Manuscript
Click here to view linked References

Standardized MRI reporting using the PRECISE criteria and MRI/TRUS-fusion with transperineal saturation biopsy to select men on active surveillance

Svenja Dieffenbacher, MD[1,2*]; Joanne Nyarangi-Dix, MD[1*], Francesco Giganti, MD[3,4], David Bonekamp, MD[2]; Claudia Kesch, MD[1,5]; Maya B. Müller-Wolf, MD[2]; Viktoria Schütz, MD[1]; Claudia Gasch, MD[1]; Gencay Hatiboglu, MD[1]; Marcus Hauffe[1]; Albrecht Stenzinger, MD[6]; Stefan Duensing, MD[1]; Heinz-Peter Schlemmer, MD[2]; Caroline M Moore, MD[4,7]; Markus Hohenfellner, MD[1] and Jan Philipp Radtke MD[1,2]

[1]Department of Urology, University Hospital Heidelberg, Heidelberg, Germany

[2]Department of Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

[3]Department of Radiology, University College London Hospital NHS Foundation Trust, London, UK

[4]Division of Surgery & Interventional Science, University College London, London, UK

[5]Vancouver Prostate Centre, University of British Columbia, Vancouver, British Columbia, Canada

[6]Institute of Pathology, University of Heidelberg, Heidelberg, Germany

[7]Department of Urology, University College London Hospitals NHS Foundation Trust, London, UK

*These authors contributed equally as first authors

Keywords: Prostate cancer, Active surveillance, MRI/TRUS-fusion-biopsy

Word count: Abstract: 287**, Manuscript:** 2891

Correspondence to:

Jan Philipp Radtke, M.D.
Department of Urology,
University of Heidelberg
Im Neuenheimer Feld 110,
69120 Heidelberg, Germany

Tel.: 0049/6221-5636321
Fax.: 0049/6221-56-5622
E-Mail: JanPhilipp.Radtke@med.uni-heidelberg.de

37

ABSTRACT

Background: Contemporary selection criteria for men with prostate cancer (PC) suitable for active surveillance (AS) are unsatisfactory, leading to high disqualification rates based on tumour misclassification. Conventional biopsy protocols are based on standard 12-core-TRUS-biopsy.

Objective: We assessed the value of MRI-/TRUS-fusion-biopsy over a 4-year follow-up in men on AS for low-risk PC.

Design, Setting and Participants: Between 2010 and 2018, a total of 273 men were included. 157 men with initial 12-core-TRUS-biopsy and 116 men with initial MRI-/TRUS-fusion-biopsy were followed by systematic and targeted transperineal MRI-/TRUS-fusion-biopsies based on PRIAS-criteria. MR-imaging from follow-up MRI/TRUS-fusion-biopsy was assessed using the PRECISE-scoring system.

Outcome measurements and statistical analysis:

AS-disqualification rates for patients on AS initially diagnosed by either 12-core TRUS biopsy or MRI/TRUS-fusion-biopsy were compared using Kaplan-Meier estimates, log-rank tests and regression analyses. We also analyzed the influence of a negative primary MRI and PRECISE-scoring to predict AS-disqualification using Kaplan-Meier estimates, log-rank tests and receiver operating characteristics (ROC) curve analysis.

Results and limitations: 59% of men diagnosed by 12-core-TRUS-biopsy were disqualified from AS based on results of subsequent MRI/TRUS-fusion-biopsy. In the initial MRI-fusion-biopsy cohort, upgrading occurred significantly less frequently (19%, $p<0.001$). ROC-curve analyses demonstrated a good discrimination for the PRECISE-score with an Area under the curve (AUC) of 0.83. No men with a PRECISE-score of 1 or 2 (demonstrating absence or downgrading of lesions in follow-up MRI) were disqualified from AS. In our cohort, a negative baseline MR scan was not a predictor of non-disqualification from AS. Limitations include transperineal approach and extended systematic biopsies used with MRI/TRUS-fusion-biopsy which may not be representative for other centres.

Conclusions: MRI/TRUS-fusion biopsies allow **a** reliable risk classification **for patients who are candidates for AS.** The application of the PRECISE scoring system demonstrated good discrimination.

Patient summary:

In this manuscript, we investigated the value of multiparametric MRI and MRI/TRUS-fusion biopsies for the assessment of AS reliability using the PRECISE criteria. Standard TRUS-biopsies lead to significant underestimation of PC. In contrast, MRI/TRUS-fusion biopsies allowed **a** more reliable risk classification. **For an appropriate inclusion into AS, men should either receive an initial or a confirmatory MRI/TRUS-fusion biopsy.**

Twitter:

Initial MRI&MRI/TRUS-fusion biopsies **allow reliable risk classification for men on** AS.

ABBREVIATIONS:

AS: Active surveillance

ASIST: Active surveillance Magnetic Resonance Imaging Study

CI: Confidence interval

DCE: Dynamic contrast enhancement

DWI: Diffusion-weighted Imaging

ESUR: European Society of Urogenital Radiology

GS: Gleason score

IQR: Interquartile range

mpMRI: Multiparametric magnetic resonance imaging

MRI: Magnetic resonance imaging

NPV: Negative predictive value

PC: Prostate cancer

PIRADS: Prostate imaging-reporting and data system

PPV: Positive predictive value

PRIAS: Prostate Cancer Research International Active Surveillance

PRECISE: Prostate Cancer Radiological Estimation of Change in Sequential Evaluation

PSA: Prostate specific antigen

RP: Radical prostatectomy

SB: Systematic transperineal saturation biopsy

STARD: Standards of Reporting of Diagnostic Accuracy

T2w: T2-weighted Imaging

TB: Targeted biopsy

TRUS: transrectal ultrasound

1. INTRODUCTION:

The introduction of prostate specific antigen (PSA)-screening for prostate cancer (PC) diagnosis has led to a substantial decrease of PC-mortality while significantly increasing detection of favorable-risk cancer [1,2]. Since half of PC diagnoses are low-risk disease, active surveillance (AS) has been recognized as an important alternative to radiation therapy or radical prostatectomy (RP) in low- and potentially intermediate-risk disease [2–4]. Current data reveal good long-term cancer specific survival and recommend AS as a safe treatment option [5]. However, despite using strict inclusion criteria, histopathological reclassification and upgrading during AS pose a great challenge in routine practice [6,7]. Drost et al. recently showed escalation of strategy to invasive treatment in over 50% of the published PRIAS men within the first years of AS [6].

Popular AS-protocols are based on a standard transrectal ultrasound (TRUS) 12-core- biopsy [8]. In the present era of magnetic resonance imaging (MRI) as a diagnostic tool for PC detection, many studies have demonstrated advantages of mpMRI as triage test reporting negative predictive values (NPV) of up to 96% [9,10]. In addition, Kasivisvanathan et al. recently **demonstrated** significantly higher detection rates of significant PC using MRI and MRI-fusion-biopsy compared to 12-core-TRUS-biopsy [11]. Similarly, Shaw et al. reported superiority of targeted biopsies (TB) over standard TRUS-biopsies [12]. Recent studies emphasize the benefit of additional MRI and MRI-guided biopsy for AS [7,13,14]. In a previously published AS-cohort with a two-year follow-up, our group demonstrated a significantly decreased AS disqualification with initial MRI/TRUS-fusion biopsy, compared to initial 12-core-TRUS-biopsy [7]. Likewise, Hamoen et al. showed additional value of initial MRI-guided-biopsy in AS [7,14]. Furthermore, the 2014 UK NICE guidelines allow the use of MRI instead of or in addition to repeat biopsies during surveillance [15]. On the contrary, the prospective, randomized Active Surveillance Magnetic Resonance Imaging Study (ASIST) showed no significant difference in the upgrading rate in men who underwent initial MRI/TRUS-fusion-biopsy compared to 12-core-TRUS-biopsy [16].

Thus, the role of MRI and MRI/TRUS-fusion-biopsy for AS is not yet clearly defined [17].

