Endemic statistical paradoxes in epidemiologic studies distort knowledge on prostate cancer: mitigation and caution of fallacies in prostate cancer causal epidemiological studies

Olivier Cussenot[a,b], *Gaelle Fromont*[b,e], *Géraldine Cancel-Tassin*[b], *Freddie C. Hamdy*[a] *and Richard M. Martin*[c,d]

Purpose of review
Many studies on epidemiology of prostate cancer (PCa) are based on a diagnosis of PCa using PSA (prostate-specific antigen) level. However, biases can distort the interpretation of the results, which in turn limits policy and decision making on public health prevention strategies or clinical guidelines. The main confusion is to interpret the posterior probability of the outcome following the exposure as a change in the prevalence of the disease outcome, whereas this change reflects only the predictive values of the PSA test induced by the exposure of interest.

Recent findings
Many studies report potential causal factors involved in PCa risk. However, the lack of integration of how physiological changes in PSA values are associated with the exposures being investigated, they explain in part contradictory and controversial results on PCa risk factors in the literature.

Summary
A strategy to perform case–control studies based on PSA stratification is suggested to avoid misinterpretation related to PSA misclassification. Real data are analysed, and we show that we can exploit the mechanism of selection biases using different modalities of controls recruitment based on biomarker stratification to distinguish real from false causal factors.

Keywords
bias, causal epidemiology, prostate cancer, prostate-specific antigen, risk factors

INTRODUCTION

There were 1.4 million new cases of prostate cancer (PCa) reported worldwide in 2020, making it the second commonest male cancer globally. Large variation in the incidence of PCa worldwide reflects multiple factors, including differences in the use of diagnostic testing, exposure to various environmental factors and genetic background. Understanding the cause of PCa, and factors that influence its progression, are major public health priorities and rely mainly on well conducted epidemiological studies to prioritize interventions for testing in randomized controlled trials or to inform policy and clinical practice in the absence of feasible trials.

The identification of PCa cases and cancer-free controls in case–control studies of the cause of PCa rely on diagnostic strategies based on blood prostate-specific antigen (PSA) levels. Thresholds (e.g. PSA values of 3 or 4 ng/ml) are used to decide who undergoes a prostate multiparametric (mp) MRI or biopsy, but it is known that up to 25% of men with values below these thresholds have PCa [1,2]. However, it is not ethical or feasible to systematically undertake prostate biopsies in men below these PSA

[a]Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK, [b]CeRePP, Paris, France, [c]Population Health Sciences, Bristol Medical School, University of Bristol, [d]NIHR Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol, Bristol, UK and [e]Department of Pathology, CHRU Tours, Tours, France

Correspondence to Olivier Cussenot, Nuffield Department of Surgical Sciences, University of Oxford, Level 6, Room 6600, John Radcliffe Hospital, Oxford OX3 9DU, UK. E-mail: olivier.cussenot@nds.ox.ac.uk

Curr Opin Urol 2023, 33:000–000
DOI:10.1097/MOU.0000000000001119

KEY POINTS

- Studies of causal factors in prostate cancer risk are often contradictory and controversial.
- In prostate cancer causal epidemiology studies, selection bias of controls based on PSA testing is the most common limitation, despite complex statistical designs, including Mendelian randomization.
- Understanding selections using directed acyclic graphs as and use of genetic markers can mitigate misinterpretation about causal risk factors.
- The deliberate introduction into the study of different opposing modalities for the selection of control populations and the use of genetic causal factors makes it possible to highlight erroneous causal factors and reduce misinterpretation.

thresholds and it is not effective to identify cases and controls using other methods, such as (mp) MRI on a large scale [3]. In some studies, the control group is based on low PSA levels without biopsy or imaging to exclude PCa. With this modality of selection of the control, PSA blood level is sometimes adjusted on age. In other studies, controls are defined by negative biopsies following an increase of PSA. Biopsy is then justifies ethically, on the medical side. A third method consists to random controls from frequency-matched to cases by age range in general men population [4].

Physiological PSA production by luminal cells of prostatic glands is androgen dependent. Factors impacting the androgen receptor pathway modulate changes in PSA secretion by prostate cells [5]. PSA secretion and consequently blood levels in healthy men are different according to various endogenous or exogenous factors [5]. Some factors increase PSA blood levels such as prostate volume, inflammation, African ancestry [6] or *PSA* gene variants that promote PSA secretion. Others decrease PSA blood levels: not only Asian ancestry [6], BMI [7,9], partial androgen deficiency or PSA gene variants that decrease PSA secretors [8] but also antiandrogens or drugs such as statins, thiazides, antiinflammatories [9] or potential environmental factors with endocrine perturbation effects [10].

