available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.ejconline.com

Review

Prostate-specific antigen (PSA) alone is not an appropriate surrogate marker of long-term therapeutic benefit in prostate cancer trials

Laurence Collette[a,*], *Tomasz Burzykowski*[b], *Fritz H. Schröder*[c]

[a]*European Organisation for Research and Treatment of Cancer (EORTC) Data Center, Avenue E. Mounier 83/11, 1200 Brusells, Biostatistics Department, Brussels, Belgium*
[b]*Center for Statistics, Hasselt University, Diepenbeek, Belgium*
[c]*Department of Urology, Erasmus MC, University Medical Center, Rotterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 27 January 2006
Received in revised form
2 February 2006
Accepted 6 February 2006
Available online 30 May 2006

Keywords:
PSA
Surrogate endpoint
Clinical trials
Prostate cancer
Prognostic factor

ABSTRACT

The prostate-specific antigen (PSA) is the most studied marker of prostate cancer. It is used for screening and as indicator of disease evolution for individual patients. PSA being a prognostic factor is however not sufficient to justify using PSA-derived endpoints as surrogate for definitive survival endpoint in phase III trials. First, we clarify the terminology and requirements for a marker to be a valid surrogate endpoint. We then review the published literature pertaining to the validation of PSA endpoints as surrogate in all disease stages. We discuss the limitations of these studies and conclude that so far, PSA is not a validated surrogate endpoint in any of the disease settings and treatment conditions considered. We give some recommendations for the planning of trials that would use PSA endpoints (in hormone refractory disease) and for the early stop of (endocrine treatment) trials on the basis of intermediate results based on PSA.

1. Introduction

Phase III cancer clinical trials that evaluate the clinical benefit of new treatment options often require large patient numbers and long follow-up. Recent advances in the understanding of the biological mechanisms of disease development have resulted in the emergence of a large number of potentially effective new agents. There is also increasing public pressure for promising new drugs to receive marketing approval as rapidly as possible, in particular for life threatening diseases such as cancer. For these reasons, there is an urgent need to find ways of shortening the duration of cancer clinical trials. The duration of phase III trials results from the use of long-term clinical endpoint (clinical progression, survival). Therefore, to replace this endpoint (the "true" endpoint) by another (a "surrogate" endpoint), that could be measured earlier, more conveniently or more frequently, and that would adequately reflect the benefit of new treatments on the clinical endpoint(s), seems an attractive solution. In the field of prostate cancer, prostate-specific antigen (PSA) has probably been the most studied biomarker.[1] It has been investigated as a prognostic factor and as a potential surrogate endpoint across disease stages.

It is a common misconception that established prognostic factors necessarily make valid surrogate endpoints. A

* *Corresponding author:* Tel.: +32 2 774 16 69; fax: +32 2 771 38 10.
E-mail address: laurence.collette@eortc.be (L. Collette).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2006.02.011

prognostic factor is an intermediate outcome that is correlated with the true clinical outcome (T) for an individual patient.[2] Its knowledge may be useful for diagnostic or prognostic assessment of an individual patient. For a prognostic factor to be a surrogate endpoint (S), it is further required that "the effect of treatment on a surrogate endpoint must be "reasonably likely" to predict clinical benefit".[2,3] In other words, a biomarker S will be a good surrogate for the true endpoint T if the results of a trial using outcome S can be used to infer the results of the trial if T had been observed and used as endpoint and this with sufficient precision. To demonstrate surrogacy, a high association between the treatment effects on the surrogate and on the true endpoint thus needs to be established across groups of patients treated with a new versus a standard intervention.

Fig. 1 shows a schematic of two situations: a) where post-treatment PSA level (S) is prognostic for mortality risk (T) (as shown by the diagonal orientation of the ovals representing the individual patient data in two treatment groups), but is not surrogate, as is indicated by the line linking the two group's averages being horizontal; and b) where post-treatment PSA level (S) is weakly prognostic for mortality risk (T) (as shown by the more horizontal orientation and more circular shape of the ovals representing the individual patient data), but is a strong surrogate of the treatment difference on the mortality risk, as shown by the bar linking the group's averages being diagonal, so that differences in average post-treatment PSA level between the treatment groups correlate with difference in average mortality risk.

