PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Risk of venous thromboembolism in men with prostate cancer compared with men in the general population: A nationwide population-based cohort study in Sweden
AUTHORS	Balabanova, Yanina; Farahmand, Bahman; Garmo, Hans; Stattin, Pär; Brobert, Gunnar

VERSION 1 – REVIEW

REVIEWER	Skajaa, Nils Aarhus University Hospital, Department of Clinical Epidemiology
REVIEW RETURNED	19-Sep-2021

GENERAL COMMENTS	Balabanova et al used Swedish registry-based data to investigate the association between prostate cancer and venous thromboembolism using a matched general population comparison cohort. The authors found that prostate cancer increased the hazard of VTE approximately 1.5-fold over the course of follow-up; the absolute risk ratios were particularly elevated short after diagnosis. Overall, the study is well-conducted and written, although the study lacks novelty. - The authors mention that several previous studies have been conducted on this topic. Could the authors elaborate on how this study adds to the literature? The authors mention that contemporary data is needed, but why? What has happened over time that may have impacted effect estimates? - What is the validty (eg. the positive predictive value) of the prostate cancer and VTE (both DVT and PE) diagnoses? This should be stated more clearly. - There seems to be some inconsistency between using either primary diagnoses only or both primary/secondary diagnoses. Could this be made more consistent? If not, further explanation is warranted. - What happened if a DVT and PE occurred on the same date? - The crude calculation of incidence proportions (ie, risks) is appropriate; however, the authors should consider computing the cumulative incidence functions, considering the competing risk of death. This estimate is a more meaningful absolute risk estimate for the individual. - Covariates included in the Cox model was chosen on a subjective basis. This needs to be elaborated further. I suggest the authors create a DAG and include covariates based on that. - Previous cancer was not an exclusion criteria, in contrast to most previous studies. Why was this not the case in the current study? I suggest, at least, a sensitivity analysis excluding previous cancer.

REVIEWER	Cannegieter, Suzanne

	Leiden University Medical Centre, Leiden, Department of Thrombosis and Hemostasis
REVIEW RETURNED	22-Sep-2021

GENERAL COMMENTS	This Swedish nationwide cohort study was designed to estimate the additional risk of venous thromboembolism (VTE) in men with prostate cancer compared with men without prostate cancer. More than half a million men with and without prostate cancer (1:5) were included and followed from 2007 for 5 years and a 50% higher VTE risk was found in the cancer group. This is a well performed study in large and unselected data which is therefore informative considering that a large number of studies on this topic is already available. I have a few comments and questions: The incidence of VTE in the non-cancer cohort is rather high, at more than 4 per 1000 person-years, whereas this generally is found to be about 2 per 1000 person-years in most registries. Can the authors explain this, and could this have affected the relatively small increased risk for the cancer-cohort? What is the rationale for using incidence proportions as a measure for the absolute risk in addition to the incidence rates? The problem with this measure is that it doesn't take different follow-up durations into account, nor does it account for the competing risk of death. It would be more insightful, also visually, to provide cumulative incidences using the life-table technique (ie Kaplan Meier), but adjusted for the competing risk of death. Figure 2 should therefore also be replaced by a Kaplan-Meier curve. Likewise, providing incidence proportion ratios is not very informative and can be left out, also considering that two other effect measures are already in place, i.e., IRRs and HRs (here I would also suggest to give only unadjusted and adjusted HRs, to keep it simple). Relate to this: the text states that "Incidence proportions over 60 months' follow-up increased in both cohorts over time, reflecting increased VTE incidence with age irrespective of cancer status". This is not correct, the incidence does not increase with age but with time, as the cases accumulate over time. Can the authors explain why the risk appears to remain increased over 5 years (although this is difficult to infer from the currently given incidence proportions)? When the cancer is cured after let's say 6 months, an additional risk is not to be expected anymore? How many patients were first in the control cohort and later switched to the cancer cohort? Were person-years in either group counted for the incidence rates? As a limitation the authors describe that there may be residual confounding due to factors like on height, weight, smoking status and alcohol intake. However, these are hardly risk factors for VTE (with the exception of BMI) and prostate cancer, as far as I am aware. Anyway, confounding does not appear to be a big issue here, from a theoretical point (prostate cancer and VTE do not have many risk factors in common except for age) and the HRs hardly change upon adjustment.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1. Dr. Nils Skajaa, Aarhus University Hospital