In this 4-year follow-up analysis we investigated AS-disqualification rates, according to PRIAS-criteria, in men who entered AS based on either initial 12-core-biopsy or initial MRI/TRUS-fusion-biopsy using MRI/TRUS-fusion-biopsies **as a** follow-up assessment. Additionally, we also performed sub-analyses on the impact of a negative primary MRI and standardized reporting of serial MRI studies over time using the Prostate Cancer Radiological Estimation of Change in Sequential Evaluation (PRECISE) criteria [18].

2. MATERIAL AND METHODS:

Study population

This retrospective single-center study was approved by the institutional review board (S-011/2011, S-157/2018). All subjects provided written informed consent. In compliance with the PRIAS-criteria, 273 men with low-risk PC (Gleason Score (GS) ≤ 3+3, PSA < 10 ng/ml, clinical stage T1c-cT2a), maximum two tumor positive cores and PSA-density ≤ 0.2 ng/ml were enrolled in a prospective database assessing MRI-targeted/TRUS-fusion-biopsy between 2010 and 2018 [8]. Two-year follow-up data of 149 men of the same cohort **has been** previously reported [7]. The primary biopsy **(in both groups)** was initiated **due** to elevated PSA-levels (≥4 ng/ml), suspicious PSA-kinetics and/or digital rectal examination.

Men with initial 12-core-TRUS-biopsy were diagnosed externally and referred to our department for follow-up MRI/TRUS-fusion-biopsy one year after AS-inclusion whereas initial MRI/TRUS-fusion-biopsy was performed by our department (**Supplementary Material 1**).

Imaging

Before MRI/TRUS-fusion-biopsy, MR images were acquired according to the European Society of Urogenital Radiology (ESUR) guidelines using two different 3-Tesla scanners (Magnetom Prisma and Biograph mMR, Siemens Healthcare, Erlangen, Germany) (Supplementary **Material 2**). Of all men, 18% had an external mpMRI-scan that was **considered** when MR-quality was adequate according to ESUR guidelines. Reporting was done prospectively at our department by two experienced uroradiologists (HPS and DB with >12 years of experience in prostate MRI) according to PI-RADSv 2.0. For men who received mpMRI in our department PI-RADS was prospectively assessed according to PI-RADS-Version v2.0 [19,20].

For scans before 2015, assessment was prospectively done using PI-RADS v1.0, but retrospective scoring using PI-RADS v2.0 was performed by two experienced radiologists (HPS and DB). Assessment of serial MRI-examinations was **done** according to the PRECISE recommendations (Supplementary **Material 3**) by a single uroradiologist (DB) blinded to clinical and histopathological data [18].

Biopsy protocol

All initial MRI/TRUS-fusion-biopsies and follow-up biopsies were performed transperineally using the BiopSee fusion-system (MedCom, Darmstadt, Germany) or the UroNav fusion-system (Invivo, Gainesville, Florida, United States of America) [21]. A median of 26 cores was obtained from each patient, with 24 SB and 2 TB cores [21]. The biopsy operator had access to all mpMRI-data and all biopsies were sampled under live TRUS-visualization**. TB were conducted on PI-RADS≥2 before 2016 and afterwards on PI-RADS≥3 lesions.**

Follow-up biopsies were conducted with a minor deviation from the PRIAS-protocol at one, two and 4 years after initial diagnosis of AS eligible PC, with knowledge of MRI-biopsy results and localization of PC [8].

Histopathology

Histopathological analyses were performed under supervision of two dedicated uropathologists (WR and AS with 8 and 12 years of experience, respectively) according to International Society of Urological Pathology standards and since 2014 according to the modified analyses scheme [22].

Statistical Analysis

Differences in patients' characteristics between initial biopsy approach subgroups were analyzed using Mann–Whitney-U-tests and Kolmogorov-Smirnov-tests.

Kaplan–Meier plots, log-rank tests, Cox-regression models and Hazard ratio were estimated to predict the proportion of ongoing AS for i) men with initial 12-core-TRUS-biopsy versus MRI/TRUS-fusion-biopsy and 12-core-TRUS-biopsy versus TB or SB components of initial MRI/TRUS-fusion-biopsy**, ii) men with initial 12-core-TRUS-biopsy who were not upgraded on confirmatory MRI/TRUS-fusion-biopsy versus initial MRI/TRUS-fusion-biopsy and iii)** for negative initial MRI (PI-RADS-score < 3) versus suspicious initial MRI (PI-RADS score ⩾ 3).

AS-disqualification was defined as GS ⩾ 3+4, PSA ⩾ 10 ng/ml, ⩾ 3 positive biopsy cores, PSA-density > 0.2 ng/ml or clinical stage ⩾ T2b.

Using McNemar`s tests, we analyzed differences between SB and TB on MRI/TRUS-fusion-biopsies.

The discrimination of PRECISE-scoring was assessed using Receiver operating characteristics (ROC) curve analyses [18]. Additionally, we gave descriptive data for PRECISE assessment results regarding AS-disqualification.

All statistical tests were performed two-sided with a significance level of 5%. Statistical analyses were performed using SPSS Statistics V20 (IBM, Armonk, NY, USA).

The reporting followed Standards of Reporting of Diagnostic Accuracy (Supplementary **Material 4**) [23].

3. RESULTS:

Descriptive results, including biopsy parameters of the initial biopsy in both groups, PI-RADS distribution at first MRI before MRI/TRUS-fusion-biopsy and clinical data are shown in Tables 1, 2 and Supplementary Figure 1.

The disqualification rate after 48-month follow-up was 59% (93 men) for initial standard-12-core-biopsy. In 66 men, AS-disqualification occurred after the first, in 19 men after the second and in 8 men after the third MRI/TRUS-fusion follow-up biopsy. The main reason for disqualification was histopathological upgrading. Minor upgrading to GS 3+4 tumors occurred in 60 men, whereas 17 men showed biopsy results with major histopathological upgrading to ≥4+3 tumors. 16 men were disqualified due to PSA-progression or number of positive GS 3+3 cores. In contrast, men with initial MRI/TRUS-fusion-biopsy showed a significantly lower (19%, 22 men) disqualification rate and fewer histopathological upgrading (minor upgrading in 15 men, no major upgrading). Data on disqualification, including RP specimen, if available, are given in Table 2.

Using Kaplan–Meier analyses, men initially diagnosed by MRI/TRUS-fusion-biopsy had a significantly higher eligibility for ongoing AS in 4-year follow-up (81% versus 41%; $p<0.001$)(Figure 1a). The Hazard ratio for AS-disqualification of 12-core-TRUS-biopsy was 2.56 (Confidence interval 1.70-3.85). Only the approach (MRI/TRUS-fusion versus 12-core-TRUS) was a significant predictor of AS eligibility. Results of multivariate Cox-regression analyses are provided in Table 3. The median time of 22 months on AS in the initial MRI/TRUS-fusion biopsy subgroup was significantly longer compared to 12 months in the initial 12-core-TRUS biopsy subgroup ($p=0.039$). **We provided additional information on men whose AS-eligibility based on initial 12-core TRUS-biopsy was confirmed by MRI/TRUS-fusion**

follow-up biopsy compared to men on AS based initial MRI/TRUS-fusion (Supplementary Material 5).

Differentiating between SB and TB as parts of the MRI/TRUS-fusion-biopsy versus 12-core-TRUS, both (TB or SB) were significant predictors of AS-qualification (Table 3). Data on TB and SB are given in Table 4. For all follow-up biopsies combined, upgrading was significantly higher on TB compared to SB (p=0.046).

Furthermore, we assessed the role of negative primary MRI during AS-follow-up **(Table 2, Supplementary Material 6).**

PRECISE-scores were available for 69 men after initial 12-core-biopsy and 89 men after initial MRI/TRUS-fusion-biopsy (Table 5). The probability of AS-disqualification was higher when serial-MRI had a PRECISE score of 4 or 5 (Table 5). The discrimination between absence and presence of AS-disqualification using PRECISE-score was demonstrated with a ROC of 0.83 (Figure 2). In this cohort, no patient on AS with a PRECISE-score of 1 or 2 on serial-MRI was disqualified from AS during 4-year follow-up (Table 5).

