THE PROGNOSTIC VALUE OF BIOCHEMICAL RECURRENCE FOLLOWING TREATMENT WITH CURATIVE INTENT FOR PROSTATE CANCER: A SYSTEMATIC REVIEW

Van den Broeck T*1,2, van den Bergh RCN*3, Arfi N*4, Gross T5, Moris L1,2, Briers E6, Cumberbatch M7, De Santis M8,9, Fossati N10,11, Gillessen S12,13, Grummet J14, Henry A15, Lardas M16, Liew M17, Rouvière O18, Pecanka J19, Mason M20, Schoots IG21, Van Der Kwast T22, Van Der Poel H23, Wiegel T24, Willemsen PP25, Yuan CY26, Lam TB27, Cornford P28, Mottet N29.

*Shared first authorship

This systematic review was performed under the auspices of:
- The European Association of Urology Guidelines Office Board
- The European Association of Urology Prostate Cancer Guideline Panel

1 Department of Urology, University Hospitals Leuven, Leuven, Belgium.
2 Laboratory of Molecular Endocrinology, KU Leuven, Leuven, Belgium.
3 Department of Urology, Netherlands Cancer Institute, Amsterdam, The Netherlands.
4 Department of Urology, Hospital Saint Luc Saint Joseph, Lyon, France.
5 Department of Urology, University of Bern, Inselspital, Bern, Switzerland.
6 Hasselt, Belgium.
7 Academic Urology Unit, University of Sheffield, Sheffield, UK.
8 Clinical Science Unit, University of Northumbria at Newcastle, United Kingdom.
9 Department of Urology, Medical University of Vienna, Austria.
10 Unit of Urology/Division of Oncology, URI, IRCCS Ospedale San Raffaele, Milan, Italy.
11 Università Vita-Salute San Raffaele, Milan, Italy.
12 Department of Oncology and Haematology, Cantonal Hospital St Gallen, St Gallen, Switzerland.
13 Department of Medical Oncology, Inselspital, Bern University Hospital, University of Bern, Berne, Switzerland.
14 Department of Surgery, Central Clinical School, Monash University, 125 Balaclava Road, Caulfield North, Victoria 3161, Australia.
15 Leeds Cancer Centre, St. James's University Hospital and University of Leeds, Leeds, UK.
16 Department of Urology, Leto Hospital, Athens, Greece.
17 Department of Urology, Wrightington, Wigan and Leigh NHS Foundation Trust, Wigan, UK.
18 Hospices Civils de Lyon, Radiology Department, Edouard Herriot Hospital, Lyon, France.
19 Prague, Czech Republic
20 Division of Cancer & Genetics, School of Medicine Cardiff University, Velindre Cancer Centre, Cardiff, UK.
21 Department of Radiology & Nuclear Medicine, Erasmus MC University Medical Center, Rotterdam, The Netherlands.
22 Department of Pathology, Erasmus Medical Centre, Rotterdam, The Netherlands.
23 Department of Urology, Netherlands Cancer Institute, Amsterdam, The Netherlands.
24 Department of Radiation Oncology, University Hospital Ulm, Ulm, Germany.
25 Department of Urology, University Hospital Groningen, Groningen, The Netherlands.
26 Department of Medicine, McMaster University, Hamilton, ON, Canada.
27 Academic Urology Unit, University of Aberdeen, Aberdeen, UK; Department of Urology, Aberdeen Royal Infirmary, Aberdeen, UK.
28 Royal Liverpool and Broadgreen Hospitals NHS Trust, Liverpool, UK.
29 Department of Urology, University Hospital, St. Etienne, France.
ABSTRACT

Context. In men with prostate cancer (PCa) treated with curative intent, controversy exists regarding the impact of biochemical recurrence (BCR) on oncological patient outcomes.

Objective. To perform a systematic review of the existing literature on BCR after treatment with curative intent for non-metastatic PCa. Objective 1 is to investigate whether oncological outcomes differ between patients with or without BCR. Objective 2 is to study which clinical factors and tumor features in patients with BCR have an independent prognostic impact on oncological outcomes.

Evidence acquisition. Medline, Medline In-Process, Embase, and the Cochrane Central Register of Controlled Trials were searched. For objective 1, prospective and retrospective studies comparing survival outcomes of patients with or without BCR following radical prostatectomy (RP) or radical radiotherapy (RT) were included. For objective 2, all studies with at least 100 participants and reporting on prognostic features in patients with BCR were included. Risk-of-bias and confounding assessments were performed according to the Quality in Prognosis Studies (QUIPS) tool. Both a narrative synthesis and meta-analysis were undertaken.

Evidence synthesis. Overall, 77 studies were included for analysis, of which 14 studies addressed objective 1, recruiting 20406 patients. Objective 2 was addressed by 71 studies with 29057, 11301 and 4272 patients undergoing RP, RT or a mixed population (mix of patients undergoing RP or RT as primary treatment) respectively. There was low risk of bias for study participation, confounders and statistical analysis. For most studies, attrition bias, prognostic and outcome measurements were not clearly reported. BCR was associated with worse survival rates, mainly in patients with a short PSA Doubling Time (PSA-DT) and high final Gleason score after RP or a short Interval to Biochemical Failure (IBF) after RT and high biopsy Gleason score.

