PEER REVIEW HISTORY

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

(This paper received three reviews from its previous journal but only two reviewers agreed to published their review.)

ARTICLE DETAILS

TITLE (PROVISIONAL)	ETS-Related Gene (ERG) Expression as a Predictor of Oncological Outcomes in Patients with High-grade Prostate Cancer treated with Primary Androgen Deprivation Therapy: a cohort study
AUTHORS	Rezk, Mark; Chandra, Ashish; Addis, Daniel; Moller, Henrik; Youssef, Mina; Dasgupta, Prokar; Yamamoto, Hide

VERSION 1 – REVIEW

REVIEWER	Rebecca E. Graff University of California, San FranciscoUSA
REVIEW RETURNED	27-Jul-2018

GENERAL COMMENTS	This manuscript describes a cohort study evaluating associations between ERG expression and outcomes among prostate cancer patients treated with primary ADT. The question is an interesting one, but the study has substantial limitations. In addition, it could benefit from addressing the following: General: 1. Gene names (and TMPRSS2:ERG) should be italicized. Abstract: 2. Given that the authors speak to a National Cancer Registry, they should specify the country in which the study was conducted. 3. The authors report a significant association between ERG positivity and Caucasian ethnicity, but the reference group is unclear. 4. Given the objective of the study, the results should focus on the association between ERG status and outcomes. Rather than report that Gleason score and metastatic status were associated with prostate cancer specific survival (and that age was associated with overall survival), the authors should address the null results for ERG status. In doing so, they should be clear about whether the Cox models described were uni- or multivariable. 5. The second sentence of the Conclusions strikes me as much too strong given the limited sample size of the study. That the authors didn't find an association, doesn't necessarily mean that one doesn't exist. Strengths and Limitations:

	6. It seems to me that the limitations of the study aren't adequately addressed. Sample size and lacking covariates come to mind. Introduction: 7. It is my understanding that the 5' untranslated end of TMPRSS2 is binds to the 3' end of ERG. 8. I don't consider TMPRSS2:ERG to be a genetic mutation so much as an alteration or aberration. 9. Is the prognostic value of TMPRSS2:ERG really still under debate? Perhaps in certain patient / treatment populations, but I think it has largely been established that the fusion isn't associated with worse outcomes (at least in those treated with radical prostatectomy). 10. I don't believe that there's any such thing as the TMPRSS2:ERG fusion protein. 11. What does it mean for ERG expression to be re-established in ERG-positive tumors after the development castration resistance? ERG isn't overexpressed throughout the clinical course? 12. The authors should be careful in their wording of the study objectives. Are they aiming to establish an association or understand whether one exists? Methods: 13. The authors speak to the exclusion of patients due to missing data. Could such an exclusion induce bias? 14. The Statistical Methods section needs to be more robust. How were the Table 1 p-values calculated? Was there a multivariable evaluation of ERG status with respect to outcomes? What variables were included in the Cox models? Etc. Results: 15. The Methods indicate that an H-score of 0-50 resulted in an ERG-negative classification, but Figure 1 shows an ERG-positive individual with an H-score seemingly <50. 16. The authors might consider a trend test for some of the variables in Table 1. 17. Figure 2 and its associated description in the manuscript strike me as somewhat beside the point of the study. So too with the description of the Table 2 results. 18. Log-rank evaluations don't account for the covariates that are likely to be important in the associations of interest. Discussion: 19. The first paragraph of the Discussion should summarize the important findings of the study. 20. Again, I don't know that the term "mutation" is appropriate. 21. Perhaps more importantly than the high prevalence of ERG expression in all grades of disease, the largest of studies haven't shown an association between ERG expression and disease outcomes. Regardless, rather than focus on the (rather settled) debate about the association between ERG expression and outcomes overall, the authors should speak to their hypothesis that ERG expression could be associated with outcomes in their particular patient population. 22. The authors should be careful to avoid using the term "effect". They've only evaluated associations. 23. The first paragraph of the Limitations indicates that the patient population wasn't as described in the Methods. This point should be surfaced earlier. In addition, the authors should run sensitivity analyses in the patient population that truly received only ADT.

	24. The authors need to discuss the limitations of their sample size. They could even consider calculating the power they had to find a reasonable magnitude of association. 25. The authors should discuss potentially important covariates for which they did not have data. Stage at diagnosis, for example, has been shown to be associated with ERG status and with prostate cancer outcomes. Conclusion: 26. The value of the future study described is unclear. 27. Given their limited sample size, the authors should soften their conclusions.

