Factors Influencing Variability in the Performance of mp-MRI in Detecting Clinically Significant Prostate Cancer: A Systematic Literature Review

Armando Stabile[1], Francesco Giganti[2,3], Veeru Kasivisvanathan[2,4], Gianluca Giannarini[5], Caroline M Moore[2,4], Anwar R Padhani[6], Valeria Panebianco[7], Andrew B Rosenkrantz[8], Georg Salomon[9], Baris Turkbey[10], Geert Villeirs[11], Jelle O Barentsz[12]

[1] Department of Urology and Division of Experimental Oncology, URI, Urological Research Institute, Vita-Salute San Raffaele University, IRCCS San Raffaele Scientific Institute, Milan, Italy

[2] Division of Surgery and Interventional Science, University College London, London, UK

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

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

[5] Urology Unit, Academic Medical Centre, Santa Maria della Misericordia Hospital, Udine, Italy.

[6] Paul Strickland Scanner Centre, Mount Vernon Cancer Centre, Northwood, UK.

[7] Department of Radiology, Sapienza Rome University, Policlinico Umberto I, Rome, Italy.

[8] Department of Radiology, NYU Langone Health, New York, NY, USA.

[9] Prostate Cancer Center, Martini-Klinik Hamburg, University Hospital Hamburg-Eppendorf, Hamburg, Germany.

[10] Molecular Imaging Program, National Cancer Institute, NIH, Bethesda, MD, USA

[11] Department of Radiology, Ghent University Hospital, Ghent, Belgium.

[12] Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands.

Corresponding author:
Armando Stabile, MD
Department of Urology and Division of Experimental Oncology
Urological Research Institute, IRCCS San Raffaele Scientific Institute,
Via Olgettina 60, 20132 Milan, Italy
Tel. +39 02 26435663; Fax. +39 02 26437298
E-mail: armando.stabile88@gmail.com

Abstract

Context: There is a lack of comprehensive data regarding the factors that influence the diagnostic accuracy of multiparametric magnetic resonance imaging (mpMRI) to detect and localize clinically significant prostate cancer (csPCa).

Objective: To systematically review the current literature assessing the factors influencing the variability of mpMRI performance in csPCa diagnosis.

Evidence acquisition: A computerized bibliographic search of Medline/PubMed database was performed for all studies assessing magnetic field strength, the use of an endorectal coil (ERC), the assessment system used by radiologists and inter-reader variability, radiologist and urologist experience, the use of a contrast agent and the use of computer-aided diagnosis (CAD) tools in relation to mpMRI diagnostic accuracy.

Evidence synthesis: A total of 77 articles were included. Both radiologist's reading experience and urologist's/radiologist's biopsy experience were the main factors that influenced diagnostic accuracy. Therefore, it is mandatory to indicate the experience of the interpreting radiologists and biopsy-performing urologists to support the reliability of the findings. The most recent PI-RADS guidelines are recommended to use, as the main assessment system for csPCa given the simplified and standardized approach as well as its particular added value for less experienced radiologists. Biparametric MRI had similar accuracy to mpMRI, however, bpMRI performed better with experienced readers. The limited data available suggests that the combination of CAD and radiologist readings may positively influence diagnostic accuracy.

Conclusions: Multiple factors affect the accuracy of mpMRI and MRI-targeted biopsy to detect and localize csPCa. The high heterogeneity across the studies underlines the need

24 to define the experience of radiologist and urologist, to implement quality control, and to

25 adhere to the most recent PI-RADS assessment guidelines. Further research needs to

26 clarify how and which factors impact the accuracy of the MRI-pathway.

27 **Patient summary:** We systematically reported the factors influencing mpMRI accuracy

28 in detecting csPCa. These factors are significantly related to each other with the

29 experience of the radiologists being the dominating factor. In order to deliver the benefits

30 of mpMRI to diagnose csPCa, it is necessary to develop expertise for both radiologists

31 and urologists, to implement quality control, and to adhere to the most recent PI-RADS

32 assessment guidelines.

33

34

35

36

37

38

39

40

41

42

43

44

45

46

1. Introduction

Over the last 10 years the diagnostic pathway of prostate cancer (PCa) is significantly changed by the advent of multiparametric magnetic resonance imaging (mpMRI) [1,2]. As shown by recently published randomized controlled trials, head-to-head comparisons and a Cochrane meta-analyses [3–8] mpMRI is the best technique to detect and localize suspicious areas for clinically significant prostate cancer (csPCa), and it allows to perform MRI targeted biopsy (MRI-TBx) [9]. The value of mpMRI and MRI-TBx over systematic transrectal ultrasound (TRUS)-biopsy is in reducing diagnoses of insignificant (is)PCa and potentially avoiding unnecessary prostate biopsies in men with negative mpMRI scans. The detection and localization of csPCa are slightly but not significantly better compared to TRUS-biopsy.

The acquisition and interpretation of prostate mpMRI are evolving with ongoing improvements, that influence its accuracy. These include magnetic field- and gradient-strength, the use of an endorectal coil (ERC)[10], different versions of assessment-systems, reader experience and inter-reader variability, the potential to avoid a contrast agent (i.e. biparametric MRI - bpMRI), and the use of computer-aided diagnosis (CAD) and deep-learning tools. Another factor that influences the accuracy of the MRI-pathway is the experience of operators performing MRI-TBx. These show a significant variation across reported series [4], affecting the risk of bias of the available data and preventing robust systematic analyses.

Despite the large number of studies reporting the accuracy of mpMRI, there is a lack of comprehensive data, that specifically address the difference of mpMRI execution and performance. Given the promising role of MRI in csPCa diagnosis, there is a need to

70 systematically review the current literature on the factors that influence the variability of

71 mpMRI in the diagnosis of csPCa.

72

73 **2.   Evidence acquisition**

74 *2.1.   Objective*

75   We aimed to systematically review the current literature assessing the factors

76 influencing the variability of mpMRI performance in detecting csPCa. The magnetic field

77 strength (1.5T vs 3.0T, including importance of the gradient-strength), the use of an

78 endorectal coil, the assessment system used by the radiologist, the inter-reader variability,

79 the radiologist and urologist experience, the use of bpMRI and the use of CAD or Deep

80 or Machine Learning (DL, ML) for mpMRI assessment were considered as potential

81 influencing factors.

82

83 *2.2.   Search strategy*

84   Data collection was conducted in accordance with the Preferred Reporting Items

85 for Systematic Reviews and Meta-analyses (PRISMA) statement [11]. A computerized

86 bibliographic search of Medline/PubMed database was searched from inception to 23rd

87 June 2019. The search terms used were (prostate cancer OR prostate adenocarcinoma)

88 AND (MRI OR magnetic resonance) AND (coil OR endorectal coil OR surface coil OR

89 magnetic field OR reporting system OR PI-RADS OR Likert OR inter-reader variability

90 OR inter-reader agreement OR biparametric OR radiologist experience OR urologist

91 experience OR learning curve OR CAD OR machine learning OR computer-aided OR

92 artificial intelligence OR neural network).

93

2.3. Inclusion criteria

95   As recommended by the PRISMA guidelines, we used the Population,

96 Intervention, Comparator and Outcome (PICO) approach to define study eligibility [11].

97 Reports were considered relevant if they provided comparative data on the relationship

98 between the aforementioned factors. MpMRI performance was defined as PCa detection

99 at either prostate biopsy or after radical prostatectomy. Thus, studies assessing one of the

100 factors without providing any comparison (e.g. providing data regarding the 1.5T mpMRI

101 or bpMRI alone accuracy) were not included in this review. Non-comparative studies,

102 case reports, editorials, letters, review articles, and meeting abstracts were not included.

103

2.4. Systematic review process

105   Two authors (AS and FG) independently reviewed a total of 2013 abstracts and

106 selected 77 studies that were finally included in the systematic review for full-text

107 evaluation. Figure 1 shows the PRISMA flowchart describing the selection process.

108

2.5. Data extraction

110 Data were independently extracted from all included studies by the same two authors. A

111 standardized data extraction form was created a priori and used to collect data such as the

112 study design, number of participants, mpMRI protocol, radiologist experience, and

113 outcome.

114

2.6. Data analysis

116 A comprehensive and narrative synthesis of included studies was performed, since

117 quantitative meta-analytic synthesis was not possible due to the heterogeneity of the

118 studies.

119

120 *2.7.   Risk of bias assessment*

121 The risk of bias and applicability concern in individual studies was assessed

122 independently by the same two authors using the Quality Assessment of Diagnostic

123 Accuracy Studies-2 (QUADAS-2) criteria [12]. The presence of baseline confounding

124 factors or selection bias, as well as the presence of any bias within mpMRI protocols,

125 mpMRI interpretation, biopsy protocol, and histopathological reference standard, were

126 assessed (Figure 2; Supplementary Figure 1).

127

128 *3.   Evidence synthesis*

129   Overall, 2013 publications were found. If it was not clear from the abstract

130 whether the paper might contain relevant data, the full paper was assessed. Seventy-seven

131 articles were included in the final analysis (Fig. 1). Single studies are described in detail

132 in Tables 1-4 and Supplementary tables 1-3.

133

134 *3.1.   Risk of bias within studies*

135   The overall risk of bias and applicability concern is given in Figure 2. The overall

136 methodological quality of the studies was moderate with 17 studies having a low risk of

137 bias and applicability concern across all domains assessed. Supplementary Fig. 1 shows

138 the risk of bias and applicability concerns for each study.