To illustrate this, let us consider the recently published secondary results of the Tax-327 study.[4] This study compared a weekly and a three-weekly schedule of docetaxel plus prednisone to mitoxantrone and prednisone in hormone refractory prostate cancer (HRPC). In this study, like often in this disease state, patients who achieved a PSA response had a 60% reduction in mortality risk compared with non-responders (hazard ratio (HR) = 0.40, 95% CI: 0.31–0.51). The reduction of the PSA by 50% or more from baseline value, which was defined as a PSA response, was a strong prognostic factor for survival. Now considering PSA response as an endpoint and as a putative surrogate for overall survival, we observe with the authors that the weekly docetaxel arms resulted in a response rate of 48% which was significantly different from the 32% response rate that was obtained with standard arm mitoxantrone plus prednisone (P < 0.0001). However, the median overall survival on the weekly docetaxel arm amounted 17.4 months and did not differ statistically significantly from the 16.5 months median survival achieved with the standard treatment (P = 0.362, Fig. 2). The benefit amounting less than a month was also not medically relevant, contrary to the difference in response rates. Thus in this study, PSA response although it was a strong prognostic factor for survival at the patient level, did not appear to be reliable as a surrogate for survival when comparing the weekly docetaxel treatment to mitoxantrone plus prednisone.

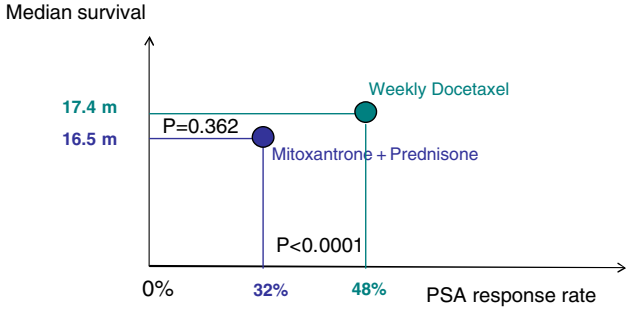
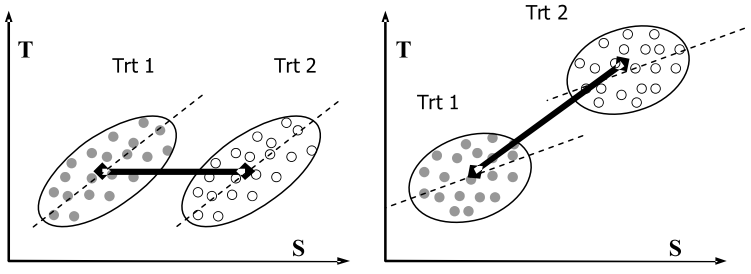


Fig. 2 – A prognostic factor does not make a surrogate endpoint–the Tax 327 trial.

2. Statistical validation of surrogate endpoints

Traditionally, the "Prentice Criteria"[5] were used for the purpose of demonstrating surrogacy on the basis of data from a single trial. The Prentice criteria require that four conditions be shown to be true in order to demonstrate the validity of a putative surrogate endpoint (here PSA), as a replacement endpoint for a true endpoint T (here survival):

(a) There must be a statistically significant treatment effect on the PSA endpoint (in univariate analysis)
(b) There must be a statistically significant treatment effect on survival (univariate analysis)



(a) S (PSA level) prognostic for T (mortality risk) but is not surrogate

(b) S (PSA level) is surrogate for T (mortality risk) and is weakly prognostic

Fig. 1 – Prognostic factor versus Surrogate endpoint. Schematic of two situations where (a) the surrogate S (PSA) is a good prognostic factor for the true endpoint T (mortality risk) in both treatment groups but is a not a surrogate for T and (b) the surrogate S (PSA) is only a weak prognostic indicator of the endpoint T (mortality) at the individual level and is a good surrogate endpoint for replacing the true endpoint T (mortality) in phase III clinical trials.

