Peer Review File

Article information: https://dx.doi.org/10.21037/tau-22-640

Reviewer A

High PSA levels are a common reason for urologic consultation request and it can cause oversaturation. This study addresses an important issue: methods to improve the referrals to the specialist. However, this research should be interpreted with caution and some considerations should be taken into account:

1) The authors stated that stratification in primary care based on the RPCRC using TAUS prostate volume would have avoided 29 out of the 41 referrals, at the expense of non-referral of 5 out of 11 men with a biopsy indication, according to the urologist.

It would be very interesting to provide information of biopsy outcomes. For example, was any clinically significant prostate cancer (Gleason score ≥7 and clinical stage >T2b) underdiagnosed due to this tool used by GPs?

In my opinion, this research should use the AUC to report the results using the biopsy outcomes as 'gold standard'.

Reply 1: We would like to thank the reviewer for this critical appraisal of our manuscript. An AUC using the biopsy outcomes would indeed be of interest. However, this was not within the scope of our study as not every participant underwent a prostate biopsy. Therefore, it is not possible to use the biopsy outcomes as the reference test. The reason why not everyone received a prostate biopsy is that the primary aim was to compare TAUS prostate volume with TRUS prostate volume within the RPCRC, and not an external validation of the RPCRC itself. An external validation of the RPCRC is reported in doi: 10.21037/tau.2017.12.11 & https://doi.org/10.1016/S0302-2838(22)00456-0 (conclusion: risk stratification with the RPCRC in primary care could prevent half of referrals, 5% considered low risk had PCa at later biopsy all GS3+3) . In addition, using the validated RPCRC (AUC is 0.78 and 0.90 for any PCa and significant PC (reference 10)), the indication for referral was compared to the indication for biopsy. In the men that did not have an referral indication according to TAUS risk stratification but had an indication for biopsy according to TRUS risk stratification (n=5), this resulted in only one clinically insignificant prostate cancer and no missed clinically significant prostate cancers against 71% avoided referrals.

Changes in the text:

In line 156 was added that no cancer was found in the 'missed referrals'.

2) A further discussion regarding the utility of the Rotterdam Prostate Cancer Risk Calculators (RPCRCs) (www.prostatecancer-riskcalculator.org) is advisable. For example, the EAU guidelines indicate that: "In asymptomatic men with a PSA level between 2–10 ng/mL and a normal DRE, use one of the following tools for biopsy indication:

• risk-calculator;

• Magnetic resonance imaging of the prostate"

Data about AUCs of the tool used (RPCRC, included such variables) should be mentioned and it should be compared with MRI.

Reply 2: It is indeed true that the guideline currently recommends to use MRI in the diagnostic pathway for prostate cancer diagnosis. In recent publications from the EAU (https://doi.org/10.1016/j.euo.2021.06.006) it is however clearly mentioned that risk stratification should precede MRI. If this step is omitted the result will be many unnecessary MRI's and potential prostate biopsies. This is confirmed by the data of the STHLM-3 screening trial (DOI: 10.1056/NEJMoa2100852) where MRI was provided in all men with a PSA >= 3.0 ng/ml, resulting in 56% of MRI being normal.

As mentioned above, the aim of this study was to compare TAUS prostate volume measurement by the general practitioner with TRUS volume measurement by the urologist. Since TRUS (and MRI) are not readily available in the primary care setting, we hypothesized that TAUS could be a tool to reduce referrals to the urologist and radiologist. In addition, this study was conducted before the wide adaption of the prostate mpMRI by urologists, so therefore the TAUS prostate volume could unfortunately not be compared to MRI prostate volume. This is also mentioned in the methods (line 94).

Changes in the text: limitations paragraph in the discussion was rewritten

3) Finally, the conclusion should be modified unless the authors would provide the required information regarding diagnosis accuracy, because the term "unnecessary referral" is not appropriate if we do not know the number of prostate cancers underdiagnosed.

This study needs to be improved before publication; especially, it needs to include AUC. It informs about a predictive tool (RPCRC using TAUS prostate volume estimated by GPs) that has not been properly evaluated.

Reply 3: We thank the reviewer for this comment. We agree that as not all patients have been biopsied, so in this study we don't know the exact number of underdiagnosed prostate cancers. However, the RPCRC which was also applied in this study, has previously showed that its ability in the primary care setting to reduce referrals and prostate biopsies while missing almost no significant prostate cancers after a median follow-up of 43 months (10.21037/tau.2017.12.11 & https://doi.org/10.1016/S0302-2838(22)00456-0). Therefore, it can be derived that TAUS may also have this effect as the prostate volume accuracy compared to TRUS increased with increased operator experience.

