Diagnostic and prognostic factors in patients with prostate cancer: a systematic review

Katharina Beyer, Lisa Moris, Michael Lardas, Anna Haire, Francesco Barletta, Simone Scuderi, Megan Molnar, Ronald Herrera, Abdul Raut, Riccardo Campi, Isabella Greco, Kirill Shiranov, Saeed Dabestani, Thomas van den Broeck, Sujenthiran Arun, Mauro Gacci, Giorgio Gandaglia, Muhammad Imran Omar, Steven MacLennan, Monique J Roobol, Bahman Farahmand, Eleni Vradi, Zsuzsanna Devecseri, Alex Asiimwe, Jihong Zong, Sara J Maclennan, Laurence Collette, James NDow, Alberto Briganti, Anders Bjartell, Mieke Van Hemelrijck and the PIONEER Consortium

ABSTRACT

Objectives As part of the PIONEER Consortium objectives, we have explored which diagnostic and prognostic factors (DPFs) are available in relation to our previously defined clinician and patient-reported outcomes for prostate cancer (PCa).

Design We performed a systematic review to identify validated and non-validated studies.

Data sources MEDLINE, Embase and the Cochrane Library were searched on 21 January 2020.

Eligibility criteria Only quantitative studies were included. Single studies with fewer than 50 participants, published before 2014 and looking at outcomes which are not prioritised in the PIONEER core outcome set were excluded.

Data extraction and synthesis After initial screening, we extracted data following the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of prognostic factor studies (CHARMS-PF) criteria and discussed the identified factors with a multidisciplinary expert group. The quality of the included papers was scored for applicability and risk of bias using validated tools such as PROBAST, Quality in Prognostic Studies and Quality Assessment of Diagnostic Accuracy Studies 2.

Results The search identified 6604 studies, from which 489 DPFs were included. Sixty-four of those were internally or externally validated. However, only three studies on diagnostic and seven studies on prognostic factors had a low risk of bias and a low risk concerning applicability.

Conclusion Most of the DPFs identified require additional evaluation and validation in properly designed studies before they can be recommended for use in clinical practice. The PIONEER online search tool for DPFs for PCa will enable researchers to understand the quality of the current research and help them design future studies.

INTRODUCTION

Prostate cancer (PCa) accounts for 15% of cancers diagnosed and is the second most common cancer in males worldwide. PCa is clinically and molecularly heterogeneous and is usually suspected based on the clinical findings of digital rectal examination and/or prostate-specific antigen (PSA) levels. However, which diagnostic or prognostic factors (DPFs) can be used to select patients for specific therapeutic options remains largely unclear. Specific biomarkers in urine or in blood are available on top of traditional PSA testing, such as PCA3, TMPRSS2-ERG fusion or kallikreins as incorporated in the Phi or 4Kscore test together with other parameters including family history. However, the European Association of Urology (EAU) guidelines (2021) currently do not provide general recommendations to implement these biomarkers into routine screening programmes due to limited data. As part of the American Society of Clinical Oncology (ASCO) guidelines, Eggener et al
recommended commercially available biomarkers, which have been shown to provide prognostic significance and additional information beyond standard clinical models in patient selection in the localised context: Oncotype Dx Prostate, Prolaris, Decipher, and ProMark. However, no guidelines have recommended DPFs for other stages of PCa. The expert panel at the Advanced Prostate Cancer Consensus Conference (APCCC) consensus meeting of advanced PCa in Basel 2019, recommended AR-V7 for mCRPC as potentially useful, which ultimately led to the inclusion of AR-V7 testing in the NCCN guidelines.

The PIONEER Consortium is an international collaboration coordinated by the EAU, which aims to establish the best evidence-based management and clinical practice of PCa across all disease stages using the power of big data analytics towards a more outcome-driven, value-based and patient-centric healthcare system. A key objective is to address one of the major challenges within the context of diagnostic or prognostic biomarkers/factors: the inability to incorporate DPFs into the management of PCa in terms of screening, diagnosis and treatment. It is therefore important to summarise and evaluate the evidence. Biomarkers can be classified into different types: diagnostic, prognostic, predictive and therapeutic—in this study we focus on the first two. A diagnostic biomarker or factor is useful when cancer is suspected and allows the early detection based on symptoms or tests. The overall aim of a diagnostic biomarker is to distinguish people with the diseases from people without the disease. A prognostic biomarker or factor is a clinical or biological characteristic which provides information on the likely course of the disease, that is, biochemical progression or disease recurrence. It enables clinicians to decide on the most suitable treatment depending on the likely course of the disease. In the sections below, we have used the terms biomarkers and factors interchangeably. Multiple DPFs can be measured in tissue, blood or urine. These come with different advantages and disadvantages and only a limited number of factors are currently available for PCa in standard clinical care.

