0.91 [0.88; 0.93]
Heterogeneity: χ^2_{35} = 295.27 (P < .01), I^2 = 88%
Setting = Previous negative
Filson, 2016 0.92 [0.82; 0.96]
Hansen, 2016 0.91 [0.84; 0.95]
Adamcova, 2017 0.67 [0.48; 0.81]
Boesen, 2017 0.95 [0.91; 0.98]
Gaunay, 2017 0.96 [0.88; 0.98]
Hansen, 2017 0.93 [0.88; 0.96]
Hoffman, 2017 0.83 [0.44; 0.97]
Karman, 2017 0.91 [0.86; 0.95]
Lu, 2017 1.00 [0.90; 1.00]
Wang, 2017 0.83 [0.66; 0.93]
Hwang, 2018 1.00 [0.80; 1.00]
Karman, 2018 0.97 [0.94; 0.98]
Morote, 2018 0.91 [0.79; 0.96]
Mortezavi, 2018 0.86 [0.71; 0.94]
Oishi, 2018 0.92 [0.82; 0.96]
Panebianco, 2018 0.96 [0.94; 0.97]
Simmons, 2018 0.78 [0.64; 0.87]
Regis, 2019 0.97 [0.92; 0.99]
Zalesky, 2019 0.76 [0.57; 0.89]
Total 0.93 [0.90; 0.95]
Heterogeneity: χ^2_{18} = 66.36 (P < .01), I^2 = 73%
Total 0.91 [0.89; 0.93]
Heterogeneity: χ^2_{54} = 378.23 (P < .01), I^2 = 86%
Residual heterogeneity: χ^2_{53} = 361.63 (P < .01), I^2 = 85%
0.5
0.6
0.7
0.8
0.9
1
(95% CI)

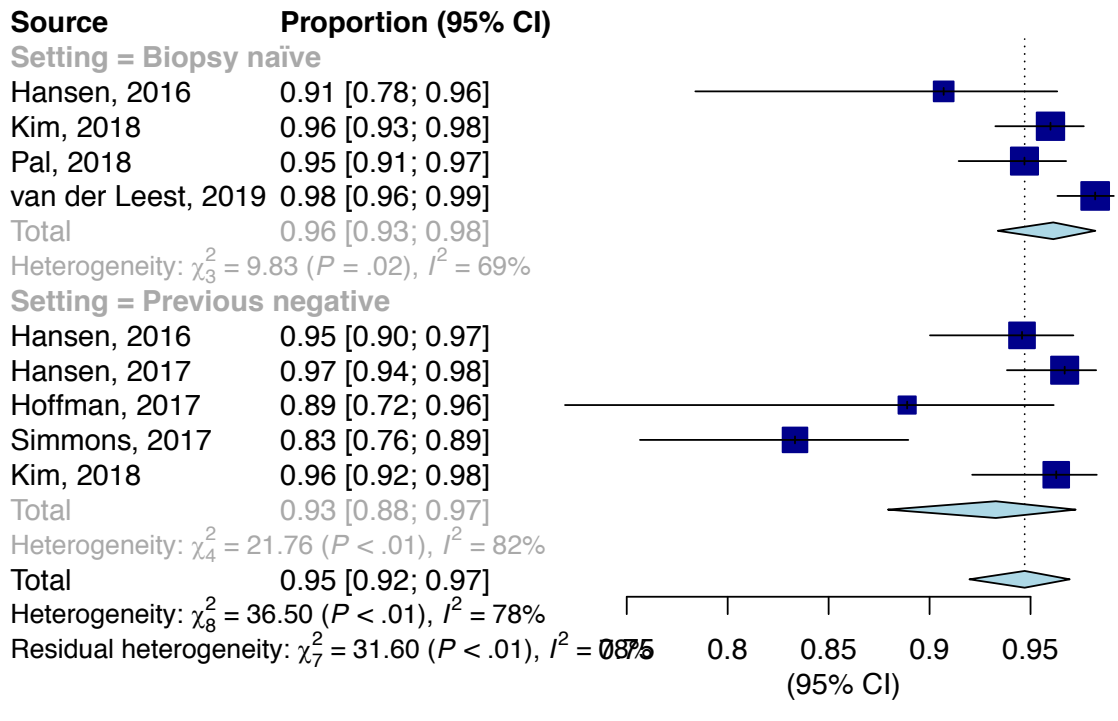
(A)