4. DISCUSSION:

To our knowledge, our study is one of the first **to compare** the outcome of men who initially underwent combined transperineal saturation and MRI/TRUS-fusion-biopsy versus initial 12-core-TRUS-biopsy in a mid-term follow-up period. All men received follow-up examinations by mpMRI and transperineal MRI/TRUS-fusion-biopsy.

MRI/TRUS-fusion-biopsies, in particular the TB component, detected upgrading to GS≥3+4 from GS=3+3 after initial 12-core-TRUS-biopsy in 59% **in contrast to recent publications where upgrading occurred in 23-44%** [14,24–26]. This variance might be explained by differences fusion-technique and more extensive SB in our cohort. **Then again,** reclassification rates **of approx.** 50% for 12-core TRUS-biopsies compared to MRI have been recently demonstrated [9].

Using Kaplan-Meier plots, we found a statistically significant lower probability of AS-disqualification in men who had initial MRI/TRUS-fusion-biopsies, not only for the previously reported 2-year follow-up, but also for longer term follow-up [7]. **The strategy of MRI-targeted and extended transperineal systematic biopsies was associated with lower rates of subsequent upgrading. Therefore, it is a more accurate strategy to rule out coexisting higher-grade disease at initial workup than in standard 12-core TRUS-biopsy.** This is contrary to the ASIST-Trial, which demonstrated no significant difference between men randomized to TRUS biopsy and those randomized to MRI-targeted biopsy using a 2 core-targeted biopsy approach. However, there were significant differences between individual centers in this study, with **two centers** showing lower rates of **upgrading** in MRI-targeted biopsy and one showing significantly higher rates, indicating that MRI/TRUS-fusion-biopsy performance varies with center specific radiologist and operator experience [16]. Additionally, in our cohort, both, initial and follow-up MRI/TRUS-fusion-biopsy on AS were performed using a higher **number** of cores for SB (24 versus 12) [16].

Performing MRI/TRUS-fusion-biopsy with a median number of 24 SB cores demands precise analyses of additional TB utility to upgrade low-risk PC as improved AS confirmation could be solely based on the difference in the number of SB cores [7]. However, on Kaplan-Meier and Cox regression analyses both SB and TB of MRI/TRUS-fusion biopsies were significant predictors of AS-confirmation. Moreover, using McNemar`s tests, we could demonstrate that reclassification on follow-up biopsy was significantly higher on TB compared to SB (p=0.046). This is in the line with Frye et al., demonstrating superiority of TB, but contrary to a recent publication **stating** that SB was superior to TB for AS-disqualification on one-year follow-up biopsy [14,27]. In our cohort, accurate risk stratification and AS-confirmation on follow-up is mainly associated with TB. Since 6% (55/85) upgrades were not detected by TB alone, SB should not yet be omitted [14,17,28]. We acknowledge that we did not analyze differences in SB approaches (12-core-TRUS-biopsy versus extended 24-core transperineal scheme). However, Pham et al. demonstrated that an extended 24-core scheme does not affect the disqualification rate significantly and it is debatable whether 12-core-TRUS-biopsy remains the standard of care in the light of the results of PROMIS and PRECISION [9,11].

We further evaluated the ability of PI-RADS- and PRECISE-scoring to appropriately predict PC and AS-disqualification. Focusing on PI-RADS, 99 men had a primarily negative MRI (PI-RADS 1 or 2) defined as negative MRI before AS inclusion in the MRI/TRUS-fusion-biopsy subgroup or as negative MRI at one-year follow-up for the initial 12-core-TRUS-biopsy subgroup. Out of those, 25 men (27%) had AS-disqualification during 4-year follow-up. Using Kaplan-Meier plots we could demonstrate, analogous to Olivier et al., that there was no statistical significant difference between men who had a negative primary MRI compared to a suspicious primary MRI (PI-RADS 3-5) [29]. However, the overall disqualification rate of 27% for men with primarily negative MRI was higher compared to Olivier et al. (12%) [29]. Differences might be explained by the longer follow-up in our cohort [29].

One strength of this study is the application of the PRECISE recommendations in an AS cohort [18]. Using PRECISE scoring the AUC in ROC curve analysis was 0.83, illustrating a high diagnostic **competence**. When only PRECISE-scores of 1-2 are taken into account, the NPV of a resolved or reduced lesion on MRI (PRECISE-score 1 or 2) for GS-upgrading was 100% in our cohort. This was comparable to recent publications, reporting NPVs of 96-100% for a pristine MRI [3,30]. However all used a different index test, performing standard 12-core-TRUS-biopsies [3,30]. In contrast, transperineal saturation biopsy as an index test may lead to decreased NPV of mpMRI for GS ≥3+4 PC, which is also demonstrated by a higher disqualification rate in our entire cohort, but results in improved patient selection, as mpMRI still misses approximately 10% of index lesions compared to final RP specimen [21].

The question arises if AS-inclusion based on the initial MRI/TRUS-fusion biopsy and an unsuspicious follow-up MRI may allow to **abandon** follow-up biopsy. **In case of a PRECISE-score** of 1-2 the risk of upgrading was 0% in our cohort. This is similar to Panebianco et al., who followed negative MRI for four years and demonstrated a NPV of 96% [10]. Thus, using MRI and PRECISE might **allow** to **monitor** AS patients without rebiopsy or **performing a rebiopsy** only **in** men with a PRECISE-**score** ≥3 on follow-up. However, the number of patients to corroborate these findings is small and longer follow-up is needed.

Our study has some limitations. We did not perform multi-rater reading of mpMRI and PRECISE-scoring which may influence detection sensitivity. However, results of NPV for MRI in our selected cohort are comparable to those reported in recent AS and large-scale publications [11,25]. Representing real-life clinical routine, radiologists and operators were not blinded to clinical data and MRI-results.

The **requirements** for performing mpMRI, transperineal fusion-biopsy and extended SB are certainly higher than those for standard biopsies. **We acknowledge that a comparison of initial 12-core TRUS-biopsy and MRI/TRUS-fusion biopsy with transperineal extended SB may cause some bias. However, at the beginning of data accrual in 2010, most AS protocols, including the PRIAS protocol (and the UCSF and Johns Hopkins protocols) included men to AS based on a standard 12-core TRUS [8,31]. At our institution we chose a different biopsy approach since 2010. Besides the fact that this comparison is not optimal in terms of a randomized prospective trial, it represents a real-life scenario. However, what was done was a test of surveillance protocol standard (12-core TRUS-biopsy) versus MRI/TRUS-fusion biopsy. More important and besides the comparison between initial MRI/TRUS-fusion and initial 12-core TRUS-biopsy, our data support that initial MRI and MRI/TRUS-fusion are reliable tools to stratify men to AS with good mid-term eligibility. To account for men who are not eligible for AS but seem to have low-risk PC based on 12-core TRUS, a confirmatory biopsy is mandatory. Based on our data, a MRI/TRUS-fusion-biopsy is an appropriate approach to do so . This is in accordance with recent updated PRIAS-criteria (https://www.prias-project.org).**

Additionally, the cost effectiveness of mpMRI and TB has been demonstrated [32]. Although our follow-up has grown to 48 months, a longer follow-up, like in the Toronto and PRIAS data, is lacking. We used only the PRIAS-protocol, while other protocols for the assessment of long-term probability of AS are available [5]. However, the performance of the PRIAS-protocol has recently been shown to be comparable with other protocols [8,33]. We acknowledge that contrary to original PRIAS-protocol we performed a second MRI/TRUS-fusion follow-up biopsy after two years.