We aimed to systematically review the evidence from 2014 onward to assess which DPFs are available in relation to previously defined outcomes for PCa.

**METHODS**

The systematic review (SR) followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A detailed protocol of the overall project was published elsewhere (please see the protocol attached as methods online supplemental appendix). Briefly, we followed the following four steps (figure 1):

1. Comprehensive systematic literature review of DPFs for all stages of PCa (localised, locally advanced, metastatic, and non-metastatic castration resistant) from 2014 onwards. DPFs developed before 2014 were not included, due to the significant changes influencing the staging of PCa (i.e., Consensus Conference on Gleason Grading of Prostatic Carcinoma [60]) that have taken place in diagnostic and prognostic practice and patient management since then.

2. Assessment and identification of final list of DPFs by a multidisciplinary expert panel.

3. Evaluation of quality of studies published using risk of bias (RoB) tools: Prediction model Risk Of Bias Assessment Tool (PROBAST), Prediction model Risk Of Bias Assessment Tool (QUADAS-2), Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for diagnostic factors.

4. Due to the heterogeneity of the studies identified no further formal quantitative assessments in the form of a meta-analyses could be performed. Hence, the findings of stages 1–3 have been reported here as the results of a SR.

**Figure 1** Overview of four stage process. CHARMS-PF, Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of prognostic factor studies; DPFs, diagnostic and prognostic factors; PROBAST, Prediction model Risk Of Bias Assessment Tool; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies 2; QUIPS, Quality in Prognostic Studies; SR, systematic review.

**Stage 1: comprehensive literature review**

We developed the search criteria for the first search with an information scientist who specialises in SR for urology. MEDLINE, Embase and the Cochrane Library were searched on 21 January 2020. The second search was developed following a consultation with an independent information scientist who excluded row 12, 14 and 16 of (see online supplemental table 1). We screened the EAU Guidelines reference list for PCa in our third search (see figure 2).

**Stage 2: multidisciplinary expert meeting**

On the 20 March 2020, we invited a group of multidisciplinary participants to discuss the identified articles on DPFs (see online supplemental table 2). The participants were presented the search criteria and the extracted data. Data extraction followed the CHARMS-PF checklist and we added author and year of publication.
Prior to the evaluation of the quality of studies, an initial pilot screening to prepare the raters for the use of PROBAST, QUADAS-2 and QUIPS was performed. This aimed to reach consensus on how to judge the domains of the assessments using the three RoB tools. Two urologists (FB and SS) and two epidemiologists (AH and KB) were involved in the pilot assessments. The group discussed any discrepancies. Articles which presented the development and validation either internal validation or external validation (i.e., the same data was used for both development and internal validation, such as bootstrapping or cross-validation; different populations were used for development and validation), of a diagnostic or prognostic model were assessed with PROBAST. Papers assessing single biomarkers or with/without validation were assessed with QUIPS for prognostic or QUADAS-2 for diagnostic biomarkers.

The RoB of diagnostic factors without validation or single validated factors was evaluated using QUADAS-2. We assessed the following four domains: patient selection, index test, reference standards and flow and timing. The first three domains are assessed looking at applicability and all four domains were assessed in terms of RoB.14 We created a summative score after the diagnostic studies were assessed by two reviewers and in case of disagreement a third reviewer assessed the study.

Evaluation of quality of studies published using PROBAST (diagnostic)
The RoB of internal or external validated diagnostic models was assessed using the PROBAST RoB tool. PROBAST includes four domains assessing the RoB (i.e., participants, predictors, outcome, and analysis) and four domains assessing applicability (i.e., participants, predictors, and outcome) (see online supplemental table 3 for scoring information).