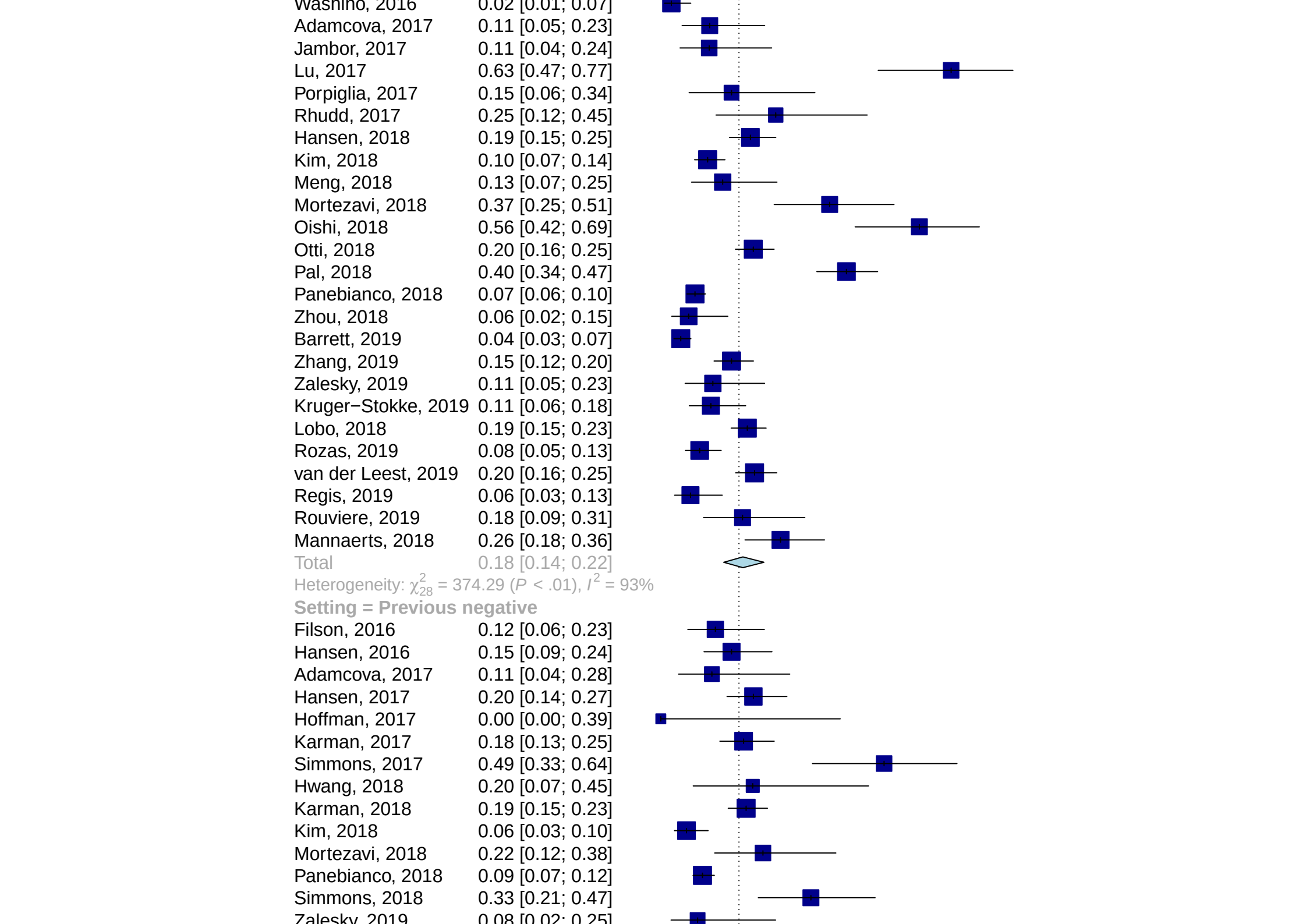

Source
Proportion (95% CI)
Setting = Biopsy naïve
Hansen, 2016 0.88 [0.76; 0.95]
Thompson, 2016 0.78 [0.68; 0.86]
Adamcova, 2017 0.84 [0.76; 0.89]
Jambor, 2017 0.92 [0.82; 0.97]
Hansen, 2018 0.76 [0.71; 0.80]
Pal, 2018 0.75 [0.70; 0.80]
Barrett, 2019 0.96 [0.94; 0.97]
van der Leest, 2019 0.94 [0.91; 0.96]
Zhang, 2019 0.90 [0.88; 0.93]
Total 0.87 [0.80; 0.92]
Heterogeneity: χ^2_8 = 149.66 (P < .01), I^2 = 95%
Setting = Previous negative
Hansen, 2016 0.90 [0.84; 0.94]
Adamcova, 2017 0.73 [0.63; 0.81]
Hansen, 2017 0.92 [0.88; 0.94]
Hoffman, 2017 0.93 [0.77; 0.98]
Total 0.87 [0.79; 0.94]
Heterogeneity: χ^2_3 = 17.17 (P < .01), I^2 = 83%
Total 0.87 [0.82; 0.91]
Heterogeneity: χ^2_{12} = 166.82 (P < .01), I^2 = 93%
Residual heterogeneity: χ^2_{11} = 166.82 (P < .01), I^2 = 93%
0.7
0.75
0.8
0.85
0.9
0.95
(95% CI)