Conclusion. BCR has an impact on survival, but this effect appears to be limited to a subgroup of patients with specific clinical risk factors. A short PSA-DT and high final Gleason score after RP and a short IBF after RT and high biopsy Gleason score are the main factors which have a negative impact on survival.

Patient summary. This review looks at the risk of dying in men who have a rising PSA blood test after curative surgery or radiotherapy. For many men a rising PSA does not mean they are at a higher risk of dying from prostate cancer in the longer term. Men with a PSA that rises shortly after they were treated with radiotherapy or a rapidly rising PSA after surgery and a high tumor-grade for both treatment modalities are at the highest risk of dying.
INTRODUCTION

Patients with non-metastatic prostate cancer (PCa) who are fit enough to receive curative treatment generally have a good prognosis. This was particularly true in historical series when many patients with low risk cancer were included [1–3]. After treatment with curative intent, the most sensitive and the only validated biomarker for disease recurrence remains prostate specific antigen (PSA) measurement [4,5]. Biochemical recurrence (BCR) defined as the return of a measurable PSA occurs in about 35% of patients undergoing treatment for localized disease [11]. However, biochemical recurrence (BCR)-based on PSA measurements does not necessarily indicate that an individual will develop clinically relevant recurrence and/or die due to his disease, with studies reporting that only + 30% of patients with BCR after primary surgery develop clinical recurrence [6] with only 16.4% dying from their disease [7]. Furthermore, for patients with BCR after radical prostatectomy (RP) or primary radiotherapy (RT), several authors have reported that only certain patient subgroups with poor tumor differentiation and PSA kinetics are at high-risk of progressive disease [8–10]. Understanding the true impact of BCR on oncological outcomes is crucial because first of all it occurs frequently, in about 35% of patients who have undergone treatment for localized PCa [11]. Secondly, when BCR occurs clinicians need to counsel decide whether or not their patients as to whom would benefit from potentially toxic salvage therapies. The objectives of this systematic review are to determine whether BCR is associated with oncological outcomes (**objective 1**) and to determine which clinical factors have independent prognostic impact for oncological outcomes after BCR has occurred (**objective 2**).

EVIDENCE ACQUISITION

The review was commissioned and undertaken by the EAU Prostate Cancer Guideline Panel as part of its guideline update for 2019. The protocol for this review has been published (http://www.crd.york.ac.uk/PROSPERO; registration number CRD42015026807). The review was performed according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines [12] and Cochrane review principles [13] (Supplementary File S1). **PICO development was performed by the EAU PCa panel, based on the expertise of urologists, oncologists, radiotherapists, radiologists, pathologists and a patient representative.** English language articles published from January 2000 to July 2018 were included. The year 2000 was used as cutoff due to the need for sufficiently long follow-up (at
least 10 years) after PSA was introduced (at the end of the 1980s). All abstracts and resulting full-text articles were independently screened in duplicate (TVDB, RVDB, NA, TG, LM) and disagreement was resolved by discussion or reference to an independent third party (NM, TL).

To investigate the research objectives on the impact of BCR on oncological outcomes (objective 1) and prognostic clinical factors in patients experiencing BCR (objective 2), studies that investigated the clinical impact of BCR after RP or RT were included. Due to the expected heterogeneity of BCR definitions used in different studies, all studies investigating the above objectives were included, irrespective of the BCR definition that was used. The study population was limited to men with histologically proven non-metastatic PCa who underwent treatment with curative intent and a minimal median follow-up of 5 years post-BCR. The primary outcome was overall mortality (OM) and secondary outcomes were prostate cancer-specific mortality (PCSM) and development of distant metastases (DM). Risk of bias (RoB) was assessed using the Quality In Prognosis Studies (QUIPS) tool [14] as recommended by the Cochrane Prognosis Methods Group. Double data extraction was performed independently in duplicate relating to the pre-specified outcomes. Sub-group analysis was pre-planned based on the following variables: T/N stage category, primary treatment, disease risk classification, GS, characteristics of BCR (PSA-DT, IBF), Surgical margin status, Salvage treatment, Extent of lymphadenectomy for RP.

A meta-analysis was performed regarding the risk effects associated with reported patient and tumor variables for three different end-points: DM, PCSM and OM. This was performed separately for patients treated with RP or RT. The meta-analysis of the observed risk effects was performed using the random-effect linear regression model to account for the expected clinical heterogeneity across studies. Since various studies included multiple co-dependent variables into one multivariable analysis (MVA) (e.g. GS 7 vs 6 and 8-10 vs 6), a study grouping variable was included as a random effect. Sensitivity analysis was performed by comparing the results of the primary analysis with the results of: a) a fixed-effect model; b) a linear regression model excluding the random effect; and c) a conservative strategy to test the primary analysis’ sensitivity. Potential publication biases were investigated using visual inspection of Funnel plots and calculated using the rank test and regression test (Egger’s test).