DEFINE THE OUTCOME, THE EXPOSURE AND THE CLASSIFIER

If the PSA blood test is used as diagnosis procedure, the endpoint or outcome observed is the posterior probability of the event evaluated such as PCa diagnosis on biopsy, aggressiveness or survival without progression and not the real changes in the prevalence of event measured. According to the Bayes theorem [11] (Probability PCa if PSA test positive = [Prevalence PCa × sensitivity PSA test]/ [Prevalence PCa × sensitivity PSA test] + [(1−prevalence PCa) × (1−specificity PSA test)], the posterior probability using a diagnosis test is a function of the prevalence of the disease (prior probability) and of the performance (sensibility and specificity) of the diagnosis procedure, which is the classifier.

As an illustration, the prevalence of PCa lesions changes according to the diagnostic procedure. This explains variations in the incidence of PCa, such as the changes in incidence before and after the PSA era, and the worldwide disparities in incidence related to the social access to PSA testing. This is also confirmed by systematic whole prostate analysis during postmortem autopsies in men who died of causes other than PCa. These report that after 50 years of age, over 30% of men have PCa lesions [12], whereas the prevalence of PCa lesions diagnosed on biopsy following PSA over 4 ng/ml is about 15% [13].

The direction of the interactions between outcome (such as result of prostate biopsy), exposure (risk or interventional factor) and classifier (marker like PSA) is presented in Fig. 1.

According to the Bayes Theorem, the posterior probability to have a cancer on biopsy if the test (PSA) is positive increases if the prevalence increases (causal effect of the risk factor on the disease) or if the performance of the test increases (no causal effect of the risk factor on the disease, but increasing of the test's positive predictive value). Introduction of instrumental variable to perform Mendelian randomization does not avoid the selection bias.

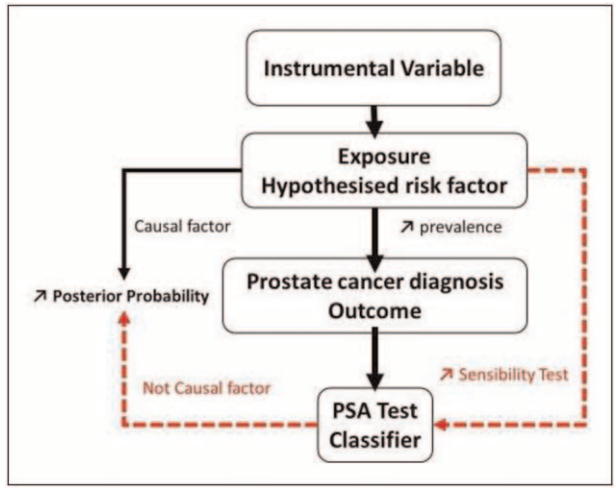

FIGURE 1. Directed acyclic graph for causal factors.

Mendelian randomization in prostate cancer causal epidemiology

Mendelian randomization is an analytic approach that utilizes genetic variation [single-nucleotide polymorphisms (SNPs)] as a randomized proxy (instrument) of the exposure of interest to provide insights into causality. Large-scale genomic information from published genome-wide association studies (GWAS) during the past decade have provided a very large number of quantitative trait loci (eQTLs; https://www.ebi.ac.uk/eqtl/) that can be used to perform Mendelian randomization studies [14]. Mendelian randomization uses germline [12] genetic variants as instrumental variables to limit some common epidemiological biases, assuming that such variants are randomly distributed with respect to postnatal confounders. To be valid Mendelian randomization analysis requires three assumptions: the genetic variants need to be associated with the exposure; the genetic variants must not be associated with confounders of the exposure-outcome association; and the genetic variants must be independent of the outcome and conditional on the exposure. In the last 10 years, there have been more than 100 publications (60% in the last 3 years) that have used Mendelian randomization to appraise potential causal risk factors for PCa. However, Mendelian randomization cannot avoid selection biases because the instrumental variable is upstream to the suspected risk factor (illustrated by the MCAR genotype and BMI in Figs. 1 and 3).

Selection bias and prostate cancer studies

The commons selection bias related to case–control studies or routine data analysis is Berkson's bias (Collider bias) (Fig. 2a). Berkson's selection bias, described for 70 years, remains present in clinical reports and has been recently pointed out on studies related to COVID-19 [15▪].