Comments to the Author:

Balabanova et al used Swedish registry-based data to investigate the association between prostate cancer and venous thromboembolism using a matched general population comparison cohort. The authors found that prostate cancer increased the hazard of VTE approximately 1.5-fold over the course of follow-up; the absolute risk ratios were particularly elevated short after diagnosis. Overall, the study is well-conducted and written, although the study lacks novelty.

Comment 1. The authors mention that several previous studies have been conducted on this topic. Could the authors elaborate on how this study adds to the literature? The authors mention that contemporary data is needed, but why? What has happened over time that may have impacted effect estimates?

Authors response: There are several reasons why more contemporary data are needed. Firstly, the case mix of men with prostate cancer has changed dramatically since the early noughties, as shown by the PSA distribution at diagnosis.(1) Secondly, the pattern of care for men with prostate cancer has changed dramatically over this time.(1) Thirdly, there have only been a few previous studies on this topic, and these have all been conducted during time periods that pre-date the introduction of direct oral anticoagulants (DOACs), which have become increasingly adopted in Swedish clinical practice over the last decade, and which could potentially be associated with VTE risk. The first DOAC was introduced in Sweden in 2008, and our large population-based study, where the start of follow-up was from 2007–2016 and the end of follow-up was up to December 2017, covered years in which DOACs were used in Sweden. Finally, in 2010, we reported on risk of VTE among 75,000 men diagnosed with prostate cancer between 1997 and 2007 and calculated standardised incidence rates to compare these risks with those seen in the general male population of Sweden. At that time, we did not have a matched control cohort; (2) however, since then, we have reported on the risk of VTE in men according to the specific treatment(s) received.(3-5) We argue that our current study based on a larger sample of men with prostate cancer, a matched control cohort with comprehensive data on comorbidities and drug use allows us to assess risk estimates with higher precision than in our first study. We have added some text to cover these points at the end of page 5/start of page 6 in our manuscript introduction.

Comment 2. What is the validty (eg. the positive predictive value) of the prostate cancer and VTE (both DVT and PE) diagnoses? This should be stated more clearly.

Authors response: Unfortunately, information on the accuracy of the ICD-10 codes for VTE used in clinical practice across Sweden are lacking, and we were unable to validate the VTE diagnoses because the results of imaging procedures are not routinely recorded in the Patient Register. This is a limitation of our study, and we have added some text to the Discussion section of our manuscript (page 13) to cover this point.

Comment 3. There seems to be some inconsistency between using either primary diagnoses only or both primary/secondary diagnoses. Could this be made more consistent? If not, further explanation is warranted.

Authors response: We excluded all men with a primary or secondary diagnosis in the National Patient Register of VTE in 10 years prior to cancer diagnosis to increase the certainty that only men with a first-ever VTE were identified during follow-up (as previous VTE is a risk factor for a subsequent VTE). To clarify this to the reader, we have now added text to cover this in the 'Study cohort and outcome follow-up' section of the Methods (page 7). However, for VTE during follow-up (the outcome of interest in the study), we chose to limit cases to those where the VTE was a primary diagnosis to increase case validity and to avoid double counting of initial primary cases (because a code for VTE could be entered at follow-up visits relating to the initial VTE event.

Comment 4. What happened if a DVT and PE occurred on the same date?

Authors response: This was not an issue we had to deal with because no man in our study had a record of DVT and PE as a main diagnosis on the same date.

Comment 5. The crude calculation of incidence proportions (ie, risks) is appropriate; however, the authors should consider computing the cumulative incidence functions, considering the competing risk of death. This estimate is a more meaningful absolute risk estimate for the individual.