(c) The PSA endpoint must be a statistically significant prognostic factor for survival (univariate analysis)
(d) The treatment effect on survival must completely vanish in a survival model with both the treatment and the PSA endpoint as explanatory variables (multivariate analysis).

Although often used, the criteria are not a proper tool to check the validity of a surrogate. They do not aim at verifying the quality of prediction of clinical benefit. Condition (b) limits the applicability of the criteria to trials that showed a statistically significant treatment effect on the true endpoint, a condition which is rarely fulfilled by clinical trials in prostate cancer. Condition (d) is impossible to verify in practice, as it amounts to "proving a null hypotheses", i.e., showing that the treatment effect is zero. Usually, it is checked by requiring that a statistical test shows the treatment effect to be statistically not significant in a model adjusted for the surrogate endpoint. Statistical tests however are designed to reject a null hypothesis and non-rejection of the null hypothesis never stands as definitive proof that the null hypothesis is true.[6] In fact, one can obtain a non-significant test result simply by having an inadequate sample size. Finally, it was shown that for time-to-event endpoints, the Prentice criteria are neither necessary nor sufficient to demonstrate that surrogacy holds true.[7] Thus, failure to demonstrate that the four criteria hold does mean a biomarker should be disregarded as a surrogate endpoint, while successful demonstration that the four criteria hold true is not sufficient to actually demonstrate that a biomarker is a surrogate for a time-to-event endpoint.

More recently, a new methodology known as the "meta-analytic validation" was developed.[8,9] Using data from several trials, this method consists in deriving a model that can predict the magnitude of the treatment effect on the true endpoint, from the treatment difference observed on the surrogate (PSA) endpoint. A surrogate is valid if the prediction is sufficiently precise. This new methodology aims directly at verifying whether "the effect of treatment on a surrogate endpoint is reasonably likely to predict clinical benefit". Furthermore, it does not require that any of the treatment effects in the individual studies be statistically significant. It does necessitate, however, large databases from multiple randomized clinical trials with similar design and treatments. Using data from several trials, the method consists of simultaneously estimating the treatment effect (e.g., hazard ratio) for the true (survival) endpoint and for the surrogate (PSA) endpoint in each trial. The association between the treatment effects on the true endpoint and the corresponding effects on the surrogate endpoint is then modelled in a way similar to standard linear regression (Fig. 3), although mathematically more sophisticated. Alike in linear regression, the strength of the association is measured by the squared correlation coefficient (R^2_{trial}) that also indicates the precision with which the treatment effect on the true (survival) endpoint can be predicted from the observed effects on the surrogate (PSA). The maximal possible value of R^2_{trial} is 1 which indicates a perfect prediction. In practice, observing $R^2_{trial} = 1$ is not possible and one rather seeks a value close to one which indicates a strong association between the treatment effects and thus a relatively precise prediction.[8,10]

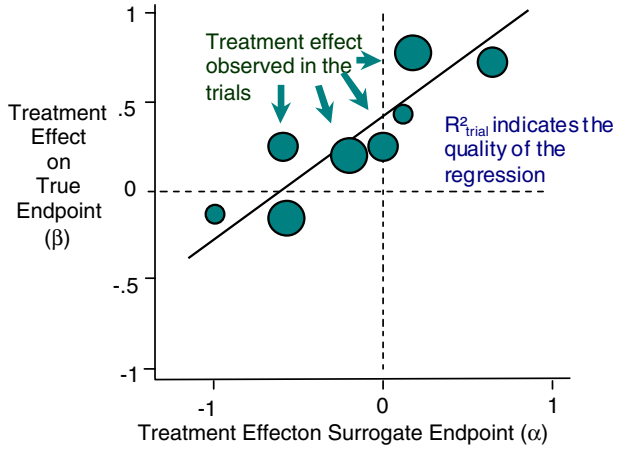


Fig. 3 – Prediction using data from several trials: the meta-analytic validation method.

3. Published results on PSA surrogacy in prostate cancer

Although the literature concerning the association between PSA and long-term outcome with prostate cancer is extensive, there are relatively few reports of true validation studies of this endpoint. We shall critically review the published evidence assessing PSA endpoints (PSA response, time to PSA progression, PSA velocity, PSA doubling time) as potential surrogate endpoint for overall or progression-free survival, for each stage of prostate cancer.