Evaluation of quality of studies published using QUIPS
To assess the articles which are single factors or were not internally or externally validated, we used the QUIPS rating procedure (see online supplemental table 4 for scoring information). To standardise the approach across raters, we used the QUIPS electronic spreadsheet (excel) from Hayden et al.15 There are no rules available for QUIPS on how to score the overall RoB of a paper. Due to the large number of papers and the need for synthesis, we followed the suggestions from Grooten et al., and categorised on the following criteria: (1) Paper was classified as low RoB if all domains were classified as having low RoB, or up to one moderate RoB; (2) Paper was classified as high RoB if one or more domains were classified as having high RoB, or ≥ 3 moderate RoB; (3) Paper was classified as having moderate RoB if all papers in between 1 or 2 (see online supplemental table 1). This assessment was based on the risk scores of individual assessments within the group. If the overall assessment was not possible due to differences in the individual category, a third assessor reviewed the assessments and the results were discussed.

Evaluation of quality of studies published using PROBAST (prognostic)
The RoB of prognostic validated models were assessed using PROBAST. As highlighted above, PROBAST includes four domains assessing the RoB (i.e., participants, predictors, outcome and analysis) and the domains assessing applicability (i.e., participants, predictors and outcome).

RESULTS

Stage 1: comprehensive literature review
Stage 1 identified 6604 citations and contained three independent searches. After removing duplicates, we screened 4215 abstracts, from which 489 met the inclusion criteria.

Stage 2: multidisciplinary expert meeting
The group discussed the results and additional literature on DPFs was suggested to help the classification of the DPFs, such as the ASCO Guideline on Molecular Biomarkers in Localised Prostate Cancer.16

Stage 3: evaluation of quality of studies published using the RoB tools
The 489 articles were equally divided between six groups. The six groups received the guidance documents which
were identified during the pilot phase. In addition, MvH and KB discussed questions with each individual group.

Evaluation of quality of studies published using QUADAS-2

The RoB of the 41 included studies was low for 10 studies, high for 23 studies and unclear for eight. RoB concerning applicability was low for 10 studies, high for 21 studies and unclear for 10 studies (see table 1). Table 2 shows the studies with an overall low RoB across both categories. Two studies were identified to have an overall low RoB.

Evaluation of quality of studies published using PROBAST (diagnostic)

We identified 20 papers to be assessed with PROBAST. The RoB of three papers was low, high for 14 and was unclear for three. The applicability of eight papers was high and was unclear for two (see table 1). Online supplemental table 1 shows the criteria on how to judge the RoB. One study had an overall low RoB across both domains. All categories except ‘predictors’ was scored to have a low RoB. There was little information available for the category predictors and therefore it was scored as ‘unclear’ (see table 3).

Evaluation of quality of studies published using QUIPS

The 12 assessors independently inserted the relevant information and assessed each domain such as participation, attrition, prognostic factor confounding and statistical analysis and reporting.

A total of 387 prognostic factors were assessed using QUIPS. A total of 307 papers were classified as high RoB. Forty-nine papers were classified as having a moderate RoB and 28 papers were scored as low RoB (see table 1). Out of the 28 papers with a low RoB, the most common moderate bias was linked to attrition (12 papers), followed by confounding (4 papers), participation (3 papers), outcome (1 paper), statistical analysis (1 paper) (see table 4).

Evaluation of quality of studies published using PROBAST (prognostic)

The assessors identified 44 papers to be assessed with PROBAST, of those three scored a low RoB, 27 a high RoB and 13 were assessed as unclear (see table 1). In terms of applicability, 15 papers scored low, 20 high and eight unclear. Two papers were scored to have an overall low RoB.

Characteristics of studies identified with low RoB

Details of the identified validated DPF models with an adequate quality are presented in table 5. We identified 32 studies with an overall low RoB (assessed with PROBAST, QUIPS, QUADAS-2). Out of these 32 studies, we identified one validated diagnostic model (assessed with PROBAST), two validated prognostic models (assessed with PROBAST), two non-validated diagnostic single factors (assessed with QUADAS-2) and 26 prognostic factors (assessed with QUIPS) which have not been validated and two single prognostic factors which have been validated (assessed with QUIPS). Prognostic factors assessed with QUIPS were identified with a low RoB for the localised PCa population. Sixty-seven per cent of the low RoB DPFs were intended to be measured after the treatment was performed. In addition,
the most commonly measured outcome was biochemical recurrence followed by overall survival. However, it is important to take into consideration that even from the studies assessed with a low RoB, only 2 out of the 32 were of a non-observational study design.