For full methodological information, consult Supplementary File S2.

EVIDENCE SYNTHESIS
Quantity of evidence identified

The study selection process is outlined in the PRISMA flow diagram (Fig. 1). In total, 10863 records were identified through database searching, and 6759 were screened after removal of duplicates. Of these, 381 articles were eligible for full-text screening. Finally, 77 studies met the inclusion criteria (34 and 71 studies dealing with oncological outcomes (objective 1) and prognostic factors (objective 2), respectively) [7–9,15–88].

Characteristics of the included studies

Table 1 presents the baseline study characteristics of included studies dealing with objective 1, recruiting a total of 20406 patients (one study did not report number of included patients). Supplementary Tables S3-S5 present the baseline study characteristics of included studies dealing with objective 2. Table S3 refers to studies reporting on patients who underwent RP (29057 patients); Table S4 refers to studies reporting on patients who underwent RT (11301 patients; 2 studies did not report the number of included patients); and Table S5 refers to a mixed population (4272 patients). For objective 1, 4, 7 and 10 studies investigated DM, PCSM and OM respectively. For objective 2, 35, 47 and 28 studies correlated prognostic factors with DM, PCSM and OM respectively. Due to the need of long term follow-up for the reported outcomes, most RT studies reported on a population with mixed RT doses. Of all included RT studies, only 5 studies reported on a patient population with a median RT dose of ≥74 Gy [15,62,65,70,77].

Risk of bias and confounding assessment of the included studies

Figure 2 summarizes the QUIPS-based RoB assessment of all included studies. Overall, there was low RoB for the domains of study participation, study confounders and statistical analysis. Most studies also had low RoB on reporting on patient characteristics and corrected for confounding factors, although these confounders were not pre-specified in any a priori protocol. Attrition bias was unclear in most studies due to lack of reporting of patient drop-outs. Prognostic factor and outcome measurements were unclear in a significant proportion of included studies mainly due to lack of reporting on the methods of data collection.

Results of evidence synthesis

Impact of BCR on oncological outcomes (Objective 1)
Fourteen studies directly compared patients with BCR vs no BCR. All studies found BCR to be an independent risk factor for the development of DM, PCSM and to a lesser extent OM (Table 2).

Three studies directly compared OM of patients with or without BCR after RP. Choueiri et al investigated 3071 patients (of whom 454 died) with a median follow-up of 7 (IQR: 3.1 – 12.0) years after primary treatment and concluded that the impact of BCR on OM rates was significant, but with a small effect size (HR 1.03; 95% CI 1.004-1.06) [57]. The second study by Jackson WC et al found BCR to have a larger effect size on OM (HR 2.32; 95% CI 1.45-3.71) with a median follow-up of 8 years [35]. Finally, the study by Fryczkowski et al remained inconclusive as well due to the limited number of events with only 11 and 20 reported cancer-specific and overall deaths out of 247 patients [46].

Studies reporting on patients who were treated with primary RT overall stated that overall BCR is an independent risk factor for the development of DM, PCSM and again to a lesser extent OM [15–19,21,79,87,88]. For cancer-related outcomes, all four studies investigating DM or PCSM as outcome showed a significant negative impact of BCR with varying degree of impact. In contrast, the impact of BCR on OM is less clear, but still five out of eight studies report an impact of BCR on OM rates [16,17,79,87,88] while the other three studies do not show this effect [15,18,19].

Not only is the impact of BCR on (cancer related) survival outcomes controversial, the PSA-based definition of BCR itself is also. Stephenson et al investigated different definitions of BCR after RP as primary treatment to determine which PSA cutoff and kinetics would best define which patients would eventually develop clinical progression. They suggested that a progressively rising PSA of at least 0.4 ng/mL is most strongly associated with metastatic progression [68] compared to definitions based on one PSA measurement (one PSA of at least 0.2, 0.4 or 0.6 for example) or PSA measurement based solely on their kinetics (2 or 3 consecutive rises for example). For RT, Abramowitz et al compared the impact of BCR based on which definition was used (ASTRO or Phoenix definition) and showed that the Phoenix definition was much more predictive for DM, CSM and OM compared to the ASTRO definition with a HR of 173 (95% CI 74-404) vs 62.7 (95% CI 27-143) for DM, 308 (95% CI 38-2483) vs 26.0 (95% CI 8.5-79) for CSM and 2.0 (95% CI 1.6-2.6) vs 1.0 (95% CI 0.8-1.3) for OM [16].

Prognostic factors for oncological outcomes in patients who develop BCR (Objective 2)

Commented [PC3]: This is true although if I was a reviewer I might argue the reason there is no correlation between BCR and risk of prostate cancer survival is because the wrong definition of BCR is being used. Only the 0.4ng/mL cut off is associated with development of metastatic disease after RP. What about considering replacing this sentence with “Part of the disparity between studies relates to the definition of BCR used?”
All investigated prognostic factors reported in each individual study are summarized in Supplementary tables S6 (after primary RP), S7 (after primary RT) and S8 (in mixed populations). Below we summarize prognostic factors that are reported most frequently in the included studies.