139

3.2.   Magnetic field

MpMRI at 3.0T has an increased signal-to-noise ratio compared to 1.5T scanners, resulting in a higher spatial resolution of T2-weighted (T2-WI) and diffusion-weighted imaging (DWI).

In total, seven studies assessed the value of the magnetic field relating to mpMRI staging accuracy (Supplementary Table 1) [13–19]. Two studies from 2004 [14,17] were among the first reporting a comparison of 1.5T and 3.0T scanners even though with slightly different results. Bloch et al.[14] reported the higher image quality at 3.0T compared to 1.5T (both with ERC) while Sosna et al.[17] reported a comparable quality for non-endorectal 3.0T and endorectal 1.5T mpMRI in forty men receiving mpMRI and subsequent biopsy. However, these studies are not comparable given that ERC was not used at 3.0T [17]. Most importantly only two studies addressed diffusion-weighted imaging (DWI) [18,19] and thus only these studies could evaluate the detection difference between insignificant (is)PCa and csPCa, as for this DWI is the most important sequence in the peripheral zone (PZ). These studies showed a similar PI-RADS assessment for 1.5 and 3T.

Overall, the majority of the studies did not investigate the detection of csPCa but instead focused on the recognition of any PCa. In this respect both magnetic field strengths performed equally [13,16,18] but the small sample size and high heterogeneity of these studies make an objective comparison difficult to conduct. In summary, no reliable information could be obtained regarding the detection of csPCa according to field strength regardless of reception coils usage.

162

3.3. Reception Coil

A total of eleven studies comparing the use of mpMRI with an external pelvic phased-array coil with or without ERC were included (Supplementary Table 2)[20–30]. Four studies did not use DWI.

The use of an ERC significantly improves the signal-to-noise ratio irrespective of the magnetic field strength, providing T2-WI with higher spatial resolution and potentially a more accurate delineation of the structures in the transition zone (TZ), which is the key factor in assessment of csPCa in this zone. However, the addition of an ERC is associated with increased costs, increased artifacts [31], organ deformation and discomfort for patients. Mirak et al.[27] investigated the performance of 3.0T mpMRI with and without ERC to detect PCa using PI-RADS v.2 guidelines, with whole-mount histopathology as reference standard. Two sub-cohorts, with (n=260) and without (n=169) ERC, were analyzed. They concluded that detection rates for any PCa, for the index lesion and for csPCa were similar in both cohorts, but there was a significantly lower detection of anterior and TZ csPCa in the ERC sub-cohort due to signal drop off in the anterior gland when ERC was used without an accompanying abdominal phased array coil.

Another study with a direct comparison of ERC vs non-ERC at 3.0T [21] showed no differences in detecting PCa using PI-RADS v2 guidelines.

Costa et al.[25] reported that using an ERC at 3.0T provides superior sensitivity (78%) for PCa detection when compared with standard and augmented protocols (i.e. those with twice as many signal averages; 43% and 60%, respectively) without ERC.

185   Turkbey et al.[22] compared the utility of T2-WI and DWI at 3.0T with and

186   without ERC in detecting PCa in twenty men receiving mpMRI before radical

187   prostatectomy. The authors demonstrated a higher sensitivity when using an ERC. The

188   sensitivity and positive predictive value (PPV) of ERC use vs non-ERC MRI were 76%

189   vs 45% and 80% vs 64%, respectively.

190   Torricelli et al.[29] reported that the image quality with ERC at 1.5T is superior to

191   3.0T without ERC in evaluating tumor conspicuity, capsular infiltration, and seminal

192   vesicle involvement. No significant differences between the two techniques for apex and

193   neurovascular bundles involvement and comparable performance for the diagnosis of

194   capsular involvement were observed. The authors concluded that during preoperative

195   PCa staging, 3.0T mpMRI can provide comparable diagnostic information to 1.5T

196   mpMRI with ERC.

197   In summary, the use of ERC improves signal reception, that slightly improved

198   sensitivity to visualize any PCa. Only one study addressed the value of the ERC with

199   respect to the PI-RADS assessment of csPCa[27]. This study showed a minor

200   disadvantage of the ERC to detect anterior TZ PCa. Due to the lack of significant

201   evidence that the ERC improves csPCa assessment, associated increased costs, duration

202   of examination and patient discomfort, the PI-RADS v. 2.1 guidelines [32] recommend

203   the ERC to be only used only for older 1.5T scanners with lower gradient-strength [29].

204   Still, the lack of standardized protocols in the available studies makes robust comparisons

205   hard to assess.

206

207   *3.4.   Assessment system*

208   The development of a simple, structured, standardized assessment system was one

209   of the most considerable challenges in prostate mpMRI. Assessment systems have been

210   refined during the years, in order to increase the inter-reader agreement, to decrease the

211   gap between differently skilled radiologists and to improve communication between

212   radiologists and urologists. Ten studies comparing the use of different assessment

213   systems were included (Table 1)[33–43]. The majority compared PI-RADS v1[44] and

214   v2[45]. Differently from PI-RADS v1, PI-RADS v2 defined dominant sequences (T2-WI

215   for the TZ and DWI for the PZ) and decreased the role of dynamic contrast-enhanced

216   (DCE) MRI-imaging[46].

217   Most studies reported a similar diagnostic accuracy for both PI-RADS v1 and v2.

218   However, three studies[35,36,38] showed a higher sensitivity of PI-RADS v2 for TZ

219   lesions and another study by Krishna et al.[39] showed that PI-RADS v1 detected

220   approximately 10% more tumors than PI-RADS v2.

221   De Visschere et al.[34] compared the performance of PI-RADS v2 and v1 in 245

222   biopsy naïve men with an elevated prostate-specific antigen (PSA). They found that PI-

223   RADS v1 and PI-RADS v2 yielded similar accuracy to detect csPCa. However, PI-RADS

224   v2 had lower sensitivity than PI-RADS v1 when a score of 3 was used as a threshold for a

225   positive mpMRI. The authors suggested that the majority of discrepancies were caused

226   by a suspicious lesion in the PZ on T2-WI but with normal DWI, scored positive with PI-

227   RADS v1 but negative with PI-RADS v2. Indeed, sensitivity of PI-RADS v2 might be

228   less when suspicious lesions in T2-WI but negative on DWI are present in the PZ.

229   Nonetheless the authors concluded that PI-RADS v2 is preferable because of the more

230   structured and standardized and simpler approach. Similar results were found by Krishna

231   et al.[39] who compared the two assessment systems for the detection of csPCa lesions in

232   47 men before radical prostatectomy. They found no difference in the overall detection of

233   csPCa but observed a higher sensitivity of PI-RADS v1 on T2-WI and DCE.

234   Feng et al. [36] compared PI-RADS v.1 and v.2 in 401 consecutive biopsy naïve

235   men with clinical suspicion of PCa at 3.0T mpMRI. Both assessment systems had a good

236   diagnostic performance for the detection of csPCa, but the diagnostic accuracy increased

237   from 0.82 to 0.88 with the use of PI-RADS v2 compared to v1 when also non-biopsy

238   naïve men were included. Interestingly, PI-RADS v2 had a better performance in the TZ

239   (0.92 vs 0.81).

240   Hoffmann et al.[37] evaluated in 58 men whether PI-RADS v2 was more accurate

241   in assessing anterior prostate csPCa compared to PI-RADS v1. PI-RADS v2 did not

242   improve the accuracy for diagnosing anterior csPCa when compared to PI-RADS v1,

243   however, PI-RADS v.2 was more reproducible between radiologists.

244   Schaudinn et al.[42] reported moderate interobserver reliability (k=0.48) for PCa

245   detection of two radiologists in 40 men before radical prostatectomy, and similar results

246   have been reported by Tewes et al.[43], with a moderate interobserver agreement

247   (k=0.56) for PI-RADS v2 and fair agreement (k=0.39) for PI-RADS v1. The authors

248   concluded that PI-RADS v2 had equivalent diagnostic accuracy to PI-RADS v1 for PCa

249   detection, but with a shorter interpretation time for PI-RADS v2.

250   Two studies compared PI-RADS v1 and two different Likert assessment

251   systems[40,41]. Both showed good inter-reader agreement, even if one study reported a

252   higher accuracy of the Likert scale for TZ lesions compared to PI-RADS v1[41].

253   In summary, it is recommended to use the most recent PI-RADS guidelines as the

254   main assessment system given the simplified, objective, and standardized approach as

255   well as its particular added value for less experienced radiologists.