In this example of a population of eight individuals: four blue dots have a survival of 80% and four red dots 40% at time T. The individuals are stratified into risk groups according to the X PSA level. Above X, the mean survival of the poor prognosis group is $[(3*40\% + 80\%)/4 = 50\%]$ at time T and under X the good prognosis group with an average survival at $[3*80\% + 40\%]/4 = 70\%$). If we compare this population of eight individuals with an identical population exposed to a factor, which reduces the PSA level by 2, and has no effect on the aggressiveness of the cancer. Using the same X cut-off (ignoring the biological effect of the factor on PSA), we observed a reduction in survival in the two

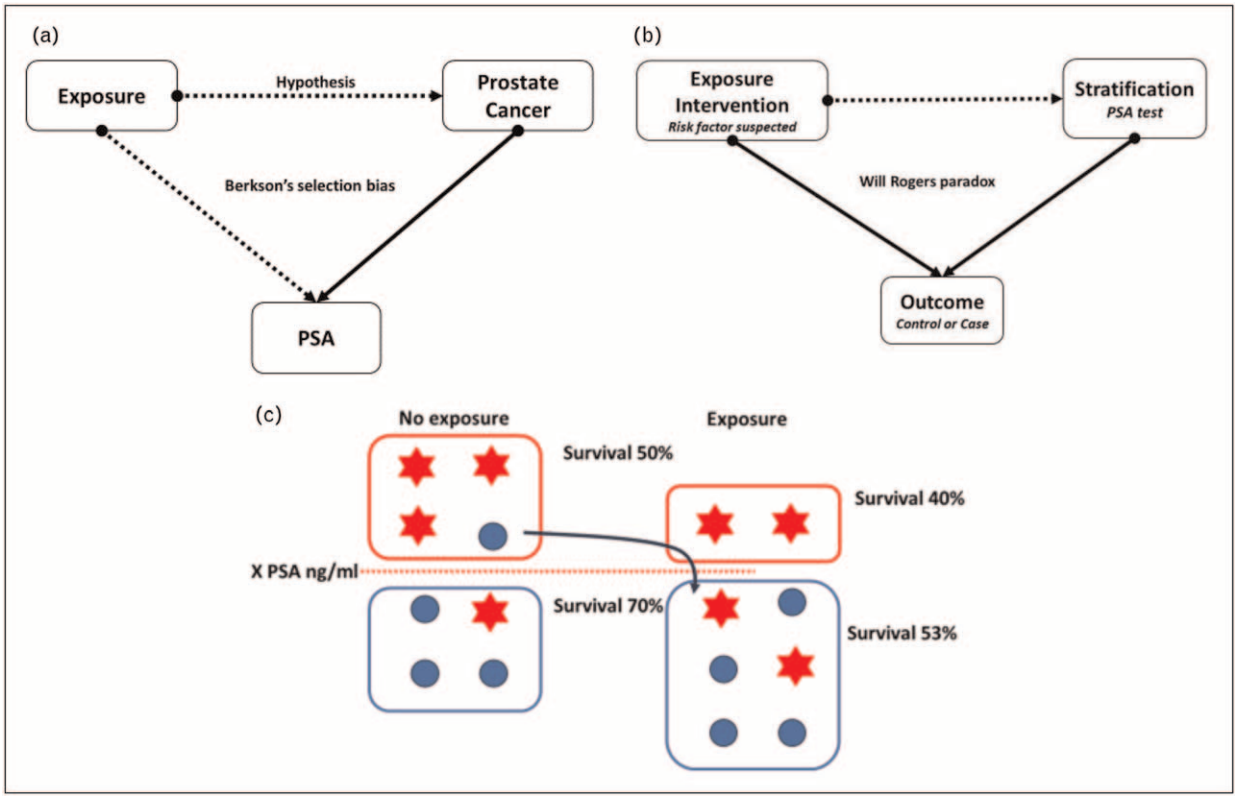

FIGURE 2. Selection biases. (a) Collider bias or Berkson's selection bias and (b) Will Rogers paradox. (c) Theoretical illustration of the Will Rogers paradox (stages migration).

groups with good prognosis [(3∗80% + 2∗40%)/6 = 53%] and poor prognosis [(2∗40%)/2 = 40%] compared with the unexposed population. The change in stratification may lead to the erroneous conclusion that exposure increases the aggressiveness of the cancer.