Authors response: The reason we chose to calculate incidence proportions without also calculating cumulative incidence functions, which considers the competing risk of death, is because of the relatively low risk of both VTE and mortality (during the first year). Therefore, estimations of the cumulative incidence function that take into account the competing risk of death would only minimally change the results, as illustrated by Häggström *et al*.(6)

Comment 6. Covariates included in the Cox model was chosen on a subjective basis. This needs to be elaborated further. I suggest the authors create a DAG and include covariates based on that.

Authors response: The covariates kept in the Cox model were those chosen on a subjective basis (those deemed to be clinically meaningful), as opposed to any pre-defined arbitrary cut-off. Very few variables altered the crude HR, and those that did, did so very minimally. Overall, very little confounding was seen in our study, and we adjusted for all potential confounders that we have information on. If the editor prefers, we would be happy to present age-adjusted HRs only. We agree that it may have been beneficial to produce a DAG during the design stage of the study, but not now that the study has been completed.

Comment 7. Previous cancer was not an exclusion criteria, in contrast to most previous studies. Why was this not the case in the current study? I suggest, at least, a sensitivity analysis excluding previous cancer.

Authors response: As the objective of the study was to compare the incidence of VTE between men living with prostate cancer and men from the general population in Sweden, we did not feel there was a strong rationale for excluding men with previous cancer. Also, this was partly the reason we did not aim to strictly define VTE cases during follow-up as cancer-associated thrombosis but rather interpreted them as VTE events that occurred in men living with prostate cancer.
We acknowledge, however, that a different study objective may have warranted exclusion of these men. Instead, we included previous cancer as a potential confounder, and we found that it had no material effect on the risk estimates observed when included in the regression models.

Reviewer 2. Dr. Suzanne Cannegieter, Leiden University Medical Centre, Leiden, Leiden University Medical Centre

Comments to the Author:

This Swedish nationwide cohort study was designed to estimate the additional risk of venous thromboembolism (VTE) in men with prostate cancer compared with men without prostate cancer. More than half a million men with and without prostate cancer (1:5) were included and followed from 2007 for 5 years and a 50% higher VTE risk was found in the cancer group.
This is a well performed study in large and unselected data which is therefore informative considering that a large number of studies on this topic is already available.
I have a few comments and questions:

Comment 1. The incidence of VTE in the non-cancer cohort is rather high, at more than 4 per 1000 person-years, whereas this generally is found to be about 2 per 1000 person-years in most registries. Can the authors explain this, and could this have affected the relatively small increased risk for the cancer-cohort?

Authors response: The likely explanation for the higher incidence in our general population comparison cohort (when compared with that seen in most registries) is that this cohort was predominantly composed of older men, which was a result of obtaining an age-matched cohort to the prostate cancer cohort. The median age of men in both cohorts was 69 years with an interquartile range of 64–76 years. In the study by Walker *et* ,(7) the incidence rate of VTE was 9.1 per 1000 person-years and the relative risk was 2.6, thus the incidence rate among in the general population was about 3.5 per 1000 person-years. Correspondingly, in the study by Cronin-Fenton *et al,*(8) the incidence rate in the general population cohort was 3.1 per 1000 person-years. In these studies, as in our study, VTE was identified in the same way in both the cancer and comparator cohorts.

Comment 2. What is the rationale for using incidence proportions as a measure for the absolute risk in addition to the incidence rates? The problem with this measure is that it doesn't take different follow-up durations into account, nor does it account for the competing risk of death. It would be more insightful, also visually, to provide cumulative incidences using the life-table technique (ie Kaplan Meier), but adjusted for the competing risk of death. Figure 2 should therefore also be replaced by a Kaplan-Meier curve. Likewise, providing incidence proportion ratios is not very informative and can be left out, also considering that two other effect measures are already in place, i.e., IRRs and HRs (here I would also suggest to give only unadjusted and adjusted HRs, to keep it simple).
Relate to this: the text states that "Incidence proportions over 60 months' follow-up increased in both cohorts over time, reflecting increased VTE incidence with age irrespective of cancer status". This is not correct, the incidence does not increase with age but with time, as the cases accumulate over time.