**Age**

For PCa specific outcomes (DM and PCSM) and irrespective of the primary PCa treatment (RP or RT), age is not a prognostic factor with HRs ranging from 0.98 (95% CI 0.95-1.01) to 1.02 (95% CI 0.99-1.04) for development of DM and 0.96 (95% CI 0.91-1.00) to 1.05 (95% CI 1.01-1.08) for PCSM with only two studies showing a biologically small, but statistically significant unfavorable effect of increasing age [7,31]. Obviously, age is a significant risk factor for OM, both in patients undergoing RP and RT. In the former group, HRs range from 1.04 (95% CI 1.02-1.07) to 1.79 (95% CI 1.48-2.17) and in the latter group, HRs range from 1.05 (95% CI 1.03-1.07) to 2.4 (95% CI 1.6-3.5) [15–17,22,42,44,50,57,64,65,70–72].

**Initial PSA**

Initial serum PSA (iPSA) levels prior to primary treatment were not uniformly found as a prognostic factor for any of the investigated outcomes. For patients undergoing RP, 2 out of 6 reported studies show a significant correlation with DM [8,26–28,43,54]. For RT, 3 out of 5 [16,17,61,62,72] and 2 out of 5 [16,17,61,62,64] studies showed a significant correlation with PCSM and OM respectively.

**Gleason Score**

For patients undergoing primary RP, a higher Gleason score identified on prostatectomy histology report (pGS) is a strong prognostic factor for oncological outcomes in most studies. For DM and PCSM, 12 out of 14 [7,8,22,23,26–28,38,39,41,43,52,53,56] and 10 out of 13 [7,22,25,31,33,36,38–40,48–50,52] studies showed a positive association with HR ranging from 1.2 (95% CI 1.1-1.5) to 14.4 (95% CI 4.3-48.8) and 1.35 (95% CI 1.07-1.71) to 10.8 (95% CI 3.1-37.9). Similarly, for patients with RT as primary treatment, 8 out of 8 [16,62,65,70,71,73,75,77], 5 out of 8 [15–17,59,61,62,65,72] and 7 out of 9 [15–17,58,59,61,62,64,71] studies showed a positive association between a higher Gleason score identified on prostate biopsies (bGS) and DM, PCSM and OM respectively with HRs ranging from 1.7 (95% CI 1.1-2.7) to 3.7 (95% CI 1.4-10.3), 2.11 (95% CI 1.03-4.34) to 14.8 (95% CI 2-110) and 1.8 (95% CI 1.3-2.4) to 17.9 (95% CI 9.6-33) respectively. Only a limited number of studies used the recently introduced ISUP grading in their MVA, allowing for a comparison.
between ISUP grade 2 and 3. For RP and RT, three studies investigated the impact of the different ISUP grades on DM, PCSM or OM. One RT study compared the impact of bGS 3+4 and bGS 4+3 to bGS ≤6, reporting HRs of 0.99 (95% CI 0.60–1.68) and 1.55 (95% CI 0.98–2.47) for DM, 1.69 (95% CI 0.85–3.40) and 2.11 (95% CI 1.03–4.34) for PCSM and 1.31 (95% CI 0.81–2.10) and 1.30 (95% CI 0.79–2.13) for OM respectively [62]. Two RP studies compared pGS 3+4 and pGS 4+3 to pGS ≤6, reporting HRs of 5.55 (95% CI 1.30–23.78) and 10.82 (95% CI 2.56–45.76) for DM [26], 1.73 (95% CI 0.59–5.07) to 4.66 (95% CI 0.59–36.73) for pGS 3+4 and 2.5 (95% 0.83–7.56) to 8.04 (95% CI 1.04–61.96) for pGS 4+3 in relation to PCSM [26,40].

T category
Increasing T category at prostatectomy histology report (pT category) is not clearly associated with any of the oncological outcomes for patients undergoing RP, with only 7 out of 13 [7,8,22,23,26–28,38,39,41,43,52,53], 4 out of 12 [7,22,26,31,33,36,38–40,48,49,52] and 2 out of 6 [22,33,38,42,44,57] studies showing a significant correlation for DM, PCSM and OM respectively. For RT, 4 out of 6 [16,62,65,70,71,77], 3 out of 7 [16,17,59,61,62,65,72] and 2 out of 6 [16,58,59,61,62,64] studies showing a significant correlation between cT category and DM, PCSM and OM respectively.

Positive surgical margin
For patients who underwent RP, 2 out of 8 studies showed an inverse relationship with DM (i.e. “a protective effect”) [23,26,27,38,39,41,43] and none of the 8 studies investigating PCSM as outcome showed a significant correlation [26,31,33,38–40,48,49].