REVIEWER	Jennifer Cullen HJF-USU
REVIEW RETURNED	15-Aug-2018

GENERAL COMMENTS	The authors contribute novel information in this study, in which they examine whether there is prognostic value of ERG in predicting prostate cancer (PCa) progression or death in a cohort of men who underwent primary ADT as monotherapy for treatment of high risk PCa (metastatic and non-metastatic). A key strength is use of a national registry to ascertain vital status. The paper is well written and the discussion section offers meaningful insight into how these data contribute uniquely to the field, while clearly recognizing the study's limitations. Table 1 should included the number of deaths (all causes and prostate cancer specific), if possible. Could age-adjusted KM curves for OS be provided? It is an interesting finding that Ethnicity does not predict any study outcome. Was comparability anticipated?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Rebecca E. Graff

ABSTRACT:
1. Gene names have now been italicized throughout the document
2. Added 'UK National Cancer Registry' pages 2 and 3.
3. We have now omitted this statement as its effect was not seen on multivariate analysis.
4 & 5. These points have been amended on the abstract.

STRENGTHS AND LIMITATIONS:
6. An additional limitation has been added under the appropriate subheading.

INTRODUCTION:
7. Corrected on page 4 of the manuscript
8. We agree with this point and has been amended
9 & 10. Please see amended statements on page 4 of the manuscript
11. ERG had been shown to persist following the development of castrate resistance. Please see amendment on page 4.
12. Amendment on page 5 of manuscript.

PATIENTS, MATERIALS AND METHODS:
13. We agree that missing data could introduce bias. Exclusion of patients in our dataset occurred only in a small proportion (20 out of 169, as stated in results) and hence did not perform correction for missing data which would introduce its own bias. This limitation has been added to the discussion.
14. The table 1 p values were calculated using the Log-rank method, as amended. Further information regarding the multivariate analysis has been provided in the statistical methods section.

RESULTS:
15. There is 1 patient with a negative H score between 0-50. To prevent misunderstanding, we have altered the figure to show only ERG positive patients i.e. those with an H score >50. The H score for the 51 patients has also been represented in supplementary figure 4.
16. We considered to include trend analysis however, felt that in the context of a small sample size may yield inconclusive results. Nevertheless, we have indicated this graphically for Gleason score and PSA in figure 2.
17. We used figure 2 to indicate the likely survival outcomes of hormone-only treated prostate cancer patients in the UK which is rarely reported for this particular geographical location. Table 2 indicates the relative impacts of the other covariates in determining survival in relation to ERG expression, which we felt was useful.
18. We entirely agree with the reviewer. In this study Cox regression was conducted using known co-variates associated with prostate cancer progression which were also reliably recorded in clinical notes.

DISCUSSION:
19. The paragraphs have been rearranged in the discussion according to the preferential order to include points 19 and 21. Please note that the order of references has subsequently been changed and all changes are highlighted with the track changes tool.
20. Amended in the discussion section also.
21. Paragraph 3 of the discussion section has been added to summarise important findings of ERG expression in prostate cancer patients treated with androgen deprivation therapy. A brief summary of the Pettersson meta-analysis is provided in the following part of our discussion.
22. This point has also been amended
23. The fact that patients may have received unplanned adjuvant therapy has now been noted in the data collection section of our methods. Unfortunately, whilst we strongly agree with the reviewer that this is a limitation of our study, we are unable to obtain this data due to incomplete follow-up. Some of the patients were regional referrals and therefore did not have local follow-up. Therefore, we are unable to run sensitivity analysis on this specific point.
24. A statement regarding this limitation has been added in our discussions and in the ‘strengths and limitations’ section following the abstract. As we have utilised the maximum number of patients we had access to in our institutions, we did not undertake any power calculation. However, we had commenced the study as an exploratory investigation.
25. Please see added statement regarding collection of other covariates.

CONCLUSION:
26 and 27. The Conclusion has been adjusted accordingly for greater clarity.

Reviewer: 2
Reviewer Name: Jennifer Cullen

We have amended Table 1 appropriately to display the number of deaths in our study population. Although age adjusted survival curves can be provided, it was not the only variable that differed between the ERG positive and negative subgroups. Therefore, we felt that a better representation of age was provided using multivariate analysis

Ethnicity in the UK is less likely to predict outcome given the equality of healthcare in the National Health Service across all socioeconomic classes and ethnicity. This paper shows that in the United States, African-American men have a 60% higher risk for developing prostate cancer with poorer prognosis than in their white counterparts. This was not observed in the UK population.
Kheirandish, P., & Chinegwundoh, F. (2011). Ethnic differences in prostate cancer. British Journal of Cancer, 105(4), 481–485. http://doi.org/10.1038/bjc.2011.273

VERSION 2 – REVIEW

REVIEWER	Rebecca Graff UCSF, USA
REVIEW RETURNED	29-Oct-2018

GENERAL COMMENTS	Note that while gene names should be italicized, protein names should remain unitalicized. The authors otherwise addressed my comments adequately.

REVIEWER	Jennifer Cullen CPDR-USU, Rockville, MD, USA
REVIEW RETURNED	31-Oct-2018

GENERAL COMMENTS	The authors have addressed all of the concerns raised.