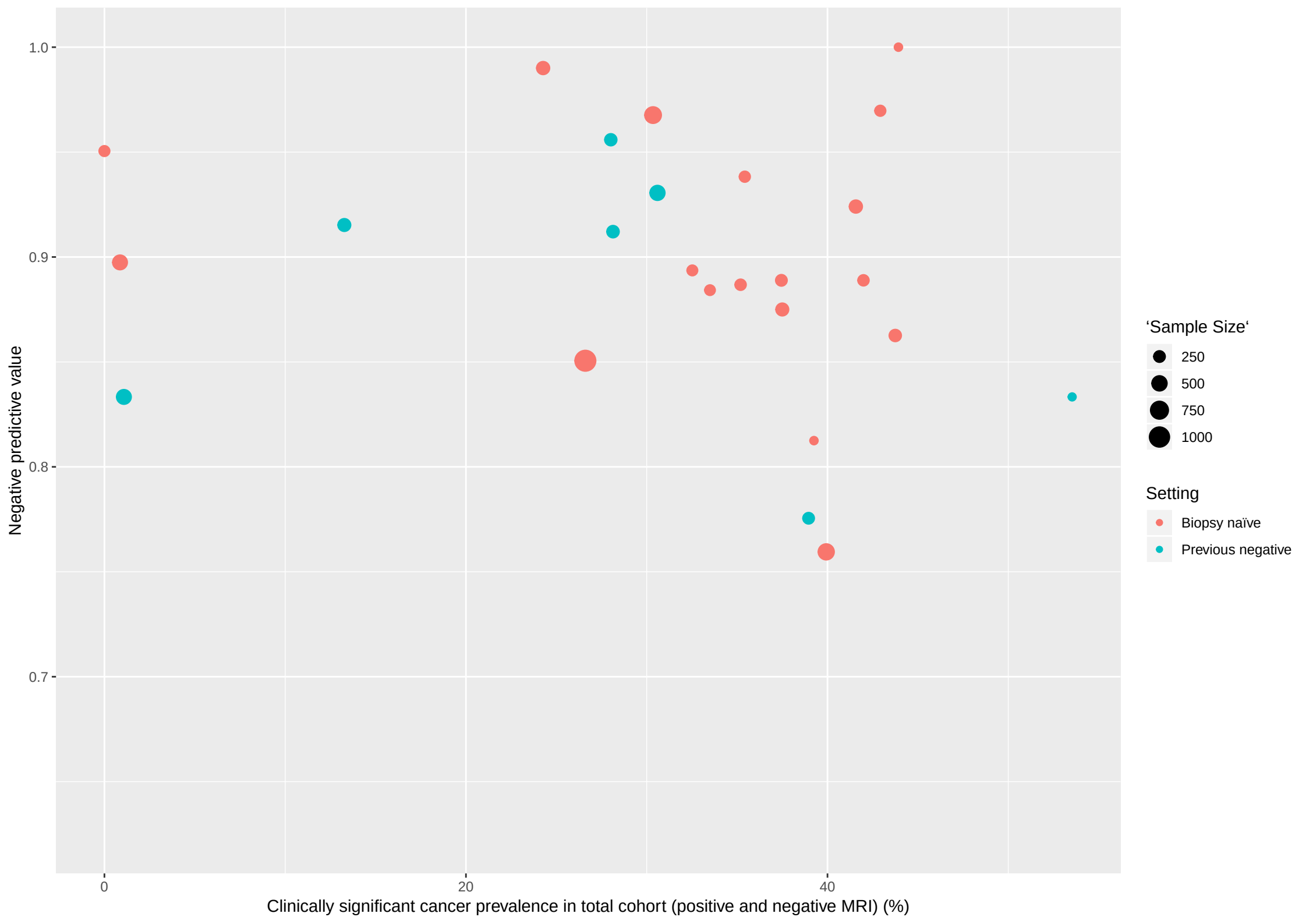
(B)