The endpoint of our analysis was AS-eligibility and not the more sophisticated disease-specific mortality rate in long-term follow-up **which** is associated with excellent results [5]. However, the endpoint of disease-specific mortality may be debatable in a cohort of low-risk PC patients exhibiting long-term disease-specific survival of at least 90%. Furthermore the number of PC-invaded cores (≤2, according to the PRIAS-protocol) was not adapted to our biopsy approach with median 26 cores [8]. **This might have implications on the tumor burden. However, the percentage rate on initial biopsy was comparable for both groups (Supplementary Material 5), core involvement caused AS-disqualification in very few cases in our cohort and the predictive value of absolute numbers of PC-positive cores is debatable [24].**

Next, performing transperineal extended SB in all patients has implications for both cost and resource utilization and is a technique that may not be widely available outside of selected academic centers. Additionally, the approach using transperineal SB does not provide information if the same results for SB could also be achieved using standard 12-core TRUS as SB.

Lastly, it is **controversial** if upgrading to GS=3+4 with minimal pattern 4 disease should automatically disqualify from AS [4,34].

5. CONCLUSION:

Initial MRI/TRUS-fusion biopsy leads to **long** AS-eligibility and **minor** upgrading in the 4-year AS period. **After initial 12-core TRUS-biopsy, confirmatory (MRI/TRUS fusion) biopsy is necessary for reliable AS-inclusion.** For detection of GS-upgrading, both TB and SB are necessary. Discrimination of the PRECISE-score for AS-disqualification was good. Using PRECISE, it might be **possible** to follow-up AS patients without rebiopsy or **only performing a rebiopsy in** men with a PRECISE ≥3 on follow-up.

6. CONFLICTS OF INTEREST:

None declared.

Francesco Giganti is funded by the University College London (UCL) Graduate Research Scholarship and the Brahm PhD scholarship in memory of Chris Adams.

7. REFERENCES:

[1] Cooperberg MR, Carroll PR. Trends in Management for Patients With Localized Prostate Cancer, 1990-2013. JAMA 2015;314:80. doi:10.1001/jama.2015.6036.

[2] Wong MCS, Goggins WB, Wang HHX, Fung FDH, Leung C, Wong SYS, et al. Global Incidence and Mortality for Prostate Cancer: Analysis of Temporal Patterns and Trends in 36 Countries. Eur Urol 2016;70:862–74. doi:10.1016/j.eururo.2016.05.043.

[3] Da Rosa MR, Milot L, Sugar L, Vesprini D, Chung H, Loblaw A, et al. A prospective comparison of MRI-US fused targeted biopsy versus systematic ultrasound-guided biopsy for detecting clinically significant prostate cancer in patients on active surveillance. J Magn Reson Imaging 2015;41:220–5. doi:10.1002/jmri.24710.

[4] Ploussard G, Isbarn H, Briganti A, Sooriakumaran P, Surcel CI, Salomon L, et al. Can we expand active surveillance criteria to include biopsy Gleason 3+4 prostate cancer? A multi-institutional study of 2,323 patients. Urol Oncol Semin Orig Investig 2015;33:71.e1-71.e9. doi:10.1016/j.urolonc.2014.07.007.

[5] Klotz L, Vesprini D, Sethukavalan P, Jethava V, Zhang L, Jain S, et al. Long-Term Follow-Up of a Large Active Surveillance Cohort of Patients With Prostate Cancer. J Clin Oncol 2014;10.1200/JC. doi:10.1200/JCO.2014.55.1192.

[6] Drost F-JH, Rannikko A, Valdagni R, Pickles T, Kakehi Y, Remmers S, et al. Can active surveillance really reduce the harms of overdiagnosing prostate cancer? A reflection of real life clinical practice in the PRIAS study. Transl Androl Urol 2018;7:98–105. doi:10.21037/tau.2017.12.28.

[7] Radtke JP, Kuru TH, Bonekamp D, Freitag MT, Wolf MB, Alt CD, et al. Further reduction of disqualification rates by additional MRI-targeted biopsy with transperineal saturation biopsy compared with standard 12-core systematic biopsies for the selection of prostate cancer patients for active surveillance. Prostate Cancer Prostatic Dis 2016;19. doi:10.1038/pcan.2016.16.

[8] van den Bergh RCN, Roemeling S, Roobol MJ, Roobol W, Schröder FH, Bangma CH. Prospective Validation of Active Surveillance in Prostate Cancer: The PRIAS Study.

Eur Urol 2007;52:1560–3. doi:10.1016/j.eururo.2007.05.011.

[9] 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. Lancet 2017;389:815–22. doi:10.1016/S0140-6736(16)32401-1.

[10] Panebianco V, Barchetti G, Simone G, Del M, Ciardi A, Domenico M, et al. Negative Multiparametric Magnetic Resonance Imaging for Prostate Cancer : What ' s Next ? Eur Urol 2018. doi:10.1016/j.eururo.2018.03.007.

[11] Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. N Engl J Med 2018:NEJMoa1801993. doi:10.1056/NEJMoa1801993.

[12] Shaw GL, Thomas BC, Dawson SN, Srivastava G, Vowler SL, Gnanapragasam VJ, et al. Identification of pathologically insignificant prostate cancer is not accurate in unscreened men. Br Cournal Cancer 2014;110:doi: 10.1038/bjc.2014.192. doi:10.1038/bjc.2014.192.

[13] Schoots IG, Osses DF, Drost F-JH, Verbeek JFM, Remmers S, van Leenders GJLH, et al. Reduction of MRI-targeted biopsies in men with low-risk prostate cancer on active surveillance by stratifying to PI-RADS and PSA-density, with different thresholds for significant disease. Transl Androl Urol 2018;7:132–44. doi:10.21037/tau.2017.12.29.

[14] Hamoen EHJ, Hoeks CMA, Somford DM, van Oort IM, Vergunst H, Oddens JR, et al. Value of Serial Multiparametric Magnetic Resonance Imaging and Magnetic Resonance Imaging-guided Biopsies in Men with Low-risk Prostate Cancer on Active Surveillance After 1 Yr Follow-up. Eur Urol Focus 2018:1–9. doi:10.1016/j.euf.2017.12.008.

[15] Bryant R, Yang B, Philippou Y, Lam K, Obiakor M, Ayers J, et al. Does the introduction of prostate multiparametric MRI into active surveillance of localised prostate cancer improve patient re-classification? BJU Int 2018:0–2. doi:10.1111/bju.14248.

[16] Klotz L, Loblaw A, Sugar L, Moussa M, Berman DM, Van der Kwast T, et al. Active Surveillance Magnetic Resonance Imaging Study (ASIST): Results of a Randomized Multicenter Prospective Trial. Eur Urol 2018. doi:10.1016/j.eururo.2018.06.025.

[17] Schoots IG, Nieboer D, Giganti F, Moore CM, Bangma CH, Roobol MJ. Is magnetic resonance imaging-targeted biopsy a useful addition to systematic confirmatory biopsy in men on active surveillance for low-risk prostate cancer? A systematic review and meta-analysis. BJU Int 2018. doi:10.1111/bju.14358.

[18] Moore CM, Giganti F, Albertsen P, Allen C, Bangma C, Briganti A, et al. Reporting Magnetic Resonance Imaging in Men on Active Surveillance for Prostate Cancer : The PRECISE Recommendations — A Report of a European School of Oncology Task Force Associate Editor : 2017;71:648–55.

[19] Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. Eur Radiol 2012;22:746–57. doi:10.1007/s00330-011-2377-y.

[20] Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider M, Macura KJ, et al. PI-RADS Prostate Imaging and Reporting and Data System: 2015, Version 2. Eur Urol 2016;69:16–40. doi:10.1016/j.eururo.2015.08.052.

[21] Radtke JP, Schwab C, Wolf MB, Freitag MT, Alt CD, Kesch C, et al. Multiparametric Magnetic Resonance Imaging (MRI) and MRI?Transrectal Ultrasound Fusion Biopsy for Index Tumor Detection: Correlation with Radical Prostatectomy Specimen. Eur Urol 2016;70. doi:10.1016/j.eururo.2015.12.052.

[22] Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma definition of grading patterns and proposal for a new grading system. Am J Surg Pathol 2016;40:244–52. doi:10.1097/PAS.0000000000000530.