As highlighted above, we identified three validated DPFs which were scored to have a low RoB and low risk concerning applicability. First, we identified the ‘Unified Prostate Cancer Risk Prediction Model Combining the Stockholm 3 Test and MRI’, a risk prediction model which combines clinical variables, genetic and protein biomarkers. Five hundred and thirty-two men were involved across three centres. Second, the DREAM challenge developed a set of five standardised raw event-level tables, using laboratory values, patients’ demographic information, medical history, lesion sites, previous treatments and vital signs of patients with mCRPC. These variables where combined by using data from four clinical trials. Third, Joniau et al., developed ‘Pretreatment Tables’ to predict the pathologic stage of locally advanced PCa after RP based on pretreatment PSA level and biopsy Gleason score.

We identified two single factors which were validated and had low RoB. First, Lara et al., assessed and validated the serum biomarkers of bone metabolism (N-telopeptide and pyridinoline) and formation (C-terminal collagen propeptide and bone alkaline phosphatase) in 778 CRPC patients as part of the randomised phase III SWOG trial (S0421) of docetaxel/prednisone with or without atrasentan. Second, Berg et al., showed that ERG expression can be used to estimate the risk of progression during AS including 265 patients at diagnosis and progression during AS.

**DISCUSSION**

Despite the large number of studies on DPFs which are published every year, there is a paucity of DPFs that are suitable to be incorporated into clinical practice. The majority of DPFs have not yet been validated and are identified in poor quality studies. Our analysis found that most identified studies had a high to moderate RoB due to poor design standards, conduct, reporting and/or analysis that is, generalisability and size of the population, poor model development (no testing or missing important confounders) or only correlation studies, missing data was rarely reported. However, we did identify a small number of validated DPFs with low RoB. We identified three validated models which combine: first, clinical variables, genetic and protein biomarkers, and improved clinical outcome performance of PCa diagnostics (The Unified Prostate Cancer Risk Prediction Model); second, laboratory values, patients’ demographic information, medical history, lesion sites, previous treatments and vital signs of patients with metastatic castration-resistant PCa (DREAM challenge); and third, pretreatment PSA level and biopsy Gleason score to predict the pathologic stage of locally advanced PCas (‘Pretreatment Tables’).

Two single factors have been validated: the serum biomarkers of bone metabolism in CRPC patients and the ERG expression, which can be used to estimate the risk of progression during AS, which has already been highlighted in the clinical guidelines.

Aladawani et al., assessed prediction models for PCa to be used in primary care settings in their SR and identified five models which met their inclusion criteria. From these identified models only one model was externally validated and only one (the Lazzari model) had the potential to be implemented in primary care. Lazzari et al., had the lowest RoB (based on PROBAST); however, it must be externally validated before it can be implemented. Hence, Aladawani et al., also concluded that the existing models have limitations concerning study design and reporting performance.