256

257   *3.5.   Radiologist and urologist experience*

258   The use of mpMRI to detect csPCa represents a multidisciplinary approach, that

259   includes both the skills in acquisition and interpretation of mpMRI and in performing

260   MRI-TBx.

261   Five studies assessed the impact of reader experience on mpMRI diagnostic

262   performance (Table 2). Akin et al. and Garcia-Reyes et al. were the first to test the

263   learning curve in prostate mpMRI reporting using whole-mount pathology as a reference

264   standard. They showed that a dedicated training curriculum is useful to improve mpMRI

265   interpretation [47,48]. Rosenkrantz et al. [49] assessed the variation of mpMRI diagnostic

266   accuracy in detecting and localizing csPCa of six 2nd-year radiology residents reporting

267   124 prostate mpMRI scans (both negatives and positives). Three out of six readers (50%)

268   received feedback after each examination showing the preceding case's solution. For both

269   readers with and without feedback, there was an initial rapid improvement, that slowed

270   down after 40 examinations. In the group receiving feedback, accuracy, sensitivity and

271   specificity improved from 58%, 59%   and 56% to 72-77%, 72-77% and 74-82%,

272   respectively. Interestingly, the presence of feedback did not significantly affect the

273   accuracy as compared to the group without feedback, showing the effects of self-directed

274   learning, even though readers with feedback felt more confident. Moreover, the feedback

275   was more useful for TZ lesions, suggesting a higher challenge in detecting these tumors,

14

276   in line with previous studies[47,48]. Pickersgill et al.[50] conducted a retrospective

277   review of 459 men receiving mpMRI according to PI-RADS (v1 implemented with v2

278   during the study period) and a subsequent MRI-TBx if necessary. They showed, that the

279   radiologist's experience did not improve the accuracy in csPCa detection. The authors

280   speculated that the use of PI-RADS might have attenuated the impact of the reader's

281   experience. However, this study had serious limitations, such as the implementation of

282   PI-RADS version 1 to version 2 during the study period and an arbitrary definition of

283   radiologist experience (i.e. more than 500 mpMRI examinations). Following the

284   widespread use of mpMRI and the need for dedicated training for radiologists, an online

285   interactive case-based website for prostate mpMRI interpretation using PI-RADS version

286   2 has been proposed [51]. This training course increased the sensitivity (58% to 73%,

287   p=0.003) and the negative predictive value (NPV) (69% to 78%, p=0.049) of three 2nd-

288   year radiology residents who evaluated separate sets of sixty MR scans before and after

289   the course. Interestingly, there were no significant improvements in the accuracy of the

290   PI-RADS assessment scores (from 46% to 53%, p=0.149) [51].

291   The quality of MRI-TBx performance plays an equally important role in defining

292   the final diagnostic accuracy of this technique. Similar to systematic ultrasound-guided

293   biopsy[52,53], the experience of the biopsy-operator significantly influences the outcome

294   of an MRI-TBx, that can be visual/cognitive (cogn), MR-US-fusion (fus), or direct-MRI-

295   TBx (in-bore)[54], using either a transrectal or transperineal approach.

296   A total of eight studies assessing the learning curve of MRI-TBx were included

297   (Table 2). Gaziev et al.[55] demonstrated a significant gradual increase in csPCa

298   detection (27% to 63%) in 70 men receiving fusion MRI-TBx. Similarly, Calio et al.[56]

299 reported data from three consecutive cohorts of biopsy-naïve men receiving MRI-TBx

300 over a study period of nine-years. There was a 13% increase in csPCa detection by MRI-

301 TBx from the early to the most recent cohort. Meng et al.[57] reported a 26% increase of

302 csPCa detection in 1500 men receiving repeat biopsy. Mager et al.[58] attempted in

303 demonstrating the presence of the learning curve effect proposing the MRI-TBx quotient,

304 defined as the ratio between the number of positive targeted cores and the total number of

305 targeted cores. The authors showed a significant learning process, in both detection-

306 quotient and biopsy times; for a novice performer a sufficient learning process occurred

307 after 42 procedures reaching a flattening after 63 biopsies. Kasabwala et al.[59]

308 calculated the distance between the planned and the actual core route in the prostatic

309 tissue during a fusion MRI-TBx, and demonstrated a significant improvement in MRI-

310 TBx accuracy after 98 cases. Halstuch et al.[60] attempted to identify a minimum number

311 of procedures to reach the best PCa detection using a mathematical algorithm. The

312 authors demonstrated that at least 104 transrectal fusion MRI-TBx and 119 transperineal

313 fusion MRI-TBx are necessary for men with visible PI-RADS 3 lesions, before reaching

314 the plateau phase of csPCa detection. In this context, Stabile et al.[61] demonstrated the

315 presence of a learning curve affecting csPCa detection rate even when accounting for

316 several confounders (such as PSA, prostate volume and PI-RADS score) for both visual

317 and fusion MRI-TBx. The authors showed a steep increase in csPCa detection during the

318 first 60 procedures and a flattening after 80 procedures. Interestingly, it was suggested

319 that the transperineal approach might be less affected by the learning curve effect, hence

320 it might be easier compared to the transrectal approach when considering MRI-TBx [61].

321 Finally, Westhoff et al.[62] proposed at least eight procedures as experience threshold.

322   However, this result should be cautiously interpreted considering the limited number of

323   MRI-TBx performed by each of the twenty-two urologists included in this study.

324   In summary, when assessing the performance of mpMRI in detecting csPCa it is

325   nowadays mandatory to indicate the experience of the interpreting radiologists and

326   biopsy performing urologists to support the reliability of the findings. Less experienced

327   readers and biopsy-operators should always be supervised by experienced readers and

328   operators. Moreover, mpMRI performance should be internally validated before

329   widespread adoption. According to Rosenkrantz et al. radiologists should have reported

330   at least 100 expert-supervised prostate mpMRI examinations after a dedicated training

331   course [49] and urologists should have performed between 60 and 100 MRI-TBx before

332   they potentially reach an acceptable csPCa detection. Most importantly additional quality

333   assurance tests are needed. Further development of quality-criteria, -assessment, and the

334   development of training platforms/courses are needed.

335

336   *3.6.   Inter-reader variability*

337   Although reader experience plays a substantial role in determining mpMRI

338   accuracy, the reporting process is affected by an almost inevitable variability among

339   different radiologists which varies across different studies and centers. Fifteen studies

340   were included for this topic (Table 3).

341   Quentin et al.[63] assessed the inter-reader agreement of mpMRI using a 5-point

342   (Likert) scale [64]. The authors showed high inter-reader reliability (PPV: 88-96%;

343   k=0.90) between three blinded radiologists scoring 108 lesions. After the introduction of

344   PI-RADS v1 guidelines, Schimmöller et al.[65] reported the agreement of three

experienced radiologists scoring 164 pre-marked lesions. The overall agreement was good to moderate and higher for malignant than for benign lesions. Nonetheless, the clinical application of PI-RADS was still premature, and a diagnostic cut-off had not yet been proposed. Rosenkrantz et al[66] published a series of 55 patients undergoing prostate mpMRI in a single institution that were retrospectively reviewed by three radiologists (two moderately experienced and one inexperienced) using both PI-RADS v1 and Likert scores. For both assessment methods the agreement was strong between the two experienced readers but poor when compared to the inexperienced reader. Interestingly, the Likert assessment scale had better inter-reader reproducibility than PI-RADS score in the TZ. This was probably due to the greater experience of the readers with their 'own' Likert assessment. Since the widespread use of PI-RADS assessment score and the introduction of PI-RADS v2, few studies have assessed its reproducibility, with conflicting results. Mussi et al.[67] reported moderate to good agreement between eight radiologists with different levels of experience using PI-RADS v2. However, this study is hardly applicable to clinical practice since each reader evaluated only one MR slice containing a single specified lesion. Similarly, Glazer et al.[68] conducted a retrospective study with three radiologists (with different levels of experience, ranging from 1 to 11 years) who scored pre-selected lesions, with a moderate agreement for PZ (k=0.46) and fair agreement for TZ (k=0.36). Moreover, the authors disclosed as PI-RADS v2 had been recently introduced in their clinical practice, potentially influencing the level of reproducibility. Girometti et al.[69] supported the higher level of agreement among experienced radiologists in a monocentric study including three radiologists analyzing forty-eight MRI scans, with a moderate agreement both for PI-RADS cut-offs

368 ≥3 (k=0.57) and ≥4 (k=0.63). Nonetheless the readers were aware of the pre-operative

369 reason of mpMRI, hence of the presence of csPCa. Müller et al.[70] reported a poor level

370 of agreement in a cohort of 126 men receiving two consecutive MRI scans at two

371 different institutions. Nonetheless, the design of this study was not devoid of many

372 limitations and significant bias. In fact, among 292 patients referred to the first

373 institution, 126 men had mpMRI lesions considered challenging to be reliably accessed

374 by systematic or cognitive biopsy. Since equipment for MRI targeted biopsy with fusion

375 technique was not available in the first institution, these men were referred to a second

376 institution where they received a second mpMRI before undergoing an MRI targeted

377 fusion biopsy. Also, all readers and authors had limited experience and training in

378 reading prostate MRI according to PI-RADS. Moreover, no information regarding the

379 used PI-RADS version was provided. For these reasons, these results should be very

380 cautiously considered. Rosenkrantz et al.[71] carried out a multi-center study with six

381 experienced radiologists reporting at two different time points (40 and 80 MRI scans per

382 session) and receiving a training session in between. The authors reported moderate

383 reproducibility of PI-RADS v2 suggesting no benefit from the training session[71].

384 However, this study was limited by suboptimal image quality in a number of the included

385 centres. Smith et al.[72] provided results regarding intra- and inter-reader agreement with

386 a multi-centre study on four differently experienced readers. Overall, intra-reader

387 agreement was moderate to substantial (60%-77% of agreement across different

388 radiologists). Inter-reader agreement was poor to moderate and higher for more

389 experienced radiologists. Hansen et al.[73] reported the value of a second opinion by a

390 sub-specialized tertiary care center in reviewing mpMRI from seven different regional

391 hospitals. Overall disagreement was 54% (86/158 MRIs). Specifically, the second

392 reading had significantly improved NPV (0.89 vs 0.72) and PPV (0.61 vs 0.28). Greer et

393 al. [74] reported an excellent agreement on index lesion identification (k=0.87) and

394 moderate on individual PI-RADS v2 category assignment (k=0.419). Other two well-

395 designed studies[75,76] reported similar results with an area under the curve (AUC) for

396 PCa ranging between 0.88 and 0.95 among six blinded readers[75]. Conversely, two

397 recent studies showed high variability in PI-RADS v2 reporting[77,78]. In particular,

398 Sonn et al.[78], in a retrospective study of real-life mpMRI reporting by taking into

399 account nine radiologists and 409 patients, while reporting a low variation in the number

400 of lesions identified, the authors demonstrated high variability in PI-RADS distribution

401 and csPCa detection. The AUC for csPCa ranged between 0.61 and 0.81[78]. Finally, van

402 der Leest at al. [79] in their prospective multicenter head-to-head comparison study

403 showed high inter-reader agreement of the participating expert radiologists. This was

404 obtained after similar training compared to Rosenkranz et al. [49].