[23] Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis C a, Glasziou PP, Irwig LM, et al. Towards complete and,accurate reporting of studies of diagnostic accuracy: the

STARD initiative. Radiology 2003;226:24–8. doi:10.1136/bmj.326.7379.41.

[24] Pham KN, Porter CR, Odem-Davis K, Wolff EM, Jeldres C, Wei JT, et al. Transperineal Template Guided Prostate Biopsy Selects Candidates for Active Surveillance—How Many Cores are Enough? J Urol 2015;194:674–9. doi:10.1016/j.juro.2015.04.109.

[25] Hu JC, Chang E, Natarajan S, Margolis DJ, Macairan M, Lieu P, et al. Targeted prostate biopsy in select men for active surveillance: do the Epstein criteria still apply? J Urol 2014;192:385–90. doi:10.1016/j.juro.2014.02.005.

[26] Bloom JB, Hale G, Gold SA, Rayn K, Smith C, Mehralivand S, et al. Predicting Gleason Group Progression for Men on Prostate Cancer Active Surveillance: The Role of a Negative Confirmatory MRI-US Fusion Biopsy. J Urol 2018. doi:10.1016/j.juro.2018.07.051.

[27] Frye TP, George AK, Kilchevsky A, Maruf M, Siddiqui MM, Kongnyuy M, et al. Magnetic Resonance Imaging-Transrectal Ultrasound Guided Fusion Biopsy to Detect Progression in Patients with Existing Lesions on Active Surveillance for Low and Intermediate Risk Prostate Cancer. J Urol 2017;197:640–6. doi:10.1016/j.juro.2016.08.109.

[28] Martin Ma T, Tosoian JJ, Schaeffer EM, If TD, Landis P, Wolf S, et al. The Role of Multiparametric Magnetic Resonance Imaging / Ultrasound Fusion Biopsy in Active Surveillance. Eur Urol 2017;71:174–80. doi:10.1016/j.eururo.2016.05.021.

[29] Olivier J, Kasivisvanathan V, Drumez E, Fantoni J-C, Leroy X, Puech P, et al. Low-risk prostate cancer selected for active surveillance with negative MRI at entry: can repeat biopsies at 1 year be avoided? A pilot study. World J Urol 2018. doi:10.1007/s00345-018-2420-6.

[30] Vargas HA, Akin O, Afaq A, Goldman D, Zheng J, Moskowitz CS, et al. Magnetic resonance imaging for predicting prostate biopsy findings in patients considered for active surveillance of clinically low risk prostate cancer. J Urol 2012;188:1732–8. doi:10.1016/j.juro.2012.07.024.

[31] Tosoian JJ, Trock BJ, Landis P, Feng Z, Epstein JI, Partin AW, et al. Active surveillance program for prostate cancer: An update of the Johns Hopkins experience. J Clin Oncol 2011;29:2185–90. doi:10.1590/S1677-55382011000200021.

[32] De Rooij M, Crienen S, Witjes JA, Barentsz JO, Rovers MM, Grutters JPC. Cost-effectiveness of Magnetic Resonance (MR) imaging and MR-guided targeted biopsy versus systematic transrectal ultrasound-guided biopsy in diagnosing prostate cancer: A modelling study from a health care perspective. Eur Urol 2014;66:430–6. doi:10.1016/j.eururo.2013.12.012.

[33] Carter HB, Kettermann A, Warlick C, Metter EJ, Landis P, Walsh PC, et al. Expectant Management of Prostate Cancer With Curative Intent: An Update of The Johns Hopkins Experience. J Urol 2007;178:2359–65. doi:10.1016/j.juro.2007.08.039.

[34] Nassiri N, Margolis DJ, Natarajan S, Sharma DS, Huang J, Dorey FJ, et al. Targeted Biopsy to Detect Gleason Score Upgrading during Active Surveillance for Men with Low versus Intermediate Risk Prostate Cancer. J Urol 2017;197:632–9. doi:10.1016/j.juro.2016.09.070.

Table legends:

Table 1:

Patients' demographics including baseline clinical parameters, initial biopsy results for both subgroups, MRI results of the initial MRI/TRUS-fusion biopsy subgroup and Mann-Whitney-U-test/Kolmogorov-Smirnov-test for differences between subgroups

Table 2:

Clinical parameters, biopsy results, histopathology of upgrading and PRECISE-score information for patients who were disqualified from AS, stratified into initial biopsy approach subgroups

Table 3:

Multivariate cox regression analysis for: (A) analysis of initial 12-core TRUS-biopsy versus MRI/TRUS-fusion biopsy to detect upgrading and (B) analysis of initial 12-core TRUS-biopsy versus separate TB and SB component of MRI/TRUS-fusion biopsy

Table 4:

Results of targeted versus systematic biopsies in MRI/TRUS-fusion biopsy

Table 5:

Detection accuracy of PRECISE-score to detect AS-disqualification

Figure legends:

Figure 1:

a) Kaplan-Meier plot for the probability of ongoing active surveillance for men with initial standard biopsy versus initial MRI/TRUS-fusion biopsy for 4-years follow-up, according to PRIAS-criteria and b) Kaplan-Meier plot for the probability of ongoing active surveillance for men with initial standard biopsy versus the TB or the SB component of the initial MRI/TRUS-fusion biopsy for 4-year follow-up, according to PRIAS-criteria

Figure 2:

ROC curve analysis for PRECISE-score in the entire cohort

Supplementary Material:

<u>Supplementary Material 1:</u>

Study flowchart with inclusion and exclusion criteria

<u>Supplementary Material 2:</u>

Sequence parameters of multiparametric MRI protocol

<u>Supplementary Material 3:</u>

Assessment of likelihood for radiologic progression on magnetic resonance imaging in men on active surveillance according to PRECISE publication

<u>Supplementary Material 4:</u>

Standards of Reporting of Diagnostic Accuracy (STARD) checklist

<u>Supplementary Material 5:</u>

Kaplan-Meier plot for the probability of ongoing active surveillance for men with initial standard biopsy versus initial MRI/TRUS-fusion biopsy for 4-years follow-up, according to PRIAS-criteria. Men who underwent initial 12-core TRUS-biopsy and suffered from Gleason score upgrading in the MRI/TRUS-fusion biopsy after one year were not included into analysis.

Supplementary Material 6:

Kaplan-Meier plot for the disqualification from active surveillance for men with initial negative MRI versus suspicious initial MRI

Take Home Message:

Compared to standard TRUS-biopsy, initial MRI/TRUS-fusion-biopsy leads to longer AS-eligibility and decreased upgrading during 4-year AS follow-up. PRECISE-scoring for AS-disqualification demonstrated high diagnostic ability on external validation. Using PRECISE, it might be possible to abstain from rebiopsies during AS follow-up.