3.1. *Non-metastatic disease*

D'Amico and colleagues[11] studied the surrogacy of a PSA doubling time less than 3 months, as a potential surrogate for prostate cancer mortality, in a non-randomized cohort of 5918 men treated with surgery and 2751 with radiation. They showed that the Prentice criteria were fulfilled, however the fourth condition was demonstrated by showing no effect of the initial treatment on the cancer specific survival after PSA relapse, in the subset of 1551 patients with PSA relapse. The value of the study is limited by the non-randomized nature of the series, the fact that the three-month cut-point is data driven and the fact that the timing of salvage hormonal treatment was not accounted for. The applicability of the results is limited by the fact that few patients actually have a PSA doubling time shorter than three months (74 of 611 cases with PSA relapse after radical prostatectomy, 12%).

Sandler and colleagues[12] showed that in the Radiation Therapy Oncology Group (RTOG) trial 92-02 that compared short-term versus long-term androgen deprivation in addition to irradiation for T2c-T4 prostate cancer, time to PSA failure (defined using the American Society for Therapeutic Radiation Oncology (ASTRO) definition) was not a surrogate for cancer-specific survival: the PSA endpoint failed the fourth Prentice criteria. In that study, time to PSA failure was longer on the long-term androgen deprivation arm but the survival time after PSA failure was shorter. The authors

postulated that on the long-term androgen deprivation arm, some patients may have already had hormone insensitive disease at the time of PSA relapse and thus decreased responsiveness to salvage treatment. They concluded that time to PSA failure should not be used as a surrogate endpoint in trials that test endocrine treatment of differing duration. Two years later, Valicenti and colleagues[13] reported from the same study, showing that post-treatment PSA doubling time (calculated using first-order kinetics on the basis of minimum three post-treatment measurements) of less than 12 months fulfilled all Prentice's criteria in respect to the endpoint of prostate cancer mortality. In their study, 142 of the 1514 eligible patients had died of prostate cancer. These two reports suggest that dynamic measures of PSA might be stronger surrogates than static measures such as the PSA increase above a threshold value.

Newling and colleagues[14] carried out a meta-analytic validation of PSA-doubling free survival (BPFS) as potential surrogate for clinical progression free survival (PFS) in over 8000 patients with localized or locally advanced M0 disease, who were randomized within AstraZeneca's Early Prostate Cancer Program between treatment with bicalutamide (Casodex) 150 mg daily versus placebo in addition to standard care (radical prostatectomy, radiotherapy or watchful waiting). PFS was defined as the time to objectively confirmed disease progression or death from any cause. They report an R^2_{trial} of 0.65 (95% CI: 0.55–0.92) for the whole group and of 0.52 (95% CI: 0.37–0.89) on only the European patients, and a lower association in prostatectomy patients ($R^2_{trial} = 0.46$) than in irradiated patients ($R^2_{trial} = 0.65$). They concluded that large positive treatment effects on BPFS are likely to reflect a clinically important benefit of bicalutamide as regards clinical PFS. They estimated the minimum reduction in the risk of a PSA doubling to yield a significant reduction ($P < 0.05$) in the risk of a PFS event to $\geqslant$20% in all patients, $\geqslant$30% in radical prostatectomy patients and $\geqslant$50% in irradiated patients. However, we must note that part of the association observed in their study may be induced by the overlap between PFS and BPFS for the patients in whom the first event is death in absence of PSA doubling.