Changes in the text: limitations paragraph in the discussion was rewritten and "could" was replaced with "may" in the conclusion.

Reviewer B

This is a report on an interesting and potentially clinically useful method to select men who have a PSA test in general practice for urology referral, based on transabdominal ultrasound

(TAUS) measurement of the prostate volume.

Major concerns:

1) The comparison between TAUS and TRUS (the reference test) is based on a reasonable sample size calculation but the numbers of study subjects in the subsequent steps are low. This is particularly problematic as they were not investigated according to current clinical practice (prostate MRI) and that only a minority had a prostate biopsy when indicated. The study therefore cannot evaluate the risk of missing clinically significant cancer with a diagnostic pathway based on TAUS for referral selection. I recommend that all parts relating to biopsy and cancer detection are omitted.

Reply 1: We would like to thank the reviewer for the critical appraisal of our manuscript. We agree that our study is limited by the unavailability to compare TAUS with MRI prostate volume measurements. However, the clinical relevance between TRUS- and MRI prostate volume seems to be marginal (doi: 10.1016/j.urology.2022.09.007).
We also agree that it would be interesting to know biopsy results of all participants. However, this was beyond the scope of research, as the primary aim is to avoid referrals to secondary care, i.e. also to circumvent unnecessary MRI's. So the assumption of the reviewer is exactly what we want: MRI should not be and very often cannot be current clinical practice (in less developed regions) (see recent publication in Eur Urol (https://doi.org/10.1016/j.euo.2021.06.006)). Additionally, a study regarding risk stratification in primary care setting and biopsy outcomes has been published earlier and that risk stratification with the RPCRC in primary care could prevent half of referrals, 5% considered low risk had PCa at later biopsy all GS3+3) (doi: 10.21037/tau.2017.12.11; this paper is also mentioned in the discussion).
Changes in the text: We rewrote the limitations paragraph in the discussion.

2) As a consequence of the above it's not justified to claim that the TAUS pathway can "reduce unnecessary referrals" (title, conclusions). Some patients who were not referred may have had a high grade cancer.
Reply 2: As also mentioned in reply 1, we agree that as not all patients have been biopsied, we don't know the exact number of underdiagnosed prostate cancers. Therefore, we changed the title and nuanced our conclusions.
Changes in the text: Title was changed and conclusions were nuanced. It was also added as a limitation in the discussion.

3) The non-inferiority limit 25% is in my opinion too high. Accepting an error such as 61 cc for an 80 cc prostate makes a big difference for PSA density. This may be discussed as a weakness of the study.
Reply 3: We thank the reviewer or this comment. We hypothesized that a non-inferiority limit 25% was sufficient, as earlier research also showed that the RPCRC showed good performance with the categorization in 3 volume classes: 25ml ((<30ml), 40ml (30–50ml), or 60ml (>50ml). DOI: 10.1111/iju.13442: Especially in lower prostate volumes a small difference due to interobserver variability results in a large percentual difference, but did not show a significant impact in cancer detection. In larger prostates the TAUS inaccuracy

was indeed decreased as is mentioned in our discussion (line 176).

4) Methods: The included study subject had an "opportunistic PCa screening request". Does this mean that none had lower urinary tract symptoms (LUTS)? This must be clarified as men with LUTS on average have bigger prostates and the positive predictive value of a raised PSA is lower in men with LUTS. If no men with LUTS were included, this should be mentioned as a limitation as the results are then not applicable to men with LUTS. If men with LUTS were included, all study subjects did probably not have an opportunistic PCa screening request but were investigated for LUTS (which is probably the most common reason for PSA testing in primary care). That 39% of the men were referred makes me believe that the study population was enriched by men with LUTS because of BPH, which affects the results and external validity.
Reply 4: The included study subjects all had a primarily opportunistic PCa screening request. In this study it was not assessed whether this request had a underlying LUTS. We agree that a PCa screening wish can coincide with or even emerge from LUTS and that in the study population both can be present. Nonetheless, we believe that this study population does reflect a real life cohort in the primary care setting, because the mean prostate volume of 45cc (TRUS) in this cohort is very close to the mean prostate volume of 39 (also TRUS) in the screening cohort of the ERSPC study, which can be seen as a reflection of the average male population aged 54-74 in the Netherlands.

5) Methods: What did the 2-hour training session include? How many TAUS procedures? Were the TAUS volumes compared with TRUS volumes?
Reply 5: The two GP's underwent two separate courses (total of 2 hours) supervised by an experienced urologist in which they each conducted at least 10 TAUS prostate volume measurements in 10 different test subjects. These test subjects did not undergo TRUS volumes as this was not ethical since they were not patients and there was no indication for TRUS. We did also mention in the limitations that as the urologist did not perform TAUS in the study subjects, no comparison could be made between TAUS prostate volume measurement by the GP's and TAUS prostate volume measurement by the urologist.