405   In summary, most of the well-designed dedicated studies reported moderate

406 agreement when PI-RADS v2 is taken into consideration. Furthermore, the radiologist's

407 experience is crucial to increase mpMRI reproducibility, with the major concerns related

408 with the variability in csPCa yield and high false-positive rates. The heterogeneity across

409 the studies is still high. Most of the studies on this topic did not provide results about

410 MRI-acquisition, or reader-experience, or training. There is thus still a need for

411 standardized mpMRI-assessment training protocols that should be widely available in

412 order to improve the general performance of mpMRI and to provide more reliable data in

413 this context. Only Rosekrantz et al.[49] and van der Leest et al. [79], describe that

414 radiologists should have reported at least 100 expert-supervised prostate mpMRI

415 examinations after a dedicated training course. Further standardization of assessment

416 systems, education, and certification will likely help in reducing the subjectivity and

417 improving the reproducibility also among less experienced readers.

418

419 *3.7.   Biparametric-MRI vs multiparametric-MRI*

420   Since the introduction of a standardized reporting system for mpMRI[44], the role

421 of DCE-MRI has been controversial. Indeed, version 2 of PI-RADS downgraded the role

422 of DCE-MRI to an additional sequence only for upgrading a PI-RADS 3 to 4 PZ lesion

423 on DWI[32]. In the light of the increasing use of mpMRI in the assessment of csPCa and

424 the need for more optimized and efficient protocols, the use of a bpMRI based only on

425 T2-WI and DWI has been proposed by multiple authors [80,81]. The benefits of omitting

426 DCE-MRI are related to reduced examination times, reduced costs and avoiding the risk

427 of adverse events related to the use of contrast agent. Results coming from prospective

428 trials assessing the diagnostic accuracy of bpMRI are promising[80,81] regarding biopsy

429 avoidance and for reducing the detection of insignificant cancers. Comparative studies of

430 mpMRI and bpMRI are mostly retrospective with significant differences in the methods

431 and not negligible risk of bias (Table 4). To the best of our knowledge, Stanzione et

432 al.[82] showed the diagnostic accuracy of bpMRI compared to mpMRI. The authors

433 reported a series of 82 men undergoing mpMRI for suspicion of csPCa then receiving

434 systematic biopsy plus eventual MRI-TBx, with 35% of patients receiving radical

435 prostatectomy. Two experienced radiologists blindly reported bpMRI first, followed by

436 mpMRI (i.e. with DCE), after an interval of 20-30 days to avoid any recall bias. The

overall AUC of bpMRI and mpMRI for csPCa detection was 0.91 and 0.93 (p>0.05). Thestrup et al.[83] reported similar accuracies of bpMRI and mpMRI in detecting csPCa, though without providing any formal statistical comparisons. Lee et al.[84] compared two cohorts undergoing mpMRI and bpMRI for a suspicion of PCa and then receiving visual MRI-TBx in addition to standard systematic biopsy. The authors reported a similar detection of csPCa among men who had suspicious lesions in the bpMRI and mpMRI group (63% and 62%, respectively). Nonetheless, these results should be carefully interpreted since the two cohorts were not randomly matched.

Further studies reported promising results though being affected by significant bias mainly concerning the MRI interpretation process (mpMRI and bpMRI read by the same radiologist during the same session)[85] and the reference standard (no biopsy result in men with negative MRI)[86]. The similar diagnostic performance of these two techniques was confirmed in other series[87–89] that attempted to better identify the eventual differences. Specifically, omitting DCE MRI was related to an increasing rate of PI-RADS 3 lesions and a slightly better specificity and worse sensitivity (though never significant)[87,89]. Furthermore, DCE-MRI was not needed for determination of the overall assessment category in 81% of patients[89].

Choi et al.[90] compared the ability of bpMRI vs mpMRI in detecting the index lesion using radical prostatectomy as reference standard. Two independent radiologists (7 and 13 years of experience) retrospectively reviewed prebiopsy MRI of 113 men using PI-RADS v2. No significant differences in csPCa diagnostic accuracy for bpMRI vs mpMRI for both readers using PI-RADS ≥3 as cut-off were found. Interestingly, both readers reported a significantly higher sensitivity for bpMRI compared to mpMRI[90].

460 Furthermore, inter-reader agreement on PI-RADS assessment score was moderate for

461 both bpMRI (k=0.540) and mpMRI (k=0.478). However, both readers of this study were

462 aware that all men underwent radical prostatectomy for csPCa and this might have

463 affected the overall results. In a similar study, Scialpi et al.[91] evaluated bpMRI and

464 mpMRI ability in detecting single lesions in a cohort of 41 men receiving radical

465 prostatectomy. For both bpMRI and mpMRI the sensitivity was similar and was 100% in

466 PZ and 98 and 95% in the entire prostate and TZ, respectively. BpMRI detected 181/131

467 lesions detected at final pathology resulting in 28% false positives and 3% false negatives

468 rates[91]. Nonetheless, no information regarding readers' background was provided and

469 specific information regarding the experience is often scarce.

470   In this context, Gatti et al.[92] compared bpMRI and mpMRI according to the

471 reader's experience. The authors conducted a retrospective study on 6 blinded

472 radiologists, divided into three groups according to the level of experience, reviewing

473 bpMRI and mpMRI protocols of 68 men receiving a prostate biopsy and eventual radical

474 prostatectomy. The authors used a modified version of PI-RADS v2 [89] for bpMRI

475 reading and a cut-off≥4 for both protocols. Interestingly, the specificity was quite stable

476 regardless of the protocol and the readers' experience. Significant differences were found

477 for sensitivity and AUC in detecting PCa index lesions, mainly related to the rate of true

478 positives. The effect of experience was more evident when considering bpMRI than

479 mpMRI. Consequently, in the highly experienced group, the performance of bpMRI vs

480 mpMRI was similar (AUC:0.86 vs 0.93, p=0.10; true positive: 82% vs 86%, p=0.13). The

481 accuracy of bpMRI became progressively less if compared to mpMRI with the decrease

482 of experience (0.68 vs 0.77 in the less experienced group). Furtherly, the authors

483 attempted to provide a rough indication on the number of necessary cases to reach an

484 AUC and sensitivity ≥0.80: 150-200 for mpMRI and 700-800 for bpMRI[92].

485 Differently, Di Campli et al.[93] showed no diagnostic differences between bpMRI and

486 mpMRI and no significant influence by readers' experience.

487 In summary, available evidence from comparative studies suggests that bpMRI

488 might be a potentially valid alternative to mpMRI, particularly for experienced readers,

489 on the condition that DWI is of excellent quality. These findings have also been

490 confirmed in a recent meta-analysis, demonstrating the non-inferiority of bpMRI and

491 showing on overall non-significant higher sensitivity and lower specificity of mpMRI

492 [94]. Moreover, a recent prospective, multi-reader, blinded direct comparison between

493 bpMRI and mpMRI showed similar diagnostic performance in ruling out the presence of

494 high-grade PCa [79].

495 That being said, the high methodological heterogeneity might have represented a

496 great confounder, and it remains unclear how the performance of bpMRI will translate to

497 less experienced centers and less quality images. Indeed, the assessment system used (i.e.

498 PI-RADS, dedicated bpMRI-PI-RADS), the choice of the cut-off, the magnetic field, the

499 choice of the outcome (i.e. PCa, csPCa) and the reference standard are the factors varying

500 the most across the studies. Ultimately, randomized prospective studies using non-

501 inferiority designs, in men with variable prevalence with clinically meaningful endpoints

502 (biopsy avoidance, detection of csPCa and cisPCa) will be needed to decide on which

503 patient groups can avoid contrast enhancement.

504

505 *3.8.* *Computer-aided diagnosis and deep learning*

506   The first study on PCa detection [95] was carried out by Chan et al. in 2003

507   (Supplementary Table 3)[96]. The CAD mainly consisted of quantitative analysis of

508   medical imaging data (i.e. mpMRI) with the aim to provide results potentially related to

509   clinical conditions (i.e. PCa). Different CAD systems have different strategies for

510   imaging feature analysis with different diagnostic accuracies[97]. The imaging analysis

511   process includes multiple steps that can change across different CAD systems [97,98].

512   The CAD has recently been assessed in aiding radiologists during mpMRI

513   interpretation and reporting. Interestingly, the evaluation of CAD as an assisting tool

514   began at a similar time to the proposal of the first standardized assessment system (i.e.

515   PI-RADS v1)[44,99]. The inclusion of CAD systems in mpMRI interpretation process

516   would potentially overcome some of the issues affecting diagnostic accuracy such as

517   reader experience, reproducibility, as well as enhancing the accuracy of mpMRI itself.