3.2. *Metastatic disease*

As a counter example to time to PSA being a surrogate for survival in metastatic disease, we can already mention trial National Cancer Institute (NCI) INT-105[15] that randomized 1382 eligible patients undergoing bilateral to additional flutamide or nil. In that study, the treatment differences in post-therapy PSA response; defined as a PSA level $\leqslant$4 ng/mL at any time after randomization did not translate into survival differences; the PSA response rates on treatment and control were 74.0% versus 61.5% ($P < 0.0001$) but there was no significant difference in overall survival ($P = 0.14$). The latter may be related to a lack of statistical power for survival in this study. However, a meta-analysis of 8275 patients later confirmed the absence of benefit of maximal androgen blockade over castration.[16]

More recently, Collette and colleagues[17] reported a meta-analytic validation of several PSA endpoints (PSA response defined as $\geqslant$50% decline from baseline PSA level, PSA normalization, time to PSA progression) as potential surrogates for overall survival in a database of 2161 patients with primary diagnosis of metastatic prostate cancer who had been treated within AstraZeneca's Casodex (bicalutamide) development program. The patients were randomized between bicalutamide monotherapy and castration or between combined androgen blockade with bicalutamide or with flutamide. The study showed that the association between the treatment effect on any PSA based endpoint and the treatment effect on overall survival was in general low ($R^2_{trial} < 0.69$ with wide confidence intervals). The association between the time to PSA progression defined as a confirmed 50% relative increase above the previously observed nadir yielded $R^2_{trial} = 0.66$ (standard error = 0.13) with a corresponding 95% confidence interval ranging from 0.30 to 0.85. Sensitivity analyses using prostate-cancer survival as the true endpoint led to similar results. Similar to Newling and colleagues,[14] they concluded that non-null treatment effects on survival would potentially be identifiable only in new trials showing a very large effect on the PSA endpoint (e.g., HR around 0.50 with standard error = 0.10) on the basis of large patient numbers. Moreover, irrespective of the size of the effect on the PSA endpoint, the prediction of the treatment effect on overall survival could not be precise, due to the large unexplained variability in the estimated prediction model (as indicated by low R^2_{trial} values). Thus, with the information at hand, a trial based on the PSA endpoint would not require fewer patients than survival trial.

3.3. *Hormone refractory disease*

D'Amico and colleagues[18] assessed whether PSA velocity (calculated by linear regression of all PSA values within one year of initially detectable and increasing PSA level) can serve as surrogate endpoint for prostate cancer specific mortality (PCSM) in 919 patients with non-metastatic hormone refractory prostate cancer (HRPC) treated with salvage hormonal treatment for PSA failure after initial radical prostatectomy or radiation therapy. They demonstrated that a PSA velocity >1.5 ng/mL yearly fulfilled the Prentice conditions of surrogacy for the endpoint PCSM. However, only 26 patients died of prostate cancer in their study, and their demonstration (in particular Prentice's fourth criteria) is therefore potentially affected by lack of statistical power. In addition, the cut-point of 1.5 ng/mL yearly was data driven and needs further validation in an independent dataset, the study is non-randomized and the models used did not control for the timing of the salvage hormonal treatment. In view of these limitations, the authors themselves conclude that they cannot claim that they have completely demonstrated surrogacy.

Crawford and colleagues[19] used Prentice's criteria to demonstrate the surrogacy of the three-month PSA change (PSA velocity) as surrogate for mortality in the SouthWest Oncology Group (SWOG) trial S9916 that compared docetaxel/estramustine to mitoxantrone/prednisone in 770 patients with HRPC. The four Prentice criteria were fulfilled and they concluded that PSA velocity measured during the first three months on study should be further studied as surrogate endpoint for mortality in future studies of chemotherapeutic regimens for HRPC.

The findings in the Tax 327 study mentioned earlier[4] conflict to some extent with those of Crawford and colleagues since the use of PSA changes would have resulted in wrong conclusions regarding the weekly docetaxel arm. Therefore the question whether PSA endpoints should be used as surrogate in chemotherapy trials or in trials involving docetaxel remains not fully answered.

The only meta-analytic validation study in HRPC we know of is the study by Buyse and colleagues[20] who assessed several PSA-based end points in androgen-independent patients treated with liarozole, cyproterone acetate or flutamide. They showed that despite a strong prognostic association neither PSA response (defined as a decline by 50% or more from baseline level), nor time to PSA progression (defined as a greater than 50% increase over nadir value) qualified as a surrogate for overall survival (R^2_{trial} was <0.45 for all tested PSA endpoints). One of the reasons for the lack of association may relate to the mode of action of liarozole which is an imidazole-like compound that causes elevation of retinoic acid, postulated to have anti-tumour activity and which effect may not be mediated by PSA. Other reasons for the lack of association might be that the patient population was very advanced and that PSA expression might be affected by tumour de-differentiation. This suggests at least that surrogacy of PSA endpoints might not be generally applicable between treatments with very different modes of actions or which effect on PSA is expected to differ substantially.