6) Methods: Figure 2 includes categorization of TAUS quality, but this assessment is not mentioned in Methods. Please add. Wouldn't it make sense to exclude study subjects with a poor TAUS quality from the TAUS pathway?
Reply 6: We thank the reviewer for pointing out this issue. Nonetheless, in real clinical practice there will also be ultrasounds with a poor quality, so we think that by including them it gives a good reflection of daily clinical practice. It also shows the importance of the learning curve.
Changes in the text: We included the categorisation of TAUS quality in the method section.

7) Methods: Why were the TAUS-measured prostate volumes categorized as 25ml (<30ml), 40ml (30–50ml), or 60ml (>50ml)? This reduces the value of the volume measurement.
Reply 7: We thank the reviewer for this comment. This was done because DRE prostate volume in statistical model of the RPCRC is also categorized to these values and has shown to hardly affect predictive capability (DOI: 10.1111/iju.13442 & DOI:

10.1016/j.eururo.2011.11.012).
Changes in the text: we clarified this in the methods section (line 114)

8) Methods: It's difficult to understand the flow of study subjects. A flow chart would be helpful.
Reply 8: We appreciate this comment.
Changes in the text: We added a study flowchart (fig 1)

9) Methods: The exclusion criteria include TAUS consultation (n=5) and urologist consultation (n=14). Please explain what these events are and why they lead to exclusion.
Reply 9: This means that these men did not show at either the TAUS consultation or the consultation at the urologist, resulting in missing information about the main outcomes (TAUS and/or TRUS). Therefore, these men could not be analysed and were excluded.
Changes in the text: these men were also clarified in the flowchart

10: Results: Proportions should be presented with 95% CIs.
Changes in the text: We added the 95% CI.

11. Results: It is not only important to report the number of men who were referred by the different pathways, but also whether they were the same men. Were the groups totally overlapping or not?
Changes in the text: We clarified this in the flowchart.

12. Conclusion: The study results do not allow for the conclusion that "The accuracy of the prostate cancer risk assessment in primary care is likely to be improved by centralizing opportunistic screening consultations within a regional care pathway" as this was not evaluated.
Reply 12: We thank the reviewer for this comment. We concluded this from the fact that we observed a learning curve. Therefore it may be concluded that the accuracy improves as the operator experience improves, which is the case when the consultations are centralized.

Minor comments:
13. "Reference test" is a better term than "golden standard", particularly as MRI is now considered the equivalent of the gold standard for prostate volume measurement.
Reply 13: We agree.
Changes in the text: We changed the term to reference test

14. Lines 42-43: "centralization to achieve a higher volume of consultations in primary care facilities." Isn't primary care decentralized per definition?
Reply 14: Primary care is indeed decentralized per general practioner's office, but there are collaborations between multiple general practioner's offices to optimize health care pathways and reduce health care costs (e.g. laboratory or radiology diagnostics).

15. Line 141: TAUS was not "inequivalent" but non-inferior.

Reply 15: We appreciate the comment of the reviewer but we respectfully disagree. The 95% confidence interval of the mean difference between TAUS and TRUS exceeded the positive equivalence margin. Thus, TAUS is non-inferior to TRUS, but rather inequivalent, because of its overestimation.

16. Lines 180-182: This statement is clearly not justified: "In this study, there were no negative clinical consequences for these non-referred men as none had clinically significant prostate cancer." The study design makes it impossible to exclude that high grade or even advanced cancers were missed in non-referred men.

Reply 16: We thank the reviewer for pointing out this relevant issue.

Changes in the text: We nuanced this proposition to: Although, no clinically significant prostate cancer was found in the "missed referrals" who were biopsied, the clinical consequences of TAUS prostate volume measurement with RPCRC-based risk stratification in such a low-risk primary care target population need to be studied further in a larger cohort with longer follow-up.

17. Lines 204-205: "Only 41 out of 105 (39%) men included for analysis had an indication for referral to the urologist according to clinical standard of care." Why "only"? The 39% are far more than I would've expected and make me believe that the study population was not representative for asymptomatic men who have a PSA test in primary care. Usually 5-15% of asymptomatic middle-aged men have a raised PSA, depending on their age. Most likely, the study population was men with LUTS because of BPH.

Reply 17: We thank the reviewer for this comment and changed this in the paragraph with the study limitations.

18. Figure 2: I'm colour blind and so are 7% of other men. I cannot discriminate the green (?) from the yellow (?) dots. Please use other colours.

Reply 18: A very relevant comment.

Changes in the text: We changed the figure.