Source
Proportion (95% CI)
Filson, 2016 0.95 [0.85; 0.98]
Filson, 2016 0.95 [0.86; 0.98]
Hansen, 2016 0.81 [0.57; 0.93]
Hansen, 2016 0.93 [0.86; 0.97]
Mahon, 2016 0.94 [0.89; 0.97]
Mahon, 2016 1.00 [0.92; 1.00]
Ahmed, 2017 0.89 [0.83; 0.93]
Boesen, 2017 0.99 [0.96; 1.00]
Hansen, 2017 1.00 [0.97; 1.00]
Hoffman, 2017 0.96 [0.82; 0.99]
Lu, 2017 0.98 [0.93; 0.99]
Lu, 2017 0.98 [0.93; 0.99]
Porpiglia, 2017 1.00 [0.87; 1.00]
Simmons, 2017 0.91 [0.78; 0.97]
Boesen, 2018 0.97 [0.95; 0.99]
Bryant, 2018 0.98 [0.95; 0.99]
Mortezavi, 2018 0.92 [0.81; 0.97]
Pal, 2018 0.97 [0.94; 0.99]
Barrett, 2019 0.99 [0.98; 1.00]
Rouviere, 2019 0.98 [0.87; 1.00]
van der Leest, 2019 1.00 [0.98; 1.00]
Total 0.97 [0.96; 0.99]
Heterogeneity: χ^2_{20} = 79.35 (P < .01), I^2 = 75%
0.6
0.7
0.8
0.9
1
(95% CI)

(C)


Source
Proportion (95% CI)
Setting = Biopsy naïve
Hansen, 2016 0.91 [0.78; 0.96]
Kim, 2018 0.96 [0.93; 0.98]
Pal, 2018 0.95 [0.91; 0.97]
van der Leest, 2019 0.98 [0.96; 0.99]
Total 0.96 [0.93; 0.98]
Heterogeneity: χ^2_3 = 9.83 (P = .02), I^2 = 69%
Setting = Previous negative
Hansen, 2016 0.95 [0.90; 0.97]
Hansen, 2017 0.97 [0.94; 0.98]
Hoffman, 2017 0.89 [0.72; 0.96]
Simmons, 2017 0.83 [0.76; 0.89]
Kim, 2018 0.96 [0.92; 0.98]
Total 0.93 [0.88; 0.97]
Heterogeneity: χ^2_4 = 21.76 (P < .01), I^2 = 82%
Total 0.95 [0.92; 0.97]
Heterogeneity: χ^2_8 = 36.50 (P < .01), I^2 = 78%
Residual heterogeneity: χ^2_7 = 31.60 (P < .01), I^2 = 78%
0.8
0.85
0.9
0.95
(95% CI)


Washino, 2016 0.02 [0.01; 0.07]
Adamcova, 2017 0.11 [0.05; 0.23]
Jambor, 2017 0.11 [0.04; 0.24]
Lu, 2017 0.63 [0.47; 0.77]
Porpiglia, 2017 0.15 [0.06; 0.34]
Rhudd, 2017 0.25 [0.12; 0.45]
Hansen, 2018 0.19 [0.15; 0.25]
Kim, 2018 0.10 [0.07; 0.14]
Meng, 2018 0.13 [0.07; 0.25]
Mortezavi, 2018 0.37 [0.25; 0.51]
Oishi, 2018 0.56 [0.42; 0.69]
Otti, 2018 0.20 [0.16; 0.25]
Pal, 2018 0.40 [0.34; 0.47]
Panebianco, 2018 0.07 [0.06; 0.10]
Zhou, 2018 0.06 [0.02; 0.15]
Barrett, 2019 0.04 [0.03; 0.07]
Zhang, 2019 0.15 [0.12; 0.20]
Zalesky, 2019 0.11 [0.05; 0.23]
Kruger−Stokke, 2019 0.11 [0.06; 0.18]
Lobo, 2018 0.19 [0.15; 0.23]
Rozas, 2019 0.08 [0.05; 0.13]
van der Leest, 2019 0.20 [0.16; 0.25]
Regis, 2019 0.06 [0.03; 0.13]
Rouviere, 2019 0.18 [0.09; 0.31]
Mannaerts, 2018 0.26 [0.18; 0.36]
Total 0.18 [0.14; 0.22]
Heterogeneity: χ^2_{28} = 374.29 (P < .01), I^2 = 93%
Setting = Previous negative
Filson, 2016 0.12 [0.06; 0.23]
Hansen, 2016 0.15 [0.09; 0.24]
Adamcova, 2017 0.11 [0.04; 0.28]
Hansen, 2017 0.20 [0.14; 0.27]
Hoffman, 2017 0.00 [0.00; 0.39]
Karman, 2017 0.18 [0.13; 0.25]
Simmons, 2017 0.49 [0.33; 0.64]
Hwang, 2018 0.20 [0.07; 0.45]
Karman, 2018 0.19 [0.15; 0.23]
Kim, 2018 0.06 [0.03; 0.10]
Mortezavi, 2018 0.22 [0.12; 0.38]
Panebianco, 2018 0.09 [0.07; 0.12]
Simmons, 2018 0.33 [0.21; 0.47]
Zalesky, 2019 0.08 [0.02; 0.25]


1.0
0.9
0.8
0.7
Negative predictive value
0
20
40
Clinically significant cancer prevalence in total cohort (positive and negative MRI) (%)
‘Sample Size‘
250
500
750
1000
Setting
Biopsy naïve
Previous negative