4. Discussion

The literature on PSA surrogacy thus far failed to satisfactorily demonstrate the value of PSA as a surrogate endpoint in prostate cancer.

From this review, one can broadly conclude that for the comparison of primary treatments, PSA is until now not proven to be a suitable replacement for a final survival endpoint. The association between PSA changes after initial treatment and survival is likely to diminish in the future, as second and third line treatments may become increasingly efficacious. As seen in the RTOG 92-02 trial,[12,13] caution is especially needed when only one of randomized treatments involves long-term hormonal manipulations because PSA will not reflect the development of hormone refractory disease, that carries poor prognosis for salvage. Vicini and colleagues[21] recently reviewed the value of monitoring PSA after initial treatment for prostate cancer. They concluded that PSA reading should not be used rapidly to judge difference in treatment efficacy in this setting.

The studies on PSA velocity and other dynamic measures of PSA changes suggest that, these might be more powerful than classical definitions of PSA changes using threshold values and the results by D'Amico and colleagues[11] need further validation. PSA doubling time and PSA velocity have otherwise been mostly studied for testing chemotherapeutic agents against HRPC. However, it is well documented that all pharmacological agents do not affect PSA in the same way:[22] drugs may decrease, increase or not change PSA, with or without a delay after treatment initiation. Therefore the PSA endpoint in phase II or phase III trials should be designed to match the anticipated effect of the tested drugs on PSA levels.[23] In addition, it is also essential to understand and document the drug's effect on tumour growth and how it correlates with PSA changes, since drugs, e.g., suramin, were shown to modify PSA production without having an impact on tumour growth.[24] For this purpose, the algorithm proposed by Schröder and colleagues[25] is very interesting. It incorporates an experimental "proof-of-concept" *in vivo* study before or in parallel with the phase II clinical trial. The design of phase II trials of targeted agents is further discussed by Stadler,[26] who shows that the Bubley definition[27] of PSA response in phase II of HRPC is not an appropriate endpoint when testing cytostatic drugs. Consequently to this and as seen in the review, further research is still needed before using measures of PSA change as the final endpoint in phase III studies in HRPC.

The meta-analytic validation studies of Newling and colleagues[14] and of Collette and colleagues[17] in hormonally treated patients, confirm only moderate correlation between effects of hormonal treatments on the final clinical endpoint and on the PSA endpoints considered. Thus, phase III trials in these settings should not be based on PSA endpoints.

However, PSA could still be used to shorten, in two ways, the duration of a phase III trial testing a new treatment the effect of which is known (from preclinical and early phase studies) to be expressed or mediated at least in part by PSA. First, early registration on the basis of a PSA endpoint could be envisaged in trials with clinical progression or survival as primary endpoint. For that purpose, the trial sample size should be determined to demonstrate a very large effect on the PSA endpoint with great precision: for example to demonstrate the presence of a hazard ratio of the order of 0.50 on the PSA endpoint, with a standard error of the order of 0.10. The trial sample size calculation ought not to be on power considerations, as these would necessarily result in very small sample size due to the large target effect, but should be based on the required precision of the estimation of the treatment effect on the PSA endpoint. An interim analysis plan should be set up with plan for one or several interim looks at the PSA endpoint as well as to the safety data and one longer term analysis on the survival endpoint. At the interim, the trial results on the PSA endpoint could be used to estimate a prediction of the survival treatment effect using the regression results from former meta-analytic validation. Whenever the prediction interval for the survival hazard ratio would exclude the null effect, the trial results could be submitted for early registration on the basis of the PSA results. In the light of the fact PSA is unlikely to capture all the potential (negative) effects of the treatment on survival and because PSA did not qualify as a surrogate endpoint, we recommend that the follow-up should continue to later document long-term safety of the treatments and their impact on survival. Of note, this procedure would likely not reduce the patient number to enter in studies. Second, even if PSA is not a surrogate, a treatment effect on the PSA endpoint might be seen, for specific drugs, as a pre-requisite for an ultimate effect on survival. Thus along the lines proposed by Royston and Parmar[28] one could design a study with survival as the final endpoint, but with planned interim looks at the PSA endpoint and a decision to stop the study for futility if insufficient