518   Hambrock et al.[99] published the first study evaluating the effect of CAD for

519   both less and more experienced radiologists on the differentiation of benign from

520   malignant lesions at mpMRI. The authors demonstrated as the addition of CAD

521   significantly improved the performance of less experienced radiologists detecting PCa,

522   and when less experienced radiologists used CAD, they achieved similar performance to

523   that of experienced radiologists. Furthermore, stand-alone CAD had similar diagnostic

524   accuracy compared to experienced readers. However, this study did not replicate a real-

525   life mpMRI diagnostic pathway, since the CAD system used was able to differentiate

526   only pre-identified regions between benign and malignant but not PCa detection in a

527   whole gland. In a similar study, Niaf et al.[100], demonstrated that CAD increased both

528   experienced and less-experienced readers' performance (AUC increase of 2% and 4%,

529   respectively, p=0.08). [100]. Litjens et al[101], in a standard clinical scenario, showed

530   that the combination of CAD and PI-RADS assessment categorization achieved higher

531   diagnostic accuracy than PI-RADS categorization alone to discriminate between both

532   benign vs PCa (0.88 vs 0.81, p=0.013) and PCa vs csPCa (0.88 vs 0.78, p<0.01)[101]. In

533   a similar study based on prostatectomy specimens, Wang et al.[102] demonstrated that

534   the combination of CAD information with PI-RADS v2 increased the clinical net benefit

535   for PCa identification as compared to PI-RADS categorization alone. Further studies

536   demonstrated the clinical utility of quantitative analyses of ADC in improving the

537   diagnostic performance when compared to the scoring system alone[103–105]. Giannini

538   et al.[106] was the first to replicate a real-life diagnostic pathway including the use of

539   CAD. On per-patient analysis, the use of CAD achieved a higher sensitivity for csPCa

540   (91% vs 81%, p=0.046), while specificity was not affected. The per-lesion analysis

541   showed a higher number of single lesions detected with CAD assistance. Interestingly,

542   the average reading time with CAD was significantly lower (60s vs 220s; p <0.001)[106].

543   Greer et al.[107] tested the effect of CAD in a multi-institutional group of differently

544   experienced and blinded readers interpreting mpMRI in a cohort of men receiving radical

545   prostatectomy for PCa. Sensitivity increased for all experience levels; however,

546   specificity was dependent on reader experience. Improved sensitivity came from lesions

547   scored as PI-RADS<3 as sensitivity for lesions scored as PI-RADS≥3 was equal. The

548   authors observed that CAD likely helped readers to overcome the "satisfaction of search"

549   limitation which stems from reduced detection for subsequent lesions after one lesion.

550   However, the improved sensitivity of CAD was balanced by decreased specificity. It is

551   noteworthy that the use of CAD improved agreement between all readers[107]. In an

552 effort to assess the clinical value of a CAD system in a "real-world" scenario, Gaur et

553 al.[108] tested a multi-institution population of 216 men receiving mpMRI. The cases

554 used in this study were diverse in terms of institution-specific acquisition, MR

555 manufacturer, and patient population and nine readers of different experiences were

556 included. Overall, sensitivity for index lesions of mpMRI without and with CAD

557 assistance was comparable. The highest benefit of CAD, as compared to mpMRI alone,

558 was observed for moderately experienced readers detecting TZ csPCa. The authors

559 concluded that CAD might help in reducing the rate of false-positivity and increase

560 sensitivity of moderately experienced readers[108]. Furthermore, recent studies

561 confirmed the clinical utility of CAD use in combination with a Likert score and its

562 generalizability to different field strengths[109,110].

563 Taken together these findings suggest that CAD represents a promising tool, that

564 generally improves mpMRI efficiency in terms of both diagnostic accuracy and reporting

565 time. The greatest benefit related to the use of CAD potentially lies in improving mpMRI

566 sensitivity for multifocal csPCa, improving diagnostic accuracy in less experienced

567 readers, improving inter-reader agreement and reducing reporting time. Nonetheless, the

568 majority of the available CAD systems are limited to site-specific pre-defined sequences

569 with most of them not considering DCE sequences, raising therefore similar caveats of

570 those raised when comparing bpMRI and mpMRI. To date CAD is limited to research

571 use only, as several aspects still need to be investigated and standardized.

572

573 *3.9. Discussion*

574 When assessing the role of mpMRI for PCa diagnosis, there are multiple

575   additional aspects that need to be taken into account before drawing conclusions

576   regarding clinical utility. In this study, the factors that might influence the diagnostic

577   accuracy of MRI have been systematically explored. On the basis of our findings, some

578   points need to be discussed.

579   First, it is clear that all the factors described in this study affect mpMRI accuracy,

580   and are extremely related to each other in determining the ability of mpMRI in detecting

581   csPCa. This is of crucial importance when comparing different studies. For instance, the

582   use of an ERC helps to increase mpMRI accuracy for anatomic image detail in the

583   posterior part of the prostate, but without DWI this does not help in better detecting

584   csPCa. The benefit of the ERC is related to the magnetic field strength, which varies

585   across the studies. In fact, according to our findings, a 3.0T scan without ERC and a 1.5T

586   mpMRI with ERC reach similar staging accuracies, and thus anatomic detail. However,

587   the ERC causes compression of the PZ, which may even result in missing small central

588   posterior clinically significant cancers [24]. Considering that the addition of ERC is also

589   associated with increased costs, duration of examination, increased artifacts, and is

590   uncomfortable for patients, the use of ERC is suggested only for older 1.5T MR scanners.

591   3.0T is the recommended magnetic field strength for prostate mpMRI [32]. However, an

592   important issue that is discussed in literature is the gradient-strengths of scanners: steep-

593   gradients are crucial for good DWI. Therefore, gradient-strength steepness may be more

594   important than the field-strength.

595   Second, the radiologist's and urologist's experience is pivotal and affects most of

596   the factors related to mpMRI accuracy, such as the inter-reader agreement and the

597   accuracy of different assessment systems. In this regard, different assessment systems

598   seem to perform differently according to the experience of the reader. Indeed,

599   standardization of the use of one assessment system (i.e. PI-RADS v2) might actually

600   help radiologists with lower experience in prostate mpMRI to easily reach acceptable

601   accuracy. Also, training and certification are important issues. E.g. following an expert

602   hands-on-course, followed by 100 supervised (double) reads may contribute to better

603   interpretation quality and lower inter-reader variability [49,79]. Similarly, the use of

604   bpMRI appears to be most effective for experienced readers, when good image quality is

605   available, whereas those with low experience might still need DCE-MRI as a backup in

606   order to achieve acceptable accuracy. The non-inferiority of bpMRI would represent a

607   great step towards the widespread use of MRI, allowing to reduce the costs (up to $300

608   spared/test [84,87]) and to increase the availability of the test (15 minutes saved per

609   patient [79,87]). With this aim in mind, well-designed prospective trials are necessary to

610   provide reliable evidence and to draw solid conclusions on bpMRI for PCa diagnosis

611   [79]. Even when assessing the use of CAD, the highest benefit in terms of diagnostic

612   accuracy is observed in low-experienced radiologists.

613   To the best of our knowledge, this is the first systematic review exploring the

614   different factors influencing mpMRI accuracy in detecting csPCa. However, this study is

615   limited by the unsatisfactory evidence of most of the studies included, especially related

616   to different MRI protocols, different outcomes, different mpMRI indications, different

617   csPCa prevalence, variable readers' experience and pathological reference standards with

618   only 22% of studies having both low risk of bias and applicability concerns. Moreover,

619   the lack of reliable pathological reference standards prevented the exploration of the

620   variability of mpMRI performance in the presence of a negative report in the majority of

621   studies. This heterogeneity prevents the possibility to conduct a proper meta-analysis on

622   the effect of these factors. Lastly, further factors might be included as potential

623   influencers of mpMRI accuracy (e.g. apparent coefficient diffusion and b-value, the

624   gradient-strength of the MRI scanner, type of scanner, communication methods between

625   radiologist and urologist and the type of study population). Nonetheless, studies

626   specifically addressing these further factors eligible for inclusion in this review are

627   scarce. These limitations are the cause that the findings of this review should be

628   interpreted with caution and within the appropriate context.

629

630

631   *4.   Conclusions*

632   Even though the role of mpMRI in predicting PCa has been widely demonstrated, several

633   factors influence its diagnostic accuracy and are affecting each other, with the

634   radiologist/biopsy-operator experience being the key-confounders. The high

635   heterogeneity across the studies underlines the need for further studies that clarify how

636   they impact on the clinical utility of mpMRI for prostate cancer diagnosis. Indeed, the

637   factors assessed in this study concur with the high variability of mpMRI performance and

638   the related clinical utility, consequently limiting the widespread use of mpMRI. In order

639   to deliver the benefits of mpMRI to diagnose csPCa, it is necessary to develop expertise

640   for both radiologists and urologists, to implement quality control, and to adhere to the

641   most recent PI-RADS assessment guidelines.

642

643

667 **Figure legend**

668 **Figure 1:** Preferred Reporting Items for Systematic Reviews and Meta-analysis flow

669 diagram showing the outcome of the initial and additional searches resulting in the full

670 studies included in the review

671 **Figure 2:** Overall summary of risk of bias and applicability concerns across studies based

672 on Quality Assessment of Diagnostic Accuracy Studies-2.

673 **Supplementary Figure 1:** Risk of bias assessment and applicability concern in included

674 studies according to Quality Assessment of Diagnostic Accuracy Studies-2.