Authors response: The objective of the study was to determine the risk an individual man with prostate cancer has in developing a VTE within a 5-year time period, and we believe this is best expressed by an incidence proportion. We believe that a Kaplan–Meier survival analysis would add little, especially because we also present HRs censoring for death. If the Editor prefers, we are happy to remove incidence proportion ratios from the results. We chose to calculate incidence proportions without also calculating cumulative incidence functions, which considers the competing risk of death due to the relatively low risk of both VTE and mortality (during the first year). Therefore, estimations of the cumulative incidence function that take into account the competing risk of death would only minimally change the results, as illustrated by Häggström *et al*.(6)
On the point of presenting IRRs and HRs, the IRRs previously presented were in fact the crude HR, and we have amended this (removing any mention of IRRs) in the 'Statistical Analysis' section of the Methods (page 8) and in the Results (page 10 as well as Table 2 and Supplementary Table 5). We acknowledge the reviewer's statement that the following sentence on page 9 of our manuscript may not be entirely accurate:

> "Incidence proportions over 60 months' follow-up increased in both cohorts over time, reflecting increased VTE incidence with age irrespective of cancer status."

We have changed this sentence to read as follows:

> "Incidence proportions over 60 months' follow-up increased in absolute terms in both cohorts over time and decreased over time in relative terms."

Comment 3. Can the authors explain why the risk appears to remain increased over 5 years (although this is difficult to infer from the currently given incidence proportions)? When the cancer is cured after let's say 6 months, an additional risk is not to be expected anymore?

Authors response: This finding was not surprising to us. The likely explanation for why the risk appears to be increased over the 5-year period is that the prostate cancer cohort was composed of men of advanced age (median age 69 years, interquartile range of 64–76 years). While they may have been a small proportion of men who were cured, many will still be living with the disease and managed with hormonal therapy. Among the latter, the cancer may have progressed over time, and other comorbidities may also have developed or worsened over time, both of which could influence the risk of VTE. Furthermore, it is worth noting that while we observed a relative risk for VTE of 1.5 over 5 years, Walker *et al* (7) observed a relative risk of 2.6 over an average follow-up of approximately 4 years, and Cronin-Fenton *et al* (8) observed a relative risk of approximately 3.3 over an approximate 2-year follow-up.

Comment 4. How many patients were first in the control cohort and later switched to the cancer cohort? Were person-years in either group counted for the incidence rates?

Authors response: In the cohort of men free of prostate cancer, a total of 3% (n=18,917) were diagnosed with prostate cancer during follow-up. These men were censored at the date of prostate cancer diagnosis and did not subsequently join the prostate cancer cohort (and so did not contribute person-time to that cohort).

Comment 5. As a limitation the authors describe that there may be residual confounding due to factors like on height, weight, smoking status and alcohol intake. However, these are hardly risk factors for VTE (with the exception of BMI) and prostate cancer, as far as I am aware. Anyway, confounding does not appear to be a big issue here, from a theoretical point (prostate cancer and VTE do not have many risk factors in common except for age) and the HRs hardly change upon adjustment.

Authors response: We agree that the variables mentioned are not obvious potential confounders to be considered in this study, yet to cover the important point in all observational studies that residual confounding could still be present, we have now modified the relevant sentence in the Discussion to read as follows:

> "…however, lack of adjustment for unknown confounders may have led to residual confounding" (page 13).

Indeed, as the reviewer mentions, confounding was not a big issue in this study.