Interval to Biochemical Failure (IBF)
For patients undergoing RP as primary treatment, IBF was associated with DM in 2 out of 6 [7,8,22,26,28,38,42], PCSM in 8 out of 11 [7,22,26,28,31,33,34,38,40,48,50] and OM in 1 out of 4 [33,38,42,50] studies respectively. For patients undergoing RT as primary treatment, IBF had a clear association with oncological outcomes, with 5 out of 5 studies [62,63,65,73,75], 5 out of 7 studies [15,61–63,65,72,75] and 5 out of 6 [15,60–64] studies showing a strong association with DM, PCSM and OM respectively. Although studies repeatedly reported this association, the effect size was less interpretable due to the use of different thresholds. However, it is clear the shorter the IBF, the higher the risk of developing DM, PCSM and OM after both RP and RT.

PSA doubling time after radical treatment (PSA-DT)
There was a clear association between PSA-DT after primary therapy and oncological outcomes, being most pronounced in patients who underwent RP as primary treatment. All reported studies showed a significant correlation between a shorter PSA-DT and DM [7,8,29,41,42,45,53,56], PCSM [29,31,33,45,48,50] and OM [29,33,42,50,57]. The shorter the PSA-DT, the higher the risk of developing worse oncological outcomes. Similarly, for patients who underwent RT, the limited number of studies consistently showed that PSA-DT was a risk factor for DM [62,70,73,77].

**Salvage radiotherapy**

Four studies included salvage RT (sRT) as a co-variable in their MVAs of which 3 studies showed a significant impact on survival outcomes. Two studies showed a protective effect of sRT on OM with HRs ranging from 0.19 (95% CI 0.09-0.38) to 0.55 (95% CI 0.38-0.78) [44,57]. Furthermore, Cotter et al compared the prognostic role of sRT on OM in relation to PSA-DT. Compared to patients with a PSA-DT <6 months who received no sRT, patients with a PSA-DT <6 months who did receive sRT had a HR of 0.35 (95% CI 0.17-0.72). Patients with a PSA-DT ≥ 6 months with or without sRT had a HR of 0.19 (95% CI 0.09-0.38) and 0.31 (95% CI 0.17-0.56) respectively [44]. Importantly though, all patients were compared to the group with a PSA-DT <6 months, so the impact of sRT in patients with a PSA-DT >6 months is hard to interpret in this study. Similarly, Trock et al investigated the impact of sRT on PCSM in relation to PSA-DT [48]. They concluded that for patients with a PSA-DT <6 months, sRT resulted in a reduction of PCSM with a HR of 0.24 (95% CI 0.07-0.77) and 0.14 (95% CI 0.05-0.39) with or without concomitant ADT. In patients with a PSA-DT > 6 months, they reported no effect of sRT with HRs of 0.66 (95% CI 0.28-1.58) and 0.85 (95% CI 0.45-1.59) with or without ADT. Furthermore, they concluded that in patients with a PSA-DT <6 months, only when started within 2 years of BCR did sRT have an impact on PCSM with a HR of 0.14 (95% CI 0.06-0.034) compared to 0.80 (95% CI 0.11-5.93) when started with a delay of at least 2 years. For patients with a PSA-DT ≥ 6 months, the delay in sRT initiation did not have any effect on outcomes [48]. In contrast, Boorjian et al failed to show an impact of sRT on DM and PCSM. However, the median PSA-DT of patients with early and late BCR was 0.7 (IQR 0.4-2.8) and 1.3 (IQR 0.7-3.5) years respectively [7]. Based on the findings by Trock et al as described above, it could be expected that due to the relatively long PSA-DT the authors failed to show a protective effect of sRT on PCSM.
Eleven studies investigated only patients who received sRT [22–24,26–28,30,33,38,39,53]. For these patients, pGS remained a strong predictor for DM and to a lesser extent for PCSM. In 2 out of 4 and 4 out of 5 studies investigating IBF in their MVA for DM and PCSM as outcomes respectively, this was significant. PSA-DT was only investigated by one study and showed a significant effect [33]. Closely related are the PSA levels before initiation of sRT which showed to be a significant predictor for DM and PCSM with 7 out of 9 and 3 out of 4 studies showing a significant impact on DM and PCSM respectively. With increasing PSA levels, the risk of having worse oncological outcomes increased. Abugharib et al showed that patients with a pre-sRT PSA level of >0.5 vs 0.01-0.2 had an increased risk of developing DM and PCSM with HRs of 4.45 (95% CI 2.45-8.06) and 4.07 (95% CI 1.69-9.81) respectively. In contrast, patients with a PSA level ranging between 0.2-0.5 µg/l (vs 0.01-0.2) did not seem to have an increased risk [26]. Two other studies compared a pre-sRT PSA level of > 0.5 vs ≤0.5 µg/l with one showing a clear effect and the other study only showing a trend [27,28]. Again, this could be due to PSA being a continuous biomarker and its associated risks increasing gradually as well, which is confirmed by the studies by Jackson, Johnson and Stish et al [22,38,39]. These data suggest that a PSA cutoff of 0.4 µg/l that keeps increasing is probably indeed a good definition for BCR and that initiating sRT based on this definition is probably desirable.