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690   **References**

691   [1]   Stabile A, Giganti F, Rosenkrantz AB, et al. Multiparametric MRI for prostate

692   cancer diagnosis: current status and future directions. Nat Rev Urol 2019.

693   [2]   Giganti F, Rosenkrantz AB, Panebianco V, et al. The Evolution of MRI of the

694   Prostate: The Past, the Present, and the Future. Am J Roentgenol 2019;213:1–13.

695   [3]   Woo S, Suh CH, Eastham JA, et al. Comparison of Magnetic Resonance Imaging-

696   stratified Clinical Pathways and Systematic Transrectal Ultrasound-guided Biopsy

697   Pathway for the Detection of Clinically Significant Prostate Cancer: A Systematic

698   Review and Meta-analysis of Randomized Controlled T. Eur Urol Oncol 2019.

699   [4]   Kasivisvanathan V, Stabile A, Neves JB, et al. Magnetic Resonance Imaging-

700   targeted Biopsy Versus Systematic Biopsy in the Detection of Prostate Cancer: A

701   Systematic Review and Meta-analysis. Eur Urol 2019;44.

702   [5]   Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-Targeted or Standard

703   Biopsy for Prostate-Cancer Diagnosis. N Engl J Med 2018;378:1767–77.

704   [6]   Drost F-JH, Osses D, Nieboer D, et al. Prostate Magnetic Resonance Imaging,

705   with or Without Magnetic Resonance Imaging-targeted Biopsy, and Systematic

706   Biopsy for Detecting Prostate Cancer: A Cochrane Systematic Review and Meta-

707   analysis. Eur Urol 2019.

708   [7]   van der Leest M, Cornel EB, Israel B, et al. Head-to-head Comparison of

709   Transrectal Ultrasound-guided Prostate Biopsy Versus Multiparametric Prostate

710   Resonance Imaging with Subsequent Magnetic Resonance-guided Biopsy in

711   Biopsy-naïve Men with Elevated Prostate-specific Antigen : A Large Prospective

712   M. Eur Urol 2019;75:570–8.

713 [8] Rouvière O, Puech P, Renard-penna R, et al. Use of prostate systematic and
714   targeted biopsy on the basis of multiparametric MRI in biopsy-naive patients
715   (MRI-FIRST): a prospective , multicentre, paired diagnostic study. Lancet Oncol
716   2019;20:100–9.

717 [9] Moore CM, Robertson NL, Arsanious N, et al. Image-Guided Prostate Biopsy
718   Using Magnetic Resonance Imaging – Derived Targets : A Systematic Review.
719   Eur Urol 2013;63:125–40.

720 [10] Engels R, Israel B, Padhani AR, Barentsz JO, Mottrie A. Multiparametric
721   Magnetic Resonance Imaging for the Detection of Clinically Significant Prostate
722   Cancer : What Urologists Needs to Know . Part 1 : Acquisition. Eur Urol 2019:1–
723   12.

724 [11] Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic
725   review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev
726   2015:1–9.

727 [12] Whiting P, Rutjes A, Westwood M, et al. QUADAS-2: a revised tool for the
728   quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–
729   36.

730 [13] Beyersdorff D, Taymoorian K, Knösel T, et al. MRI of prostate cancer at 1.5 and
731   3.0 T: Comparison of image quality in tumor detection and staging. Am J
732   Roentgenol 2005;185:1214–20.

733 [14] Bloch BN, Rofsky NM, Baroni RH, et al. 3 Tesla magnetic resonance imaging of
734   the prostate with combined pelvic phased-array and endorectal coils: Initial
735   experience. Acad Radiol 2004;11:863–7.

736   [15]   Sertdemir M, Schoenberg SO, Sourbron S, et al. Interscanner comparison of

737   dynamic contrast-enhanced mri in prostate cancer: 1.5 versus 3 T MRI. Invest

738   Radiol 2013;48:92–7.

739   [16]   Shah ZK, Elias SN, Abaza R, et al. Performance Comparison of 1.5-T Endorectal

740   Coil MRI with 3.0-T Nonendorectal Coil MRI in Patients with Prostate Cancer.

741   Acad Radiol 2015;22:467–74.

742   [17]   Sosna J, Pedrosa I, Dewolf WC, et al. MR imaging of the prostate at 3 tesla:

743   Comparison of an external phased-array coil to imaging with an endorectal coil at

744   1.5 tesla. Acad Radiol 2004;11:857–62.

745   [18]   Ullrich T, Quentin M, Oelers C, et al. Magnetic resonance imaging of the prostate

746   at 1.5 versus 3.0 T: A prospective comparison study of image quality. Eur J Radiol

747   2017;90:192–7.

748   [19]   Thompson JE, Moses D, Shnier R, et al. Multiparametric magnetic resonance

749   imaging guided diagnostic biopsy detects significant prostate cancer and could

750   reduce unnecessary biopsies and over detection: A prospective study. J Urol

751   2014;192:67–74.

752   [20]   Barth BK, Cornelius A, Nanz D, Eberli D, Donati OF. Comparison of image

753   quality and patient discomfort in prostate MRI: pelvic phased array coil vs.

754   endorectal coil. Abdom Radiol 2016;41:2218–26.

755   [21]   Barth BK, Rupp NJ, Cornelius A, et al. Diagnostic Accuracy of a MR protocol

756   acquired with and without endorectal coil for detection of prostate cancer: A

757   multicenter study. Curr Urol 2019:88–96.

758   [22]   Turkbey B, Merino MJ, Gallardo EC, et al. Comparison of endorectal coil and

759   nonendorectal coil T2W and diffusion-weighted MRI at 3 Tesla for localizing

760   prostate cancer: Correlation with whole-mount histopathology. J Magn Reson

761   Imaging 2014;39:1443–8.

762 [23] Baur ADJ, Daqqaq T, Wagner M, et al. T2- and diffusion-weighted magnetic

763   resonance imaging at 3 T for the detection of prostate cancer with and without

764   endorectal coil: An intraindividual comparison of image quality and diagnostic

765   performance. Eur J Radiol 2016;85:1075–84.

766 [24] Chernyak V, Flusberg M, Kurteva T, Ghavamian R, Rozenblit AM. Accuracy of

767   prostate measurements on MRI with and without an endorectal coil. Clin Imaging

768   2015;39:85–8.

769 [25] Costa DN, Yuan Q, Xi Y, et al. Comparison of prostate cancer detection at 3-T

770   MRI with and without an endorectal coil: A prospective, Paired-patient study. Urol

771   Oncol Semin Orig Investig 2016;34:255.e7-255.e13.

772 [26] Gawlitza J, Reiss-Zimmermann M, Thörmer G, et al. Impact of the use of an

773   endorectal coil for 3 T prostate MRI on image quality and cancer detection rate.

774   Sci Rep 2017;7:1–8.

775 [27] Mirak SA, Shakeri S, Bajgiran AM, et al. Three Tesla Multiparametric Magnetic

776   Resonance Imaging : Comparison of Performance with and without Endorectal

777   Coil for Prostate Cancer Detection , PI-RADS TM version 2 Category and Staging

778   with Whole Mount Histopathology Correlation 2019;201:496–502.

779 [28] O'Donohoe RL, Dunne RM, Kimbrell V, Tempany CM. Prostate MRI using an

780   external phased array wearable pelvic coil at 3T: comparison with an endorectal

781   coil. Abdom Radiol 2019;44:1062–9.

782 [29] Torricelli P, Cinquantini F, Ligabue G, et al. Comparative evaluation between

783 external phased array coil at 3 T and endorectal coil at 1.5 T: Preliminary results. J

784 Comput Assist Tomogr 2006;30:355–61.

785 [30] De Visschere P, Nezzo M, Pattyn E, et al. Prostate magnetic resonance

786 spectroscopic imaging at 1.5tesla with endorectal coil versus 3.0tesla without

787 endorectal coil: Comparison of spectral quality. Clin Imaging 2015;39:636–41.

788 [31] Husband JE, Padhani AR, Macvicar AD, Revell P. Magnetic Resonance Imaging

789 of Prostate Cancer : Comparison of Image Quality Using Endorectal and Pelvic

790 Phased Array Coils. Clin Radiol 1998;53:673–81.

791 [32] Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging -

792 Reporting and Data System: 2015, Version 2. Eur Urol 2015;69:16–40.

793 [33] Auer T, Edlinger M, Bektic J, et al. Performance of PI-RADS version 1 versus

794 version 2 regarding the relation with histopathological results. World J Urol

795 2017;35:687–93.

796 [34] Visschere P De, Pattyn E, Ost P, Claeys T, Lumen N. Comparison of the Prostate

797 Imaging Reporting and Data System (PI-RADS) Version 1 and 2 in a Cohort of

798 245 Patients with Histopathological Reference and 2016;100:1–10.

799 [35] Wang X, Bao J, Ping X, et al. The diagnostic value of PI-RADS V1 and V2 using

800 multiparametric MRI in transition zone prostate clinical cancer. Oncol Lett

801 2018;16:3201–6.

802 [36] Feng ZY, Wang L, Min X De, et al. Prostate cancer detection with multiparametric

803 magnetic resonance imaging: Prostate imaging reporting and data system version 1

804 versus version 2. Chin Med J (Engl) 2016;129:2451–9.

805   [37]   Hoffmann R, Logan C, O'Callaghan M, et al. Does the Prostate Imaging-

806   Reporting and Data System (PI-RADS) version 2 improve accuracy in reporting

807   anterior lesions on multiparametric magnetic resonance imaging (mpMRI)? Int

808   Urol Nephrol 2018;50:13–9.

809   [38]   Polanec S, Helbich TH, Bickel H, et al. Head-to-head comparison of PI-RADS v2

810   and PI-RADS v1. Eur J Radiol 2016;85:1125–31.