Tian et al., conducted a review on biomarkers for CRPC patients, however, their quality assessment was focused on study design (RCT vs. observational study), whereas we focused on biomarker specific tools. While Tian et al., and our review identified similar factors and quality scores, there were slight discrepancies between the overall RoB assessments. Tian et al., used an overall quality assessment scale from 1 to 6 instead of low, medium and high. In their assessment the validated prognostic study by Lara et al., and the non-validated prognostic factor by Pei et al., were scored on the quality scale as 4 (medium quality). We assessed Lara et al., to have a low RoB with a moderate risk of confounding and Pei et al., with a moderate RoB concerning the prognostic factor itself. This might explain the discrepancies between the two.
<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Population</th>
<th>Study design</th>
<th>Timing</th>
<th>Index</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palsdottir</td>
<td>2019</td>
<td>Localised PCa</td>
<td>Observational study</td>
<td>Pre treatment</td>
<td>S3M-MRI (Stockholm3 +PI-RADS)</td>
<td>csPCa diagnosis</td>
</tr>
<tr>
<td>Guinney</td>
<td>2017</td>
<td>mCRPC</td>
<td>Prog. PROBAST</td>
<td>Post treatment</td>
<td>ePCR model</td>
<td>OS</td>
</tr>
<tr>
<td>Joniau</td>
<td>2017</td>
<td>Locally advanced PCa</td>
<td>Prog. PROBAST</td>
<td>Post-treatment</td>
<td>Gleason score +PSA</td>
<td>Adverse pathological features at RP; LNI</td>
</tr>
<tr>
<td>Hagiwara</td>
<td>2017</td>
<td>Localised PCa</td>
<td>QUADAS</td>
<td>Pre-treatment</td>
<td>WFA-reactive glycan-carrying PSA-Gi</td>
<td>PCa diagnosis, PSA-free survival</td>
</tr>
<tr>
<td>Kelly</td>
<td>2017</td>
<td>Localised PCa</td>
<td>QUADAS</td>
<td>Pre-treatment</td>
<td>mIR-141, −145, −155, let7a</td>
<td>PCa diagnosis</td>
</tr>
<tr>
<td>Aguilera</td>
<td>2015</td>
<td>High risk PCa</td>
<td>Observational study</td>
<td>Pre and post treatment</td>
<td>Age, rectal examination, PSA, biopsy Gleason score, uni/bilateral tumour, affected cylinder percentage and postoperative</td>
<td>BCR</td>
</tr>
<tr>
<td>Alvim</td>
<td>2019</td>
<td>Metastatic PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>PSA response (PSA reduction ≥50%)</td>
<td>OS, PFS</td>
</tr>
<tr>
<td>Bramhecha</td>
<td>2019</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>PTEN deletion</td>
<td>BCR</td>
</tr>
<tr>
<td>Bruce</td>
<td>2016</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>AZGP1 expression</td>
<td>BR-free survival, CR-free survival, PC-specific death</td>
</tr>
<tr>
<td>Francini</td>
<td>2018</td>
<td>mHSPC</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>Volume</td>
<td>OS, time to CRPC</td>
</tr>
<tr>
<td>Hamada</td>
<td>2016</td>
<td>High risk PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>PSA, PSA density (PSAD), PSAD of the transition zone, percentage of positive cores (PPC), prostate volume, TZ volume, Gleason score, PPC from the dominant side</td>
<td>BCR</td>
</tr>
<tr>
<td>Hashimoto</td>
<td>2020</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>Micro-lymphatic invasion, Gleason</td>
<td>BCR</td>
</tr>
<tr>
<td>Hung</td>
<td>2017</td>
<td>mCRPC</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>Neurovascular bundle preservation, blood loss, pT stage, pN stage, pGS, PNI, angiolymphatic invasion, tumour amount in specimen, ECE, PSM, SW, Bladder neck invasion, Foley duration, post-op undetectable PSA</td>
<td>BCR</td>
</tr>
<tr>
<td>Kato</td>
<td>2018</td>
<td>High risk PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>LC/IDC</td>
<td>PFS, CSS</td>
</tr>
<tr>
<td>Kluth</td>
<td>2014</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post-treatment</td>
<td>No of lymph nodes</td>
<td>BCR</td>
</tr>
<tr>
<td>Lara</td>
<td>2014</td>
<td>mCRPC</td>
<td>QUIPS</td>
<td>RCT</td>
<td>Post treatment</td>
<td>Bone resorption and formation</td>
</tr>
<tr>
<td>Lee</td>
<td>2016</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post treatment</td>
<td>Positive surgical margin status and bilateral seminal vesicle invasion</td>
<td>BCR</td>
</tr>
<tr>
<td>Lévesque</td>
<td>2019</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post treatment</td>
<td>UGT2B17 expression</td>
<td>BCR</td>
</tr>
<tr>
<td>Lin</td>
<td>2017</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post treatment</td>
<td>Aberrant Promoter Methylation of Protocadherin8 (PCDH8)</td>
<td>BCR-free survival</td>
</tr>
<tr>
<td>Löffeler</td>
<td>2015</td>
<td>mCRPC</td>
<td>QUIPS</td>
<td>Anytime</td>
<td>PSA doubling time, PSA nadir during ADT, haemoglobin and alkaline phosphatase levels at CRPC</td>
<td>OS</td>
</tr>
<tr>
<td>Narang</td>
<td>2017</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Anytime</td>
<td>PSA: End-of-radiation PSA</td>
<td>BCR-free survival, MFS, CSS, OS</td>
</tr>
<tr>
<td>Ozden</td>
<td>2017</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post treatment</td>
<td>Age</td>
<td>RRP specimen, BCR, and BCR-free survival rates</td>
</tr>
<tr>
<td>Pei</td>
<td>2016</td>
<td>CRPC</td>
<td>QUIPS</td>
<td>Pre and during treatment</td>
<td>Neutrophil-to-lymphocyte ratio</td>
<td>OS, PFS</td>
</tr>
<tr>
<td>Qu</td>
<td>2016</td>
<td>mPCa and CRPC</td>
<td>QUIPS</td>
<td>Post treatment</td>
<td>AR-V7</td>
<td>Time to CRPC / CRPC: CSS</td>
</tr>
<tr>
<td>Qu</td>
<td>2017</td>
<td>PCa</td>
<td>QUIPS</td>
<td>Pre and during treatment</td>
<td>AR-V7</td>
<td>OS</td>
</tr>
<tr>
<td>Rüenauer</td>
<td>2014</td>
<td>Localised PCa</td>
<td>QUIPS</td>
<td>Post treatment</td>
<td>YWHAZ</td>
<td>OS</td>
</tr>
</tbody>
</table>