benefit or a negative effect of the experimental treatment (e.g. HR > 0.5) was seen on the PSA endpoint. The assumption underlying this design is that a significant treatment effect on the final endpoint (survival) would be very unlikely if no beneficial effect was seen on the intermediate endpoint (PSA). The information on the correlation between the treatment effects on the PSA endpoint and on survival that is necessary to the set up of stopping rules according to Royston and Parmar is available from the value of R^2_{trial} obtained from meta-analytic validations, whenever they exist.

PSA is the most widely available marker for prostate cancer. However, PSA is not tumour specific. Prognostic studies have also shown that, in hormone independent disease, only 17% of the variability in survival is explained by PSA.[29] Therefore, it is unlikely that endpoints based solely on the marker PSA can make valid surrogate endpoints for long-term clinical outcome with prostate cancer. New serum and urine markers in prostate cancer are currently being studied, noticeably within the European Community project P-MARK (http://www.p-mark.org). A large number of these markers show promise to overcome the limitations of PSA[30] and may in the future offer more solid surrogate endpoints to shorten the duration of phase III trials in prostate cancer.

Conflict of interest statement

The authors have no conflict of interest to declare.

REFERENCES

1. Small EJ, Mack III Roach. Prostate-specific antigen in prostate cancer: a case study in the development of a tumour marker to monitor recurrence and assess response. *Semin Oncol* 2002;**19**:264–73.
2. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;**69**:89–95.
3. Johnson JR, Williams G, Pazdur R. Endpoints and United States Food and Drugs Administration Approval of Oncology Drugs. *J Clin Oncol* 2003;**21**:1401–11.
4. Roessner M, de Wit R, Tannock IF, et al. Prostate-specific antigen (PSA) response as surrogate endpoint for overall survival (OS): Analysis of the TAX 327 Study comparing decetaxel plus prednisone to mitoxantrone plus prednisone in advanced prostate cancer. *J Clin Oncol* 2005;**23**:391s.
5. Prentice RL. Surrogate endpoints in clinical trials: definitions and operational criteria. *Stat Med* 1989;**8**:431–40.
6. Altman DG, Bland JM. Absence of evidence is not evidence of absence. *Br Med J* 1995;**311**:485.
7. Buyse M, Molenberghs G. Criteria for the validation of surrogate end-points in randomized experiments. *Biometrics* 1998;**54**:1014–29.
8. Daniels MJ, Hugues MD. Meta-analysis for the evaluation of potential surrogate markers. *Stat Med* 1997;**16**:1515–27.
9. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* 2000;**1**:49–68.
10. Burzykowski T, Molenberghs G, Buyse M, Geys H, Renard D. Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. *J R Stat Soc Ser C Appl Stat* 2001;**50**:405–22.
11. D'Amico AV, Moul JW, Carroll PR, Sun L, Lubeck D, Chen MH. Surrogate end point for prostate cancer specific mortality after radical prostatectomy or radiation therapy. *J Natl Canc Inst* 2003;**95**:1376–83.
12. Sandler HM, Pajak TF, Hanks GE, Porter AT, DeSilvio M, Shipley WU. Can biochemical failure (ASTRO definition) be used as a surrogate endpoint for prostate cancer survival in phase III localized prostate cancer clinical trials? Analysis of RTOG protocol 92–02. *J Clin Oncol* 2003;**22**:381.
13. Valicenti R, Deslivio M, Hanks G, et al. Surrogate endpoint for prostate cancer-specific survival: Validation from an analysis of the Radiation Therapy Group Protocol 92-02. *J Clin Oncol* 2005;**23**:4549.