811   [39]   Krishna S, McInnes M, Lim C, et al. Comparison of prostate imaging reporting

812   and data system versions 1 and 2 for the detection of peripheral zone gleason score

813         $3 + 4 = 7$ cancers. Am J Roentgenol 2017;209:W365–73.

814   [40]   Renard-Penna R, Mozer P, Cornud F, et al. Prostate Imaging Reporting and Data

815   System and Likert Scoring System: Multiparametric MR Imaging Validation

816   Study to Screen Patients for Initial Biopsy. Radiology 2015;275:458–68.

817   [41]   Rosenkrantz AB, Kim S, Lim RP, et al. Prostate cancer localization using

818   multiparametric MR imaging: Comparison of Prostate Imaging Reporting and

819   Data System (PI-RADS) and Likert scales. Radiology 2013;269:482–92.

820   [42]   Schaudinn A, Gawlitza J, Mucha S, et al. Comparison of PI-RADS v1 and v2 for

821   multiparametric MRI detection of prostate cancer with whole-mount histological

822   workup as reference standard. Eur J Radiol 2019;116:180–5.

823   [43]   Tewes S, Mokov N, Hartung D, et al. Standardized reporting of prostate MRI:

824   Comparison of the prostate imaging reporting and data system (PI-RADS) version

825   1 and version 2. PLoS One 2016;11:1–13.

826   [44]   Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012.

827   Eur Radiol 2012;22:746–57.

828   [45]   Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS Prostate Imaging -

829   Reporting and Data System: 2015, Version 2. Eur Urol 2016;69:16–40.

830   [46]   Barentsz JO, Weinreb JC, Verma S, et al. Synopsis of the PI-RADS v2 Guidelines

831   for Multiparametric Prostate Magnetic Resonance Imaging and Recommendations

832   for Use. Eur Urol 2015;69:41–9.

833   [47]   Akin O, Riedl CC, Ishill NM, et al. Interactive dedicated training curriculum

834   improves accuracy in the interpretation of MR imaging of prostate cancer. Eur

835   Radiol 2010;20:995–1002.

836   [48]   Garcia-Reyes K, Passoni NM, Palmeri ML, et al. Detection of prostate cancer with

837   multiparametric MRI (mpMRI): effect of dedicated reader education on accuracy

838   and confidence of index and anterior cancer diagnosis. Abdom Imaging

839   2015;40:134–42.

840   [49]   Rosenkrantz AB, Abimbola A, Anunita K, et al. The Learning Curve in Prostate

841   MRI Interpretation: Self-Directed Learning Versus Continual Reader Feedback.

842   AJR Am J Roentgenol 2017;208:92–100.

843   [50]   Pickersgill NA, Vetter JM, Raval NS, et al. The Accuracy of Prostate Magnetic

844   Resonance Imaging Interpretation: Impact of the Individual Radiologist and

845   Clinical Factors. Urology 2019;127:68–73.

846   [51]   Rosenkrantz AB, Begovic J, Pires A, et al. Online Interactive Case-Based

847   Instruction in Prostate Magnetic Resonance Imaging Interpretation Using Prostate

848   Imaging and Reporting Data System Version 2: Effect for Novice Readers. Curr

849   Probl Diagn Radiol 2018;48:132–41.

850   [52]   Benchikh El Fegoun A, El Atat R, Choudat L, et al. The learning curve of

851   transrectal ultrasound-guided prostate biopsies: Implications for training programs.

852   Urology 2013;81:12–6.

853   [53]   Tadtayev S, Hussein A, Carpenter L, Vasdev N, Boustead G. The association of

854   level of practical experience in transrectal ultrasonography guided prostate biopsy

855   with its diagnostic outcome. Ann R Coll Surg Engl 2017;99:218–23.

856   [54]   Bjurlin MA, Meng X, Le Nobin J, et al. Optimization of prostate biopsy: the role

857   of magnetic resonance imaging targeted biopsy in detection, localization and risk

858   assessment. J Urol 2014;192:648–58.

859   [55]   Gaziev G, Wadhwa K, Barrett T, et al. Defining the learning curve for

860   multiparametric magnetic resonance imaging (MRI) of the prostate using MRI-

861   transrectal ultrasonography (TRUS) fusion-guided transperineal prostate biopsies

862   as a validation tool. BJU Int 2016;117:80–6.

863   [56]   Calio B, Sidana A, Sugano D, et al. Changes in prostate cancer detection rate of

864   MRI-TRUS fusion vs systematic biopsy over time: evidence of a learning curve.

865   Prostate Cancer Prostatic Dis 2017;20:436–41.

866   [57]   Meng X, Rosenkrantz AB, Huang R, et al. The Institutional Learning Curve of

867   Magnetic Resonance Imaging-Ultrasound Fusion Targeted Prostate Biopsy:

868   Temporal Improvements in Cancer Detection in 4 Years. J Urol 2018;200:1022–9.

869   [58]   Mager R, Brandt MP, Borgmann H, et al. From novice to expert: analyzing the

870   learning curve for MRI-transrectal ultrasonography fusion-guided transrectal

871   prostate biopsy. Int Urol Nephrol 2017;49:1–8.

872   [59]   Kasabwala K, Patel N, Cricco-lizza E, et al. The Learning Curve for Magnetic

873   Resonance Imaging / Ultrasound Fusion-guided Prostate Biopsy Image fusion. Eur

874   Urol Oncol 2018;2:135–40.

875   [60]   Halstuch D, Baniel J, Lifshitz D, et al. Characterizing the learning curve of MRI-
876   US fusion prostate biopsies. Prostate Cancer Prostatic Dis 2019.

877   [61]   Stabile A, Dell'Oglio P, Gandaglia G, et al. Not All Multiparametric Magnetic
878   Resonance Imaging – targeted Biopsies Are Equal : The Impact of the Type of
879   Approach and Operator Expertise on the Detection of Clinically Significant
880   Prostate Cancer. Eur Urol Oncol 2018;1:120–8.

881   [62]   Westhoff N, Haumann H, Kriegmair MC, et al. Association of training level and
882   outcome of software-based image fusion-guided targeted prostate biopsies. World
883   J Urol 2018.

884   [63]   Quentin M, Arsov C, Röhlen S, et al. Inter-reader agreement of multi-parametric
885   MR imaging for the detection of prostate cancer: Evaluation of a scoring system.
886   RoFo 2012;184:925–9.

887   [64]   Dickinson L, Ahmed HU, Allen C, et al. Magnetic resonance imaging for the
888   detection, localisation, and characterisation of prostate cancer: Recommendations
889   from a European consensus meeting. Eur Urol 2011;59:477–94.

890   [65]   Schimmöller L, Quentin M, Arsov C, et al. Inter-reader agreement of the ESUR
891   score for prostate MRI using in-bore MRI-guided biopsies as the reference
892   standard. Eur Radiol 2013;23:3185–90.

893   [66]   Rosenkrantz AB, Lim RP, Haghighi M, et al. Comparison of interreader
894   reproducibility of the prostate imaging reporting and data system and likert scales
895   for evaluation of multiparametric prostate MRI. Am J Roentgenol 2013;201:612–
896   8.

897   [67]   Mussi TC, Yamauchi I, Tridente FC, et al. Interobserver Agreement and Positivity

898   of PI-RADS Version 2 Among Radiologists with Different Levels of Experience.

899   Acad Radiol 2018:1–6.

900   [68]   Glazer DI, Mayo-Smith WW, Sainani NI, et al. Interreader agreement of prostate

901   imaging reporting and data system version 2 using an in-bore mri-guided prostate

902   biopsy cohort: A single institution's initial experience. Am J Roentgenol

903   2017;209:W145–51.

904   [69]   Girometti R, Giannarini G, Greco F, et al. Interreader agreement of PI-RADS v. 2

905   in assessing prostate cancer with multiparametric MRI: A study using whole-

906   mount histology as the standard of reference. J Magn Reson Imaging

907   2019;49:546–55.

908   [70]   Müller S, Lilleaasen G, Sand TE, et al. Poor reproducibility of PIRADS score in

909   two multiparametric MRIs before biopsy in men with elevated PSA. World J Urol

910   2018;36:687–91.

911   [71]   Rosenkrantz AB, Ginocchio LA, Cornfeld D, et al. Interobserver Reproducibility

912   of the PI-RADS Version 2 Lexicon: A Multicenter Study of Six Experienced

913   Prostate Radiologists. Radiology 2016;000:152542.

914   [72]   Smith CP, Harmon SA, Barrett T, et al. Intra- and Interreader Reproducibility of

915   PI-RADSv2: A Multireader Study. J Magn Reson Imaging 2018:1–10.

916   [73]   Hansen NL, Koo BC, Gallagher FA, et al. Comparison of initial and tertiary centre

917   second opinion reads of multiparametric magnetic resonance imaging of the

918   prostate prior to repeat biopsy. Eur Radiol 2017;27:2259–66.

919   [74]   Greer MD, Shih JH, Lay N, et al. Interreader Variability of Prostate Imaging

920   Reporting and Data System Version 2 in Detecting and Assessing Prostate Cancer

921   Lesions at Prostate MRI. AJR Am J Roentgenol 2019;27:1–8.

922   [75] Ke Z, Wang L, Min X De, et al. Diagnostic performance and interobserver

923   consistency of the prostate imaging reporting and data system version 2: A study

924   on six prostate radiologists with different experiences from half a year to 17 years.