Continued
quality assessments. The reports by Alvim et al., Qu et al.,
were assessed to have the highest quality by Tian et al.,
similar to our review. This illustrates that different quality
assessment tools emphasise different criteria, which may
result in small discrepancies. However, the overall conclu-
sion for prognostic single factors was similar in our review
and to the work of Tian et al.,

Similar issues have been identified for other urological
cancers. For example, in kidney cancer, a large body of
research was identified by Harrison et al., with very few vali-
dated studies and lots of heterogeneity. Schmitz-Dräger
et al., published an International Consultation of Urologic
Disease/WHO Consensus manuscript where they identified
that in bladder cancer one of the main limitations
for the lack of incorporation of modern bladder cancer
tests into clinical practice decision making is linked to the
scarcity of ‘good clinical practice guidelines’ for the eval-
uation of diagnostic markers.

There is a need for improved guidance on develop-
ment and validation of diagnostic markers. To meet that
need, we are developing the PIONEER DPF search tool,
which will help researchers and clinicians to get a better
understanding of the DPFs for PCa. The tool will not only
summarise all relevant studies, but also provide informa-
tion on the use and results of different RoB assessment
tools, which will enable an understanding of the quality
of published studies.

Future research should, therefore, focus on addressing
the identified shortcomings such as heterogeneity, vali-
dation and poor RoB by designing more robust studies
which consistently include RoB assessments such as
PROBAST, QUIPS or QUADAS-2.

With the growing number of various therapeutic
options, diagnosis and management of PCa requires
an individualised approach to patient care. There is an
unmet need for DPFs to guide decisions for optimal
treatment and to predict which patients will benefit the
most, from a particular management strategy. DPFs could
potentially enhance the quality of patient counselling, but
currently most need additional evaluation and validation
in properly designed studies. Our SR highlights the need
for well-designed Real-World Evidence studies, while the
PIONEER online search tool can inform the design of
new research studies, through providing a rigorous eval-
uation of the methodological quality of the studies.

The main strength of this study are the extensive and
comprehensive search and screening of the studies
included. In addition, we are developing an online search
tool which showcases the identified and assessed studies.
It provides an overview of the available DPFs and enables
interested stakeholders to search for DPFs. To our knowl-
dge, this is the first study which has been performed with
this extensive amount of literature.

**Patient and public involvement**

This project has been overseen by a multistakeholder
group part of the PIONEER Consortium. PIONEER
brings together 35 key stakeholders from academic insti-
tutions, patient advocacy groups, European organisa-
tions, experts in legal data management, clinicians and
pharmaceutical companies, as well as regulatory agencies,
economics and ethics, and information and technology
specialists. Patients and their family members are there-
fore involved and actively participate as an integral part
of all research conducted by the PIONEER Consortium.