References

1. National Prostate Cancer Registry (NPCR) Report for 2019 (Published September 2020). https://npcr.se/wp-content/uploads/2020/09/ 20200907_npcr_ nationell_ rapport_ 2019.pdf
2. Van Hemelrijck M, Adolfsson J, Garmo H, Bill-Axelson A, Bratt O, Ingelsson E, Lambe M, Stattin P, Holmberg L. Risk of thromboembolic diseases in men with prostate cancer: Results from the population-based PCBaSe Sweden. Lancet Oncol. 2010;11(5):450-8.
3. Bosco C, Garmo H, Adolfsson J, Stattin P, Holmberg L, Nilsson P, Gunnlaugsson A, Widmark A, Van Hemelrijck M. Prostate Cancer Radiation Therapy and Risk of Thromboembolic Events. Int J Radiat Oncol Biol Phys. 2017 Apr.
4. O'Farrell S, Sandström K, Garmo H, Stattin P, Holmberg L, Adolfsson J, Van Hemelrijck M. Risk of thromboembolic disease in men with prostate cancer undergoing androgen deprivation. BJU Int. 2015 Oct 24.
5. O'Farrell S, Sandström K, Garmo H, Stattin P, Holmberg L, Adolfsson J, Van Hemelrijck M. Risk of thromboembolic disease in men with prostate cancer undergoing androgen deprivation. BJU Int. 2015 Oct 24.
6. Häggström et al. Interpretation of conventional survival analysis and competing-risk analysis: an example of hypertension and prostate cancer. BJU Int 2016;118(6):850–852.
7. Walker et al. Incidence of venous thromboembolism in patients with cancer - a cohort study using linked United Kingdom databases. *Eur J Cancer* 2013;49(6):1404–13.
8. Cronin-Fenton et al. Hospitalisation for venous thromboembolism in cancer patients and the general population: a population-based cohort study in Denmark, 1997-2006. *Br J Cancer* 2010;103(7):947–53.

VERSION 2 – REVIEW

REVIEWER	Skajaa, Nils Aarhus University Hospital, Department of Clinical Epidemiology
REVIEW RETURNED	Skajaa, Nils Aarhus University Hospital, Department of Clinical Epidemiology

GENERAL COMMENTS	Overall, the authors have adequately answered my initial points. However, I have a few comments: 1. The authors now describe the lack of validation of the VTE diagnosis in the limitations section. Could the authors add 1-2 sentences about the presumed direction of any misclassification bias? 2. If the authors wish to keep presenting incidence proportions, they should consider at least presenting mortality risks as well? Without an understanding of the magnitude of mortality, the effect estimates are difficult to contextualize.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1: Overall, the authors have adequately answered my initial points. However, I have a few comments:

Comment 1. The authors now describe the lack of validation of the VTE diagnosis in the limitations section. Could the authors add 1-2 sentences about the presumed direction of any misclassification bias?

Author response: To address this point, we have added the following sentence to this section on the discussion of our revised manuscript (page 13):

> "The likely under-recording of VTE cases (not capturing those whose imaging results were unobtainable or those who were asymptomatic) would not affect the HRs under the assumption of non-differential misclassification. If, hypothetically, there was an over-recording of VTE cases (inclusion of false positives), this would bias the HRs towards the null, assuming non-differential misclassification."

Comment 2. If the authors wish to keep presenting incidence proportions, they should consider at least presenting mortality risks as well? Without an understanding of the magnitude of mortality, the effect estimates are difficult to contextualize.

Author response: Aside from mortality being beyond the scope of the specific study objectives, it is well-established that men with prostate cancer, on average, have an increased risk of death compared with men in the general population, albeit this is strongly dependent on prostate cancer risk category.[1,2] Furthermore, please note that date of death was taken into account as a censoring criterion in our study.

References

1. Rider et al. Long-term outcomes among non-curatively treated men according to prostate cancer risk category in a nationwide, population-based study. Eur Urol. 2013 Jan;63(1):88-96. doi: 10.1016/j.eururo.2012.08.001. Epub 2012 Aug 10.
2. Van Hemelrijck et al. Causes of death in men with localized prostate cancer: a nationwide, population-based study. BJU Int. 2016 Mar;117(3):507-14. doi: 10.1111/bju.13059. Epub 2015 May 15.