Salvage androgen deprivation therapy

A limited number of studies report on the prognostic effect of salvage androgen deprivation therapy (sADT) after RP [7,40,44,57] or RT [18,61,70,78,80]. After RP as primary therapy, for PCSM HRs range from 0.83 (95% CI 0.45-1.55) to 1.09 (95% CI 0.5-2.36) [7,40,57] and for OM HRs range from 0.54 (95% CI 0.31-0.94) to 0.55 (95% CI 0.38-0.78) [44,57]. After RT, 4 studies report on the prognostic effect of sADT on DM based on a MVA and state it to have a protective effect [18,70,78,80], but only one study reports the actual HR of 0.43 (95% CI 0.28-0.69) [70]. For PCSM, two studies report contradictory findings [61,78].

Meta-Analysis

With multiple variables for each of the 2 therapy types, 3 different end-points and 2 different study types (continuous/categorical), the meta-analysis yielded a total 58 models (i.e. 58 p-values). After combining the p-values for each variable/end-point combination (combining p-value for categorical and continuous study types using the Fisher’s exact test) and correcting
for multiple testing using the Bonferroni test, there were 38 p-values of which 19 were significant (11 for RP, 8 for RT) (Fig 3, 4). The combined p-values for each variable/end-point combination are shown in Table 3 and significant prognostic factors were extracted from this table and summarized below.

For patients with BCR after RP, the following outcomes are associated with significant prognostic factors:

- Distant Metastatic recurrence: PSM, pGS, pT, PSA-DT, Pre-sRT PSA (Fig. 5)
- Prostate Cancer Specific Mortality: pGS, IBF, PSA-DT (Fig. 6)
- Overall Mortality: pGS, IBF and PSA-DT (Fig. 7)

For patients with BCR after RT, the corresponding outcomes are:

- Distant Metastatic recurrence: bGS, cT, IBF (Fig. 8)
- Prostate Cancer Specific Mortality: IBF (Fig. 9)
- Overall Mortality: Age, bGS, IBF, iPsa (Fig. 10)

The results of the sensitivity analysis are shown in Supplementary Table S9 (fixed-effect model), Supplementary Table S10 (random-effect model without accounting for dependence) and Supplementary Table S11 (conservative strategy). The first two models yielded 35 (21 for RP, 14 for RT) and 26 (14 for RP and 12 for RT) significant results respectively. Lastly, the conservative analysis yielded 17 significant results (10 for RP, 7 for RT). Potential publication biases in this meta-analysis were investigated using visual inspection of funnel plots and the Rank test and regression test (Egger’s test) for homogeneity of reported data. For RP (Fig S12, S13 and S14), funnel plots of pGS for outcomes DM and PCSM are asymmetrical with both a significant rank and regression test. Upon visual inspection, this is due to a limited number of studies that reported a high observed effect and high standard deviation. When these studies are visually ignored, the funnel plot would look symmetrical again. Excluding these studies doesn’t alter the conclusions of the meta-analysis. None of the other RP funnel plots are clearly asymmetrical, in addition all of the for RT (Fig S15, S16, S17) as well, none of the funnel plots are asymmetrical.

DISCUSSION

Principal findings
The available data demonstrated that patients experiencing BCR have an increased risk of developing distant metastases and dying due to PCa, and to a lesser extent have higher overall
mortality rates. However, the effect size of BCR as a risk factor for mortality was highly variable, ranging from HR 1.03 (95% CI 1.004-1.06) to HR 2.32 (95% CI 1.45-3.71) after primary RP [35,57]. Similarly, for patients who underwent primary RT, there is a clear correlation between BCR and clinically significant endpoints. Due to differences in statistical reporting, it is hard to define a precise generalized effect size of BCR on OM. However, as an example, two recently published studies reported absolute overall survival data and approximately showed approximately 20% lower survival rates at 8-10 years follow-up, even in men with minimal comorbidity [79,87]. Nevertheless, the variability in reported effect sizes of BCR remains high and suggests that only certain patient subgroups with BCR might be at an increased risk of mortality. It is for this reason that we further investigated which patient and tumor characteristics in patients experiencing BCR were prognostic for oncological endpoints.