**Limitations**

Even though this review included three searches and
assessments by a multidisciplinary group of fourteen
researchers, we recognise potential limitations. Studies
were only included from 2014 onwards and DPFs devel-
oped before 2014 were not included. However, signif-
ificant changes which influence the staging of PCa (i.e.,
Consensus Conference on Gleason Grading of Prostatic
Carcinoma) have taken place in diagnostic and prog-
nostic practice and patient management. This changed

---

**Table 4 Continued**

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>RoB</th>
<th>Population</th>
<th>Study design</th>
<th>Timing</th>
<th>Index</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shimodaira</td>
<td>2020</td>
<td>QUIPS</td>
<td>Metastatic PCa</td>
<td>Observational study</td>
<td>Post treatment</td>
<td>Value of platelet counts</td>
<td>Disease specific survival</td>
</tr>
<tr>
<td>Strand</td>
<td>2015</td>
<td>QUIPS</td>
<td>Localised PCa</td>
<td>Observational study</td>
<td>Post treatment</td>
<td>5-hydroxymethylcytosine score</td>
<td>BCR</td>
</tr>
<tr>
<td>Takagi</td>
<td>2017</td>
<td>QUIPS</td>
<td>Localised PCa</td>
<td>Observational study</td>
<td>Post treatment</td>
<td>Age, T stage, % of pos cores, Gleason score, PSA, Total ADT</td>
<td>BCR-free survival</td>
</tr>
<tr>
<td>Wang</td>
<td>2016</td>
<td>QUIPS</td>
<td>PCa</td>
<td>Observational study</td>
<td>Post treatment</td>
<td>Platelet to lymphocyte ratio (PLR)</td>
<td>PLR with PFS, CSS and OS n/a</td>
</tr>
<tr>
<td>Zacho</td>
<td>2017</td>
<td>QUIPS</td>
<td>Localised PCa</td>
<td>Observational study</td>
<td>Anytime</td>
<td>Bone scan index</td>
<td>Time to CRPC</td>
</tr>
<tr>
<td>Berg</td>
<td>2014</td>
<td>QUIPS validated</td>
<td>Under Active Surveillance</td>
<td>Observational study</td>
<td></td>
<td>ERG immunohisto-chemical staining</td>
<td>Overall AS progression, histopathologic progression</td>
</tr>
</tbody>
</table>

ADT, androgen deprivation therapy; AS, Active Surveillance; BCR, biochemical recurrence; CSS, cancer-specific survival; DPFs, diagnostic and prognostic factors; mCRPC, metastatic castration resistant prostate cancer; n/a, not available; OS, overall survival; PCa, prostate cancer; PFS, progression-free survival; PI-RADS, Prostate Imaging Reporting and Data System; PROBAST, Prediction model Risk Of Bias Assessment Tool; PSA, Prostate Specific Antigen; PTEN, Phosphatase and tensin homolog; QUADAS, Quality Assessment of Diagnostic Accuracy Studies; QUIPS, Quality in Prognostic Studies; RCT, Randomised control trial; RoB, risk of bias; WFA, Wisteria floribunda agglutinin.
In addition, there is a potential of subjectivity in the evaluation of the studies. Even though the studies have been assessed in duplicate, there might be variation across groups. However, given the overall moderate to high RoB, this does not influence the overall recommendation of the project.

**CONCLUSION**

At present DPFs that are capable of significantly improving diagnosis and prognosis in PCa are an unmet need as most of the DPFs identified require additional evaluation and validation in properly designed studies before they can be recommended for use in clinical practice. Well-designed real world evidence (RWE) studies can help to increase quality. Our SR aims to inform clinicians and patients about this rapidly evolving field, while the PIONEER online search tool for DPFs for PCa will enable researchers to perform future research, and to understand the quality of the current available studies.

**Author affiliations**

1. Translational and Oncology Research (TOUR), King’s College London, Faculty of Life Sciences and Medicine, London, UK
2. Department of Urology, University Hospitals Leuven, Leuven, Belgium
3. Department of Urology, Metropolitan Hospital, Athens, Greece
4. Unit of Urology, IRCCS Ospedale San Raffaele, Milan, Italy
5. Bayer AG, Berlin, Germany
6. Department of Urology, Mid Cheshire Hospitals, NHS Foundation Trust, Crewe, UK
7. Department of Minimally Invasive and Robotic Urologic Surgery and Kidney Transplantation, University of Florence, Florence, Italy