14. Newling D, Carroll K, Morris T. Is prostate-specific antigen progression a surrogate for objective clinical progression in early prostate cancer? *J Clin Oncol* 2004;**22**:4652.
15. Eisenberger MA, Blumenstein BA, Crawford ED, et al. Bilateral orchiectomy with or without Flutamide for metastatic prostate cancer. *N Engl J Med* 1998;**339**:1036–42.
16. Prostate Cancer Trialists' Collaborative Group. Maximum androgen blockade in advanced prostate cancer: An overview of the randomized trials. *Lancet* 2000;**355**: 1491–1498.
17. Collette L, Burzykowski T, Caroll KJ, Newling D, Morris T, Schröder FH. PSA is not a valid surrogate endpoint for overall survival in patients with metastatic prostate cancer. A joint research of the European Organisation for Research and Treatment of Cancer, the Limburgs Universitair Centrum and AstraZeneca Pharmaceuticals. *J Clinl Oncol* 2005;**23**:6139–48.
18. D'Amico AV, Moul J, Carroll PR, Sun L, Lubeck D, Chen MH. Surrogate end point for prostate cancer specific mortality in patients with non- metastatic hormone refractory prostate cancer. *J Urol* 2005;**173**:1572–6.
19. Crawford ED, Pauler DK, Tangen CM, et al. Three-month change in PSA as a surrogate endpoint for mortality in advanced hormone-refractory prostate cancer (HRPC): Data from the Southwest Oncology Group Study S9916. *J Clin Oncol* 2004;**22**:4505.
20. Buyse M, Vangeneugden T, Bijnens L, et al. Validation of Biomarkers ad Surrogates for Clinical Endpoints. In: Bloom JC, editor. *Biomarkers in Clinical Drug Development*. New-York: Springer-Verlag; 2003.; 2003 p. 149–68.
21. Vicini FA, Vargas C, Abener A, Kestin L, Horwitz E, Martinez A. Limitations in the use of serum prostate specific antigen levels to monitor patients after treatment for prostate cancer. *J Urol* 2005;**173**:1456–62.
22. Dixon SC, Knowpf KB, Figg WD. The control of prostate-specific antigen expression and gene regulation by pharmacological agents. *Parmacol Rev* 2001;**53**:73–91.
23. Scher HI, Eisenberger M, D'Amico AV, et al. Eligibility and outcomes reporting guidelines for clinical trials for patients in the state of a rising prostate-specific antigen: Recommendations from the Prostate-Specific Antigen Working Group. *J Clin Oncol* 2004;**22**:537–56. (with erratum in J Clin Oncol 2004, 22, 3205).
24. Thalmann GN, Sikes RA, Change SM, Johnston DA, van Eschenbach AC, Chung LWK. Suramin-induced decrease in prostate-specific antigen expression with no effect on tumour growth in the LNCaP model of human prostate cancer. *J Natl Cancer Inst* 1996;**88**:794–801.
25. Schroder FH, Kranse R, Baret N, Hop WCJ, Kandra A, Lassus M. Prostate-specific antigen: a surrogate endpoint for screening new agents against prostate cancer? *Prostate* 2000;**42**:107–15.
26. Stadler W. New trial designs to assess antitumor and antiproliferative agents in prostate cancer. *Invest New Drugs* 2002;**20**:201–8.

27. Bubley GJ, Carducci M, Dahut W, et al. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: Recommendations from the Prostate-Specific Antigen Working Group. *J Clin Oncol* 1999;**17**:3461–7.
28. Royston P, Parmar M, Qian W. Novel designs for multi-arm clinical trials with survival outcomes with an application in ovarian cancer. *Stat Med* 2003;**22**:2239–56.
29. Verbel DA, Heller G, Kelly WK, Scher HW. Quantifying the amount of variation in survival explained by prostate-specific antigen. *Clin Canc Res* 2002;**8**:2576–9.
30. van Gils MPMQ, Stenman UH, Schalken JA, et al. Innovations in serum and urine markers in prostate cancer current European Research in the P-Mark project. *Eur Urol* 2005;**48**:1031–41.