Table 1

Parameter	Initial 12-core TRUS biopsy subgroup with biopsy-proven low-risk prostate cancer (n=157)	Initial MRI/TRUS-fusion biopsy (n=116)	Mann-Whitney-U-test, Kolmogorov-Smirnov-test *p*-value
Median Age (IQR), years	69 (64-75)	69 (64-74)	0.94
Median PSA level (IQR), ng/ml	6.2 (4.7-7.7)	5.8 (4.5-7.0)	0.09
Men with MRI-suspicious lesions (PI-RADS Likert ≥2), no.		92	
Men without MRI-suspicious lesions		24	
Men with ≥2 suspicious lesions		41	
Maximum PI-RADS Likert score on mpMRI per patient			
2		15	
3		35	
4		30	
5		12	
Median No. of cores in initial biopsy (IQR)	12 (10-12)	26 (24-28)	**<0.001**
Median No. of systematic cores (IQR)	12 (10-12)	24 (24-24)	**<0.001**
Median No. of targeted cores (IQR)		2 (1-4)	
Median No. of core involvement in initial biopsy (IQR)	1 (1-2)	2 (1-3)	0.87
Median interval from first diagnosis to presentation (IQR), months	12 (9-15)		
No. of patients with positive digital rectal examination (DRE)(cT2a)	22	4	**<0.001**
Median prostate volume (IQR), in ml	41 (31-54)	43.0 (31.0-57.0)	0.97
Median PSA density level (IQR), in ml	0.15 (0.09-0.22)	0.15 (0.10-0.20)	0.24

IQR- Interquartile range, mpMRI- multiparametric Magnetic resonance imaging, DRE- digital rectal examination

PI-RADS- Prostate imaging reporting and data system, PC- prostate cancer, TRUS- transrectal ultrasound, PSA- prostate specific antigen

Table 2

Men with disqualification of active surveillance dichotomized according to initial biopsy method	**Initial diagnosis by 12-core TRUS biopsy**	**Initial diagnosis by MRI/TRUS-fusion biopsy**	***Mann-Whitney-U-test, Kolmogorov-Smirnov-test, p-value***
No. of patients (% of all diagnosed patients in subgroup)	93 (59)	22 (19)	**<0.001**
Median PSA level at disqualification in ng/ml (IQR)	6.2 (4.7-7.7)	5.8 (4.7-7.0)	0.633
Biopsy results in follow-up biopsies			
Median No.of cores (IQR)	26 (24-28)	26 (24-28)	0.995
Median No.of systematic cores (IQR)	24 (22-24)	24 (24-24)	0.889
Median No.of targeted cores (IQR)	2(2-4)	2 (1-5)	0.614
Median No. of involved cores per patient (IQR), % of positive cores	2 (1-4), 17	2 (1-3), 8	**0.001**
Type of reclassification/disqualification			
Pathologic progression, Grade only (GS≥3+4)	77	15	**<0.001**
Grade-related progression GS=3+4	60	15	**<0.001**
Grade-related progression GS=4+3	11	0	**0.018**
Grade-related progression GS≥4+4	6	0	**0.025**
Disqualification by number of cores (≥ 3 cores)	9	1	**0.038**
Disqualification by PSA, PSA-Progression	7	6	0.822
Median time to disqualification in months (IQR)	12 (9-16)	22 (20-25)	**0.039**
Disqualification in first follow-up MRI/TRUS-fusion biopsy	66	9	**<0.001**
Disqualification in second follow-up MRI/TRUS-fusion biopsy	19	11	**0.043**
Negative initial MRI			
No. of patients with negative initial MRI	15	10	0.427
PRECISE distribution of disqualified patients on confirmation mpMRI			
No. of disqualified patients eligible for PRECISE (serial MRI available)	7	22	
PRECISE Score 1	0	0	1.000
PRECISE Score 2	0	0	1.000
PRECISE Score 3	3	9	**<0.001**
PRECISE Score 4	2	13	**<0.001**
PRECISE Score 5	2	0	**<0.001**
Results of radical prostatectomy after AS disqualification (n=62)	n=49	n=13	
Gleason Score			**0.008**
3+3	2	2	
3+4	33	11	
4+3	7	0	
4+5	7	0	
T-Stage			**0.043**
pT2	36	11	
pT3a	10	2	
pT3b	3	0	
N-Status			0.299
N0	47	13	
N1	2	0	

GS- Gleason Score, IQR- Interquartile range, mpMRI- multiparametric Magnetic resonance imaging, SB- systematic biopsies, TB- targeted biopsies, PI-RADS- Prostate imaging reporting and data system, PC- prostate cancer, TRUS- transrectal ultrasound, PSA- prostate specific antigen

Table 3

A Cox regression for initial standard biopsy versus initial MRI/TRUS-fusion biopsy			
Characteristics	**Upgrading from Active surveillance**		
	Hazard ratio	**(95%-CI)**	***p* value**
Initial MRI/TRUS-fusion biopsy versus initial 12-core TRUS	0.25	0.10;0.61	**0.002**
No. of cores in intial biopsy	0.99	0.94;1.05	0.83
Age	1.02	0.99;1.04	0.09
PSA	0.97	0.90;1.04	0.33
PSA density	3.00	0.11;29.40	0.34
B Cox regression for initial standard biopsy versus initial TB or SB component of MRI/TRUS-fusion biopsy			
Characteristics	**Upgrading from Active surveillance**		
	Hazard ratio	**(95%-CI)**	***p* value**
Active surveillance confirmed by initial TB	0.35	0.26;0.84	**0.011**
Active surveillance confirmed by initial SB	0.75	0.63;0.96	**0.047**
Active surveillance confirmed by initial 12-core biopsy	1.82	0.83;4.67	0.31
No. of cores in initial biopsy	0.99	0.95;1.04	0.70
Age	1.01	0.99;1.04	0.62
PSA	0.97	0.81;1.04	0.39
PSA density	1.63	0.18;14.91	0.67

CI- confidence interval, MRI- Magnetic resonance imaging, PC- prostate cancer, PSA- prostate specific antigen, TRUS- transrectal ultrasound

Table 4

Results of targeted versus systematic biopsies in MRI/TRUS-fusion biopsy

Initial MRI/TRUS-fusion biopsy (n=116) prior to AS inclusion

		Detected by MRI-targeted biopsies			
		No PCa	GS = 3+3	GS ≥ 3+4	Sum
Detected	No PCa	0	13	0	13
by	GS = 3+3	44	59	0	103
Systematic biopsies	GS ≥ 3+4	0	0	0	0
	Sum	44	72	0	116

Mc Nemar's test for systematic versus targeted biopsy $p<0.001$

All follow-up biopsies

		Detected by MRI-targeted biopsies			
		No PCa	GS = 3+3	GS ≥ 3+4	Sum
Detected	No PCa	79	12	12	103
by	GS = 3+3	74	100	6	180
Systematic biopsies	GS ≥ 3+4	4	1	80	85
	Sum	157	113	98	368

Mc Nemar`s test p-value for systematic versus MRI-targeted biopsy $p=0.046$

Table 5

Detection accuracy of PRECISE score to detect AS-disqualification Entire cohort	
PRECISE Score	69 men in the initial 12-core TRUS and 89 men in the MRI/TRUS-fusion subgroup

	AS disqualification			
PRECISE Score		no	yes	Sum
	1	14	0	14
	2	43	0	43
	3	59	12	71
	4	13	15	28
	5	0	2	2
	Sum	99	29	158