A meta-analysis was undertaken to explore which prognostic factors were most strongly associated with oncological outcomes in patients that underwent RP or RT. Due to the heterogeneity of the studied populations and measurement of prognostic factors, we did not report on a pooled hazard ratio for the different prognostic factors. Rather, the meta-analysis is an exploratory analysis to attempt a more complete insight into the most strongly associated prognostic factors. Based on these analyses, in patients that underwent RP as primary treatment, the main prognostic factor for DM, PCSM and OM was a short PSA-DT. Due to the heterogeneity in the PSA-DT cutoffs that were investigated, no cutoff can be identified as being the most significant for disease recurrence. However, most studies associate a PSA-DT cutoff of < 12 month with an increased risk of clinical disease recurrence. This risk increases further with a decreasing PSA-DT (a more rapidly rising PSA). The shortest cutoff investigated is a PSA-DT < 3 months, which has the strongest association with clinical disease recurrence. To a lesser extent, a shorter IBF and an increasing pGS is associated with PCSM and OM. Again, the higher the pGS, the higher the reported risk of clinical disease recurrence with pGS 8-10 being most associated with poor prognosis. Similar to PSA-DT, different studies investigated different cutoffs, but most studies associated an IBF < 18 months with an increased risk of disease recurrence. For patients that underwent primary RT, clearly the strongest prognostic factor for DM, PCSM and OM was a short IBF. Again, the cutoffs used between studies were heterogeneous, however most studies use IBF < 18 months as a conservative cutoff that is
associated with an increased risk of clinical disease recurrence. Compared to RP patients, IBF is probably more prognostic for patients who underwent RT because it unifies two prognostic PSA kinetic factors after RT: the PSA nadir and the PSA-DT. A short IBF could therefore be caused by reaching the PSA nadir early (with higher PSA nadir levels reached early as a sign of poor response to RT) and/or a short PSA-DT after the PSA nadir has been reached (as a sign of cancer progression). Increasing bGS is also associated with DM and OM. Although only 5 studies reported on a population that received a median radiation dose >74 Gy, data reported in these studies are highly similar to those reported in the meta-analysis (which was performed on all studies). However, we believe the reported data remain generalizable to today’s treatment regimens. Although the radiation dose affects treatment effectiveness, the investigated population is a patient group that has already failed their treatment. Furthermore, treatment biology and therefore cancer cell response probably remains similar irrespective of the dose. Age determines OM as expected, both for patients receiving RP or RT.

As primary meta-analysis, a random-effect model was used, which assumes that the investigated studies were drawn from populations that differ in ways that could impact the outcome, for example the age of the subjects or disease stage. Due to multiple co-variables being introduced in the same meta-analysis (for example pGS 7 vs pGS 6 and pGS 8-10 vs pGS 6 of the same study), these co-variables were treated as being dependent on each other. Therefore, under the random effects model the goal is not to estimate one true effect (one HR summarizing all data), but to estimate the mean of a distribution of effects. In contrast to the random-effect model, the fixed-effect model assumes that all studies are more or less homogeneous and share a common true effect size. For sensitivity analysis of the meta-analysis, we therefore compared the results of the fixed-effect model with the primary random-effect model. The former has a lot more significant results, suggesting that the assumption of homogeneity between the studies would be inappropriate and the results obtained via the random-effect model are likely more reliable compared to the fixed-effect model. Repeating the random-effect model without accounting for study dependence did not severely decrease the number of identified associations while being a considerably more appropriate method of analysis. Lastly, comparing our primary analysis to a conservative analysis strategy (an analysis designed to avoid the detection of false positives) yielded almost
the same set of significant results. The high degree of agreement between these two types of analyses provides further evidence of robustness (insensitivity) of the outcome of the primary analysis. To exclude potential publication bias of the meta-analysis results, funnel plots were developed for each co-variable. The only clearly asymmetrical funnel plots were for the co-variable pGS for the outcomes DM and PCSM. This effect is due to a limited number of studies with a high observed effect and large standard error (Fig S12, Fig S13). We are convinced this does not compromise the validity of the results, since overall most studies clearly show a prognostic role for pGS (12/14 studies for DM and 10/13 studies for PCSM, see above) and excluding these studies would not change the impact of pGS. Furthermore, a priori we decided not to report on a pooled effect size for the reported meta-analysis because we expected study heterogeneity. Therefore, the inclusion of a limited number of studies with a very high effect size does not affect our study outcomes.