**Table 5**

<table>
<thead>
<tr>
<th>Study</th>
<th>Time</th>
<th>Participation</th>
<th>Attrition</th>
<th>Prognostic factor</th>
<th>Outcome</th>
<th>Confounding</th>
<th>Statistical analysis and reporting</th>
<th>Overall score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aguilera25</td>
<td>2015</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Alvim27</td>
<td>2019</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Bramhecha26</td>
<td>2019</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Bruce28</td>
<td>2016</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Francini29</td>
<td>2018</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Hamada30</td>
<td>2016</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Hashimoto31</td>
<td>2020</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Hung57</td>
<td>2017</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Kato32</td>
<td>2018</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Kluth33</td>
<td>2014</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Lara44</td>
<td>2014</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Lee35</td>
<td>2016</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Levesque36</td>
<td>2019</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Lin37</td>
<td>2017</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Loffler38</td>
<td>2015</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Narang39</td>
<td>2017</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Ozden40</td>
<td>2017</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Pei41</td>
<td>2016</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Qu42</td>
<td>2016</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Qu43</td>
<td>2017</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Rizzare45</td>
<td>2015</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Ruenauer46</td>
<td>2014</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Shimodaira47</td>
<td>2020</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Strand48</td>
<td>2015</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Takagi49</td>
<td>2017</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Wang49</td>
<td>2016</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Zacho49</td>
<td>2017</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Berg50</td>
<td>2014</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

DPFs, diagnostic and prognostic factors; QUIPS, Quality in Prognostic Studies.
Department of Urology, Erasmus University Medical Center, Rotterdam, The Netherlands

Academic Urology Unit, University of Aberdeen, Aberdeen, UK

Department of Urology, University Vita e Salute-San Raffaele, Milan, Italy

Department of Translational Medicine, Division of Urological Cancers, Lund University, Sweden

EORTC Headquarters, Brussels, Belgium

Sanofi, Paris, France

Global Medical Affairs Oncology, Real World Evidence, Bayer HealthCare

Pharmaceuticals Inc, Whippany, New Jersey, USA

ORCID iDs

Katharina Beyer http://orcid.org/0000-0002-8450-8850

Riccardo Campi @Ric_Campi, Kirill Shiranov @endourologist, Muhammad Imran Omar @dirnomanar, Alex Aslimwe @aleohabs and Sara J MacIennan @SajMacIennan

Acknowledgements

PIONEER Consortium

Contributors

KB, LM, ML, AH, FB, SS, MM, MIO, SM, MJR, BF, EV, ZD, AA, JZ, SJM, LC, JN, Abr, ABj and MvH conceptualised designed the review. Abstracts and full texts were reviewed and data extracted by KB, LM, ML, AH, FB, SS, MM, RH, AR, RC, IG, KS, TvDb and SA. Authors resolved disagreement by discussion where necessary. The risk of bias was assessed by KB, LM, ML, AH, FB, SS, MM, RH, AR, RC, IG, KS and SD. The manuscript was drafted by KB, LM, ML, Abr, MvH and reviewed by KB, LM, ML, AH, FB, SS, MM, RH, AR, RC, IG, KS, SD, TvDb, MG, GG, Mio, SM, MJR, BF, EV, ZD, AA, JZ, SJM, LC, JN, Abr, ABj and MvH. The project was supervised and guided by JZ, SJM, LC, JN, Abr, ABj and MvH. PIONEER Consortium acts as a guarantor.

Funding

PIONEER is funded through the IM2 Joint Undertaking and is listed under Grant Agreement No. 774792 and is part of the Big Data for Better Outcomes Programme (BD4BO). IM2 receives support from the European Union’s Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations (EFPIA). The views communicated within are those of PIONEER. Neither the IM2 nor the European Union, EFPIA, or any Associated Partners are responsible for any use that may be made of the information contained herein.

Competing interests

None declared.

Patient consent for publication

Not applicable.

Ethics approval

This study does not involve human participants.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data availability statement

No data are available. not applicable.

Supplemental material

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs

Katharina Beyer http://orcid.org/0000-0002-8450-8850

Francesco Barletta http://orcid.org/0000-0003-1265-3217

Megan Molnar http://orcid.org/0000-0003-4537-2987

Riccardo Campi http://orcid.org/0000-0001-5237-0888

Muhammad Imran Omar http://orcid.org/0000-0002-1597-3126

Steven MacIennan http://orcid.org/0000-0002-2691-8421

REFERENCES