Figure 1

Click here to download high resolution image

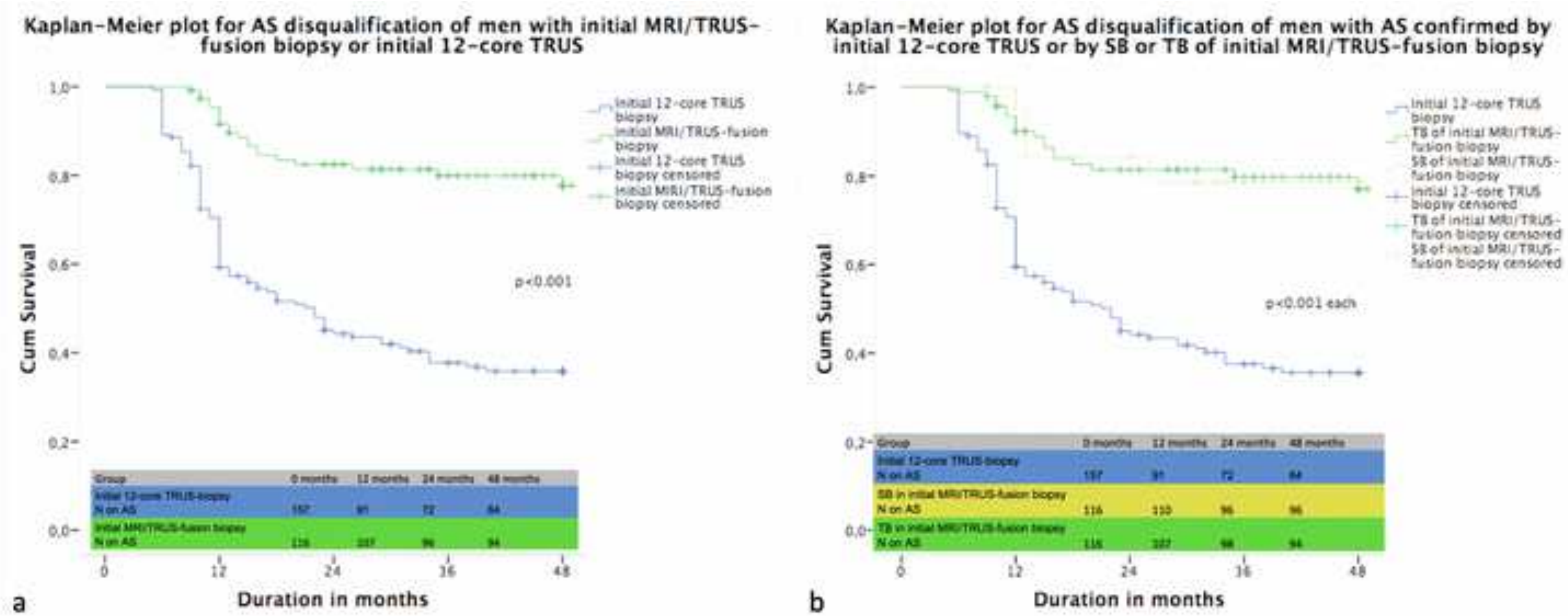
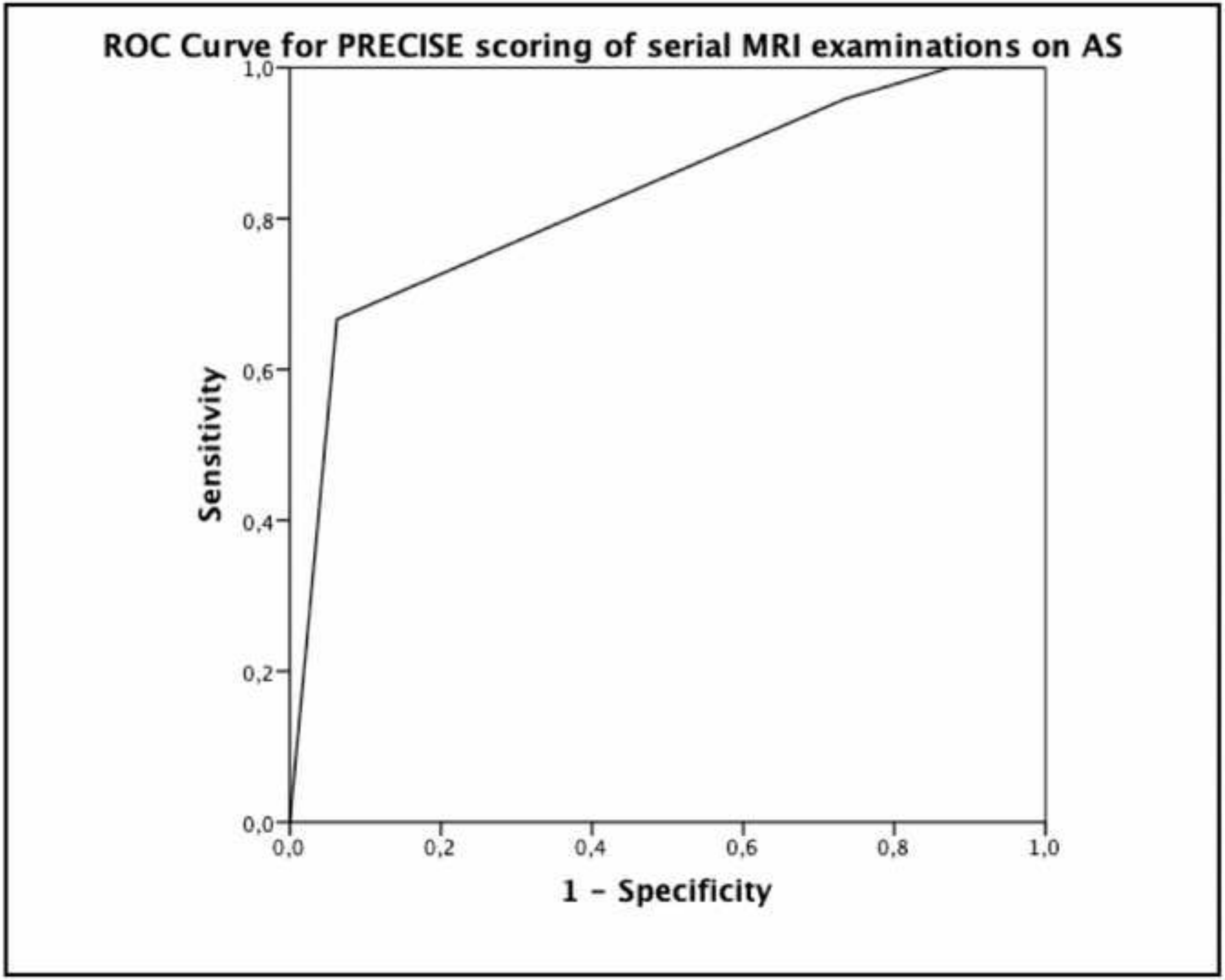