Implications for clinical practice and further research

Currently, the EAU Prostate Cancer Guidelines present a list of risk factors associated with oncological outcomes based on a limited number of large studies. Based on these extensive analyses, we can refine this list to a limited number of risk factors, which should be taken into consideration when following up patients who develop BCR. Therefore, we propose the introduction and further investigation of a novel BCR risk stratification. For patients experiencing BCR with a PSA-DT > 1 year and a pGS <8 for RP or an IBF > 18 months and bGS <8 for RT, patients have a significantly lower risk of clinical progression. These men could be classified as having 'EAU Low Risk BCR'. In contrast, patients with a shorter PSA-DT ≤ 1 year OR a pGS 8-10 for RP or an IBF ≤ 18 months OR a bGS 8-10 for RT have a higher risk of clinical progression and could be classified as having 'EAU High Risk BCR'. Therefore, both its prognostic value and the potential benefit/toxicities of initiating salvage treatment(s) should be discussed with the patient. Although data are limited, the most robust analysis regarding the impact of sRT based on patients' risk profile currently available is performed by Trock et al [48]. In this study, in patients with a PSA-DT <6 months, sRT +/- ADT resulted in a reduction of PCSM with HRs of 0.14 (95% CI 0.05-0.39) - 0.24 (95% CI 0.07-0.77). In contrast, in patients with a PSA-DT ≥ 6 months, sRT +/- ADT had no prognostic value with HRs of 0.66 (95% CI 0.28-1.58) to 0.85 (95% CI 0.45-1.59). For sADT, the level of evidence is too low to suggest any recommendations. Although we acknowledge that
prognostic studies are not the appropriate study design to investigate the therapeutic effect of a certain treatment, it does show that not all patients with BCR are equal in outcome and therefore should not receive the same treatment. As an example, if a PSA rise >0.2 ng/mL is confirmed in a patient after primary RP, the pGS is 7 or lower and if PSA kinetics are favorable (PSA-DT >12 months, IBF >18 months), patients should be informed that they have low-risk BCR and the potential toxicities of additional treatments should be discussed to allow the patient to make a well-informed decision. The included study designs do not allow us to make any further recommendations if or when to start salvage therapies when BCR occurs as this was not the goal of this review. Furthermore, the introduced risk stratification is only proposed to make clinicians aware that not all patients with BCR have equally poor outcomes. Its predictive and prognostic power needs to be confirmed based on individual patient data. Due to clinical heterogeneity and heterogeneous study design, a more in-depth interpretation of the available data (such as reporting on a pooled hazard ratio for the investigated risk factors) would be inappropriate. In an ideal situation, access to individual patient data from each study would allow us to perform a robust meta-analysis, with standardization of patient criteria, outcomes and subgroup analyses to explore interactions between variables. To this end, the EAU Guidelines Office has established the PIONEER consortium with the purpose of combining data from large organizations across different countries allowing us to work with big data, answering critical questions in PCa care. Finally, with the introduction of novel, more sensitive imaging techniques such as PSMA PET-CT scan, more patients with biochemical recurrence will be diagnosed with metastatic disease at lower PSA levels. Although the diagnostic landscape will change (and in many countries has changed already) due to these imaging techniques, currently it still remains unclear what the therapeutic implications of these metastatic lesions are. Only future research will be able answer these questions.

Limitations and strengths
The current study represents the first systematic review addressing the impact of BCR on clinically important endpoints (i.e. development of distant metastases, prostate cancer specific mortality and overall mortality). The review elements were developed in conjunction with a multidisciplinary panel of experts (EAU Prostate Cancer Guidelines Panel), including a patient representative, and the review was performed robustly in accordance with recognized standards. Important prognostic factors for cancer-related outcomes in patients with BCR
were identified and the review highlighted further areas of research that could help define which patients are at the highest and lowest risk of clinical progression. Another strength of the review is the robustness of the analysis, with additional sensitivity and publication bias analysis being performed to test the integrity of the results. Limitations include the retrospective nature of the clear majority of studies, the reliability of PSA-DT measurements in low-ranged PSA levels, heterogeneity in PSA-DT measurement (number of data points, interval between PSA data points to calculate PSA-DT) and the overall significant clinical and methodological heterogeneity across studies which limited the quality of the data and precluded further strong recommendations.

CONCLUSION

BCR is an independent risk factor for the development of distant metastases, cancer-specific mortality and to a lesser extent overall mortality. In patients who underwent RP as primary treatment and who subsequently developed BCR, the main prognostic factor for DM, PCSM and OM is a short PSA-DT (i.e. < 1 year) and an increasing pGS and to a lesser extent a short IBF. The main prognostic factor for patients developing BCR following primary RT is a short IBF (<18 months) and an increasing bGS. After primary RP, sRT protects patients from experiencing disease progression, but this effect is limited to a subpopulation of patients at the highest risk of progression with a PSA-DT < 6 months. Therefore, based on this systematic review, we recommend that patients experiencing BCR should not all be treated equally. Rather, we propose patients to be stratified into having EAU Low-Risk BCR (PSA-DT >1 year and pGS <8 for RP, IBF > 18 months and bGS <8) or EAU High-Risk BCR (PSA-DT ≤1 year or pGS 8-10 for RP, IBF ≤18 months or bGS 8-10), raising awareness that not each patient with BCR has similar outcomes. The potential benefit of initiating salvage treatment(s) should be discussed with each individual patient. In contrast, in the absence of risk factors, the non-aggressive course of the disease and the toxicities of salvage treatments should be discussed to allow patients to make a well-informed decision. Before this new classification could have therapeutic implications, future research needs to validate the predictive and prognostic value of this novel risk stratification using individual patient data.

REFERENCES


Commented [MDS4]: Not sure if increasing is the right word here as it implicates that the Gleason score would increase in one patient. Maybe “high” gleason score would be clearer and mentioning the minimum number (as you did with PSA-DT)

Commented [PCS5]: Consider switching for “the presence of a more aggressive tumour in the prostatectomy specimen”

Commented [MDS6]:

Commented [MDST6]: Just a suggestion which can be dropped of course. Maybe explain in the conclusion that in this paper EAU xx -risk BCR has been “invented”:

...we propose using newly defined EAU risk BCR groups,...
Calculated prior to PSA Recurrence, Correlate with Prostate Cancer Outcomes?