925   Chin Med J (Engl) 2018;131:1666–73.

926   [76] Purysko AS, Bittencourt LK, Bullen JA, et al. Accuracy and interobserver

927   agreement for prostate imaging reporting and data system, version 2, for the

928   characterization of lesions identified on multiparametric MRI of the prostate. Am J

929   Roentgenol 2017;209:339–45.

930   [77] Pickersgill NA, Vetter JM, Andriole GL, et al. Accuracy and Variability of

931   Prostate Multiparametric Magnetic Resonance Imaging Interpretation Using the

932   Prostate Imaging Reporting and Data System: A Blinded Comparison of

933   Radiologists. Eur Urol Focus 2018:6–11.

934   [78] Sonn GA, Fan RE, Ghanouni P, et al. Prostate Magnetic Resonance Imaging

935   Interpretation Varies Substantially Across Radiologists. Eur Urol Focus 2018:1–8.

936   [79] Leest M Van Der, Israe B, Cornel EB, et al. High Diagnostic Performance of Short

937   Magnetic Resonance Imaging Protocols for Prostate Cancer Detection in Biopsy-

938   naïve Men : The Next Step in Magnetic Resonance Imaging Accessibility. Eur

939   Urol 2019;76:582–3.

940   [80] Jambor I, Boström PJ, Taimen P, et al. Novel biparametric MRI and targeted

941   biopsy improves risk stratification in men with a clinical suspicion of prostate

942   cancer (IMPROD Trial). J Magn Reson Imaging 2017;46:1089–95.

943   [81]   Boesen L, Norgaard N, Logager V, et al. Assessment of the Diagnostic Accuracy

944   of Biparametric Magnetic Resonance Imaging for Prostate Cancer in Biopsy-Naive

945   Men: The Biparametric MRI for Detection of Prostate Cancer (BIDOC) study.

946   Jama 2018;1:1–28.

947   [82]   Stanzione A, Imbriaco M, Cocozza S, et al. Biparametric 3T Magentic Resonance

948   Imaging for prostatic cancer detection in a biopsy-naïve patient population: a

949   further improvement of PI-RADS v2? Eur J Radiol 2016;85:2269–74.

950   [83]   Thestrup KCD, Logager V, Baslev I, et al. Biparametric versus multiparametric

951   MRI in the diagnosis of prostate cancer. Acta Radiol Open

952   2016;5:205846011666304.

953   [84]   Lee DH, Nam JK, Lee SS, et al. Comparison of multiparametric and biparametric

954   MRI in first round cognitive targeted prostate biopsy in patients with PSA levels

955   under 10 ng/mL. Yonsei Med J 2017;58:994–9.

956   [85]   Kuhl CK, Bruhn R, Krämer N, et al. Abbreviated Biparametric Prostate MR

957   imaging in Men with elevated Prostate-specific antigen. Radiology 2017;285:493–

958   505.

959   [86]   Nieuwenhove S, Saussez T, Thiry S, Trefois P, Annet L. Prospective comparison

960   of a fast 1.5-T biparametric with the 3.0-T multiparametric ESUR magnetic

961   resonance imaging protocol as a triage test for men at risk of prostate cancer. BJU

962   Int 2018:0–2.

963   [87]   Junker D, Steinkohl F, Fritz V, et al. Comparison of multiparametric and

964   biparametric MRI of the prostate: are gadolinium-based contrast agents needed for

965   routine examinations? World J Urol 2019;37:691–9.

966 [88] Sherrer RL, Glaser ZA, Gordetsky JB, et al. Comparison of biparametric MRI to

967   full multiparametric MRI for detection of clinically significant prostate cancer.

968   Prostate Cancer Prostatic Dis 2019;22:331–6.

969 [89] De Visschere P, Lumen N, Ost P, et al. Dynamic contrast-enhanced imaging has

970   limited added value over T2-weighted imaging and diffusion-weighted imaging

971   when using PI-RADSv2 for diagnosis of clinically significant prostate cancer in

972   patients with elevated PSA. Clin Radiol 2017;72:23–32.

973 [90] Choi MH, Kim CK, Lee YJ, et al. Prebiopsy Biparametric MRI for Clinically

974   Significant Prostate Cancer Detection With PI-RADS Version 2: A Multicenter

975   Study. Genitourin Imaging 2019;212:1–8.

976 [91] Scialpi M, Prosperi E, Andrea AD, et al. Biparametric versus Multiparametric

977   MRI with Non-endorectal Coil at 3T in the Detection and Localization of Prostate

978   Cancer. Anticancer Res 2017;37:1263–72.

979 [92] Gatti M, Faletti R, Calleris G, et al. Prostate cancer detection with biparametric

980   magnetic resonance imaging (bpMRI) by readers with different experience :

981   performance and comparison with multiparametric (mpMRI). Abdom Radiol

982   2019;44:1883–93.

983 [93] Di Campli E, Delli Pizzi A, Seccia B, et al. Diagnostic accuracy of biparametric vs

984   multiparametric MRI in clinically significant prostate cancer: Comparison between

985   readers with different experience. Eur J Radiol 2018;101:17–23.

986 [94] Alabousi M, Salameh JP, Gusenbauer K, et al. Biparametric vs multiparametric

987   prostate magnetic resonance imaging for the detection of prostate cancer in

988   treatment-naïve patients: a diagnostic test accuracy systematic review and meta-

989   analysis. BJU Int 2019:0–1.

990   [95]   Chan H, Sahiner B, Helvie M, et al. Improvement of radiologists' characterization

991   of mammographic masses by using computer-aided diagnosis: an ROC study.

992   Radiology 1999;212:817–27.

993   [96]   Chan I, Wells W, Mulkern R, et al. Detection of prostate cancer by integration of

994   line-scan diffusion, T2-mapping and T2-weighted magnetic resonance imaging; a

995   multichannel statistical classifier. Med Phys 2003;30:2390–8.

996   [97]   Lemaître G, Martí R, Freixenet J, et al. Computer-Aided Detection and diagnosis

997   for prostate cancer based on mono and multi-parametric MRI: A review. Comput

998   Biol Med 2015;60:8–31.

999   [98]   Sun Y, Reynolds HM, Parameswaran B, et al. Multiparametric MRI and radiomics

1000   in prostate cancer: a review. Australas Phys Eng Sci Med 2019;42:3–25.

1001   [99]   Hambrock T, Vos PC, Hulsbergen–Van de Kaa CA, Barentsz JO, Hulsman HJ.

1002   Computer-aided diagnosis with multiparametric 3-T MR imaging - Effect on

1003   observer performance. Radiology 2013;266:521–30.

1004   [100]   Niaf E, Lartizien C, Bratan F, et al. Prostate Focal Peripheral Zone lesions:

1005   Characterization at Multiparametric MR Imaging — Influence of a Computer-

1006   aided Diagnosis System. Radiology 2014;271:761–9.

1007   [101]   Litjens GJS, Barentsz JO, Karssemeijer N, Huisman HJ. Clinical evaluation of a

1008   computer-aided diagnosis system for determining cancer aggressiveness in prostate

1009   MRI. Eur Radiol 2015;25:3187–99.

1010   [102]   Wang J, Wu CJ, Bao ML, et al. Machine learning-based analysis of MR radiomics

1011   can help to improve the diagnostic performance of PI-RADS v2 in clinically

1012   relevant prostate cancer. Eur Radiol 2017;27:4082–90.

1013   [103] Park SY, Shin S, Jung DC, Cho NH. PI-RADS version 2 : quantitative analysis

1014   aids reliable interpretation of diffusion-weighted imaging for prostate cancer. Eur

1015   Radiol 2016.

1016   [104] Bonekamp D, Kohl S, Wiesenfarth M, Schelb P, Philipp J. Radiomic Machine

1017   Learning for Characterization of Prostate Lesions with MRI : Comparison to ADC

1018   Values. Radiology 2018;289:128–37.

1019   [105] Moraes MO, Roman DH, Copetti J, et al. Effects of the addition of quantitative

1020   apparent diffusion coefficient data on the diagnostic performance of the PI - RADS

1021   v2 scoring system to detect clinically significant prostate cancer. World J Urol

1022   2019.

1023   [106] Giannini V, Mazzetti S, Armando E, et al. Multiparametric magnetic resonance

1024   imaging of the prostate with computer-aided detection : experienced observer

1025   performance study. Eur Radiol 2017;27:4200–8.

1026   [107] Greer MD, Lay N, Shih JH, et al. Computer-aided diagnosis prior to conventional

1027   interpretation of prostate mpMRI: an international multi-reader study. Eur Radiol

1028   2018.

1029   [108] Gaur S, Lay N, Harmon SA, et al. Can computer-aided diagnosis assist in the

1030   identification of prostate cancer on prostate MRI ? a multi-center , multi-reader

1031   investigation. Oncotarget 2018;9:33804–17.

1032   [109] Antonelli M, Johnston EW, Dikaios N, et al. Machine learning classifiers can

1033   predict Gleason pattern 4 prostate cancer with greater accuracy than experienced

1034   radiologists. Eur Radiol 2019;29:43–5.

1035   [110] Dikaios N, Giganti F, Sidhu HS, et al. Multi-parametric MRI zone-specific

1036   diagnostic model performance compared with experienced radiologists for

1037   detection of prostate cancer. Eur Radiol 2019;29:4150–9.
1038