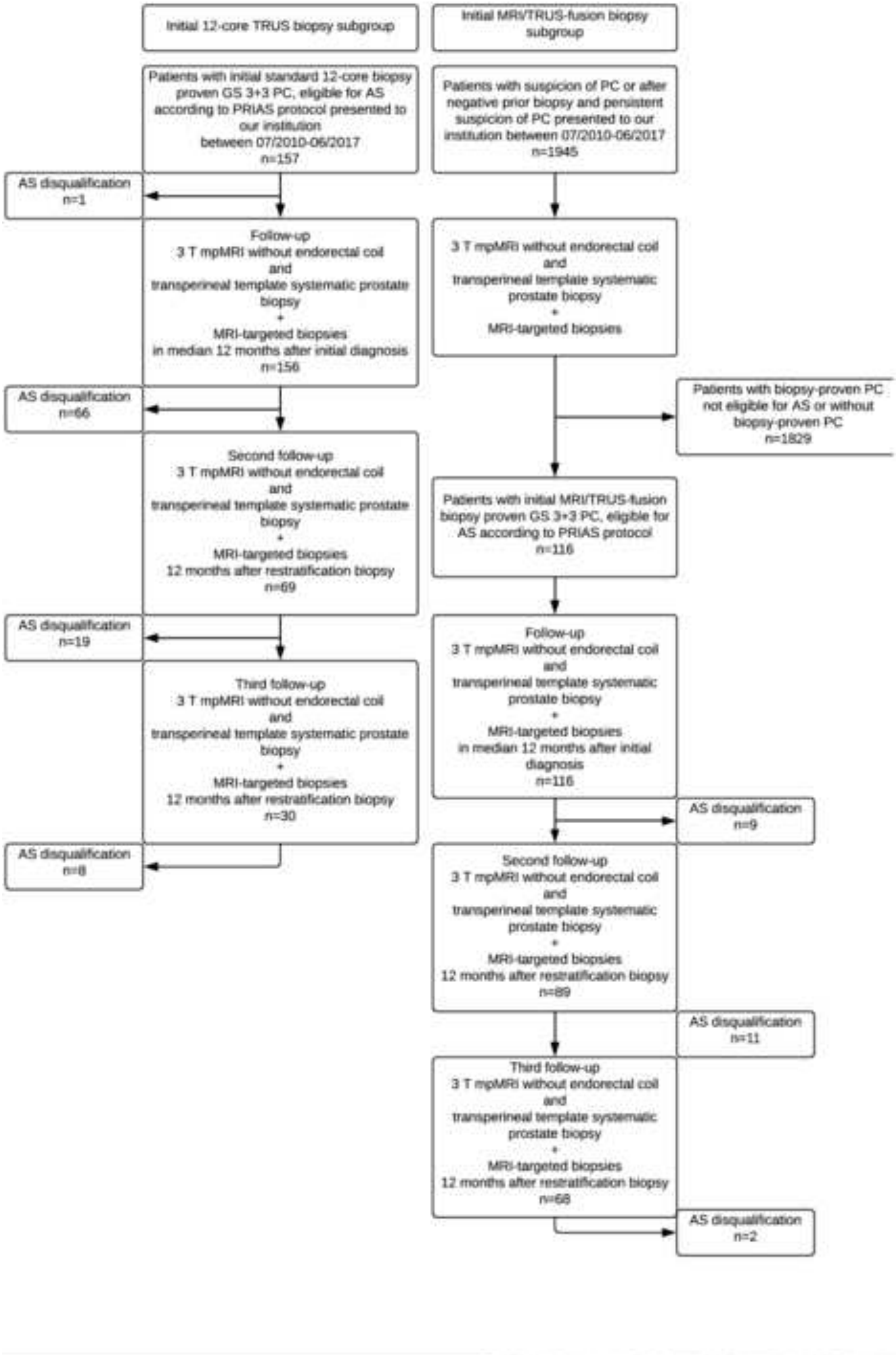
Figure 2

Click here to download high resolution image



Supplementary Material 1

Click here to download high resolution image



Supplementary Material 2

Parameter	T1 TSE	T2 TSE	epi-2D	DCE TWIST
TR ms/ TE ms	792/11	5120/143	3100/52	4.42/2.2
Flip angle (°)	90	90	90	15
ETL length/ Epi- factor	72	12	96	x
Averages	2	4	5	x
b-value	x	x	0, 50, 100, 150, 200, 250, 800, 1000	x
Section thickness (mm)	5	3	3	1,5
FOV (mm)	320	300	280	400
Resolution	1.1 x 1.0	0.8 x 0.7	2.2 x 2.2	1.6 x 1.6
Acquisition time (min)	03:51	04:14	05:04	05:18

TR- Repetition Time, TE- Echo Time, ETL- Echo Train Length, FOV- Field of View, epi- Echo Planar Imaging, TSE- Turbo Spin Echo, TWIST- Time-resolved angiography With Interleaved Stochastic Trajectories, SE- Spin Echo, DCE- Dynamic contrast enhancement

Supplementary Material 3

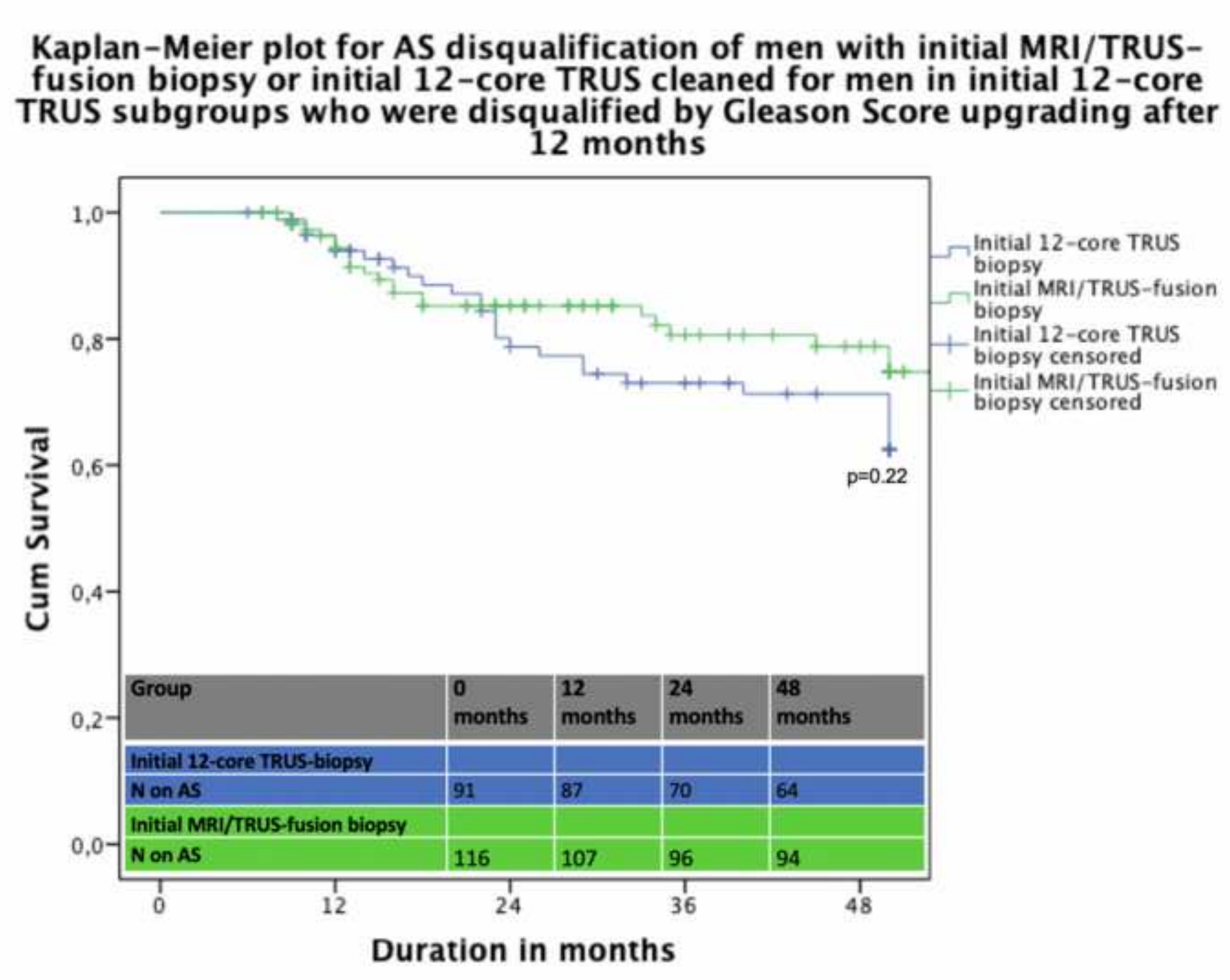
Assessment of likelihood for radiologic progression on magnetic resonance imaging in men on active surveillance		
Likert	Assessment of likelihood of radiologic progression	Example
1	Resolution of previous features suspicious on MRI	Previously enhancing area no longer enhances
2	Reduction in volume and/or conspicuity of previous features suspicious on MRI	Reduction in size of previously seen lesion that remains suspicious for clinically significant disease
3	Stable MRI appearance: no new focal/diffuse lesions	Either no suspicious features or all lesions stable in size and appearance
4	Significant increase in size and/or conspicuity of features suspicious for prostate cancer	Lesion becomes visible on diffusion-weighted imaging; significant increase in size of previously seen lesion
5	Definitive radiologic stage progression	Appearance of extracapsular extension, seminal vesicle involvement, lymph node involvement, or bone metastasis

STARD checklist for reporting of studies of diagnostic accuracy

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	1
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	6, 7
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	8, Figure 1
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	8, Figure 1
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	8, Figure 1
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	8, 9, 10
Test methods	7	The reference standard and its rationale.	8, 9
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	8, 9
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	8, 9
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	8, 9
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	8, 9
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	9, 10
	13	Methods for calculating test reproducibility, if done.	N/A
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	11
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	11, Table 1
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	11, Table 1, Figure 1
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	11, 12, Table 1, Table 2
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	11
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Tables 1,2 Figures 1,2 Supplementary Material 5
	20	Any adverse events from performing the index tests or the reference standard.	Figure 1
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	Tables 3, 4, 5
	22	How indeterminate results, missing data and outliers of the index tests were handled.	Figure 1
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	Table 4
	24	Estimates of test reproducibility, if done.	N/A
DISCUSSION	25	Discuss the clinical applicability of the study findings.	13-17, 18

Supplementary Material 5

Click here to download high resolution image



Supplementary Material 6

Click here to download high resolution image