Applied to PCa, the bias is illustrated by the situation wherein the exposure is associated with changes in PSA levels. If the exposure (risk factor) changes the performance of the stratification method, which define cases and controls, a statistical paradox (Will Rogers paradox) occurs (Fig. 2b). Independent of PCa, prostate volume increase PSA blood level, so the increase of prostate volume increases the probability of finding a PCa on biopsy if biopsy is performed only for men with elevated blood PSA values. On the contrary, men with a high PSA and a high prostate volume have a decreased probability of being diagnosed with PCa if biopsy is performed. An inverse situation is observed for factors such as high BMI, which are associated with a decrease in blood PSA levels. This effect has also been reported for genetic polymorphisms regulating the expression of the PSA. The Will Rogers paradox is met in interventional studies when the exposure/intervention changes the classifier [16,17] (Fig. 2b).

In case–control or cohort studies, an erroneous classification of the case group and the control group may lead to a biased interpretation of the associations between potential risk factors and the disease (Fig. 2). In studies on PCa, this misclassification appears if the risk factor studied influences, for example, the level of PSA, is used to stratify the outcome [18]. The direction of the association between exposure and cancer reverses when the sensitivity of the test (PSA) in the exposed group is greater than in the nonexposed group. This misclassification can change both the size and direction of an association. It has been demonstrated in routine data and in case–control studies [18], suggesting the utmost caution in the interpretation of the results when an exposure factor modifies the parameter used for the classification of the study group and the reference group. This has been illustrated, among other things, by prevention studies using finasteride [19]. Finasteride has an antiandrogenic effect such as estrogenic endocrine disruptor, which lowers PSA levels. This biological effect 'paradoxically' leads to an increase in the detection of the disease or the most progressive forms by the PSA test (higher sensitivity of the PSA test to detect the disease and its severe forms in the exposed group). Finasteride improves the performance of the PSA test because it lowers the PSA level less if cancer is present, especially if the cancer is aggressive. A man with a PSA above the threshold that indicates biopsies is more likely to have PCa in the exposed group than in the unexposed group. Thus, the PSA test has a better performance in identifying men with PCa in the finasteride group. Similarly, in prognostic terms, a man with PCa classified as low risk of progression with a risk classification (determining the type of treatment) is more likely to have aggressive PCa (recurrence after treatment, for example) in the exposed group than in the unexposed group. Indeed, finasteride causes a relative enrichment (suppression of a certain number of cases with a good prognosis) in case of poor prognosis in the exposed group. Thus, the rate of false positives (for predicting the risk of recurrence) in the exposed group is higher than that of the unexposed.

In summary, factors which decrease blood PSA levels, increase the positive predictive value of PSA and should not be interpreted as an increase of the posterior probability of the disease. In studies using PSA as a prognostic marker, factors which decrease PSA blood levels translate a set of aggressive disease in the group at low risk; this effect corresponding to the Will Rogers paradox increases artificially the aggressiveness of the group exposed to the factor, which decreases PSA blood level. (For instance, a factor which reduces biologically by 2x the secretion of the PSA antigen increases the probability to detect a PCa and an aggressive disease for a same trigger cutoff to perform biopsy.) Because as the probability (predictive value) to detect a PCa increases with the blood level of PSA, a blood level of 4 ng/ml for a patient exposed is equivalent to a blood level at 8 ng/ml for a patient non exposed.

HOW TO MITIGATE BIAS IN CASE–CONTROL STUDIES BASED ON PROSTATE-SPECIFIC ANTIGEN STRATIFICATION

The use of instrumental variables such as germline genetics' approaches and different methods of recruitment of PCa cases and controls can help to improve estimates of causal effects.

(1) Do not use PSA as stratification marker or test its independence with the exposure in the general or control population.
The effect of the factors of interest must be tested for their independence to the exposure factor before using it to select controls. As an example in the case study below, we show that BMI, height and prostate volume correlate with the PSA blood levels (LogPSA).
(2) In the same study design, use different methods of selection of controls.
Using different methods of control selection can unmask selection bias due to the impact of the

exposure of interest on the stratification biomarker. We show in the case study below that the use of controls based on having a low PSA level (a group wherein systematic prostate biopsies are not feasible or ethical on a population basis) and a control group with negative prostate biopsy when the PSA level is high, can change the direction of the association between the risk factor of interest and the outcome (diagnosis of PCa), if the risk factor is not causal but changes the predictive value of the PSA test. Conversely, a true causal risk factor shows the same direction of association regardless of the selection of the control groups.

(3) Use multiple genetics instrumental variables as control causal parameters.
The causal impact of germline genetic variation on phenotypic traits is widely recognized. This causal property is used in Mendelian randomization [14], but it does not avoid selection bias (Fig. 3). The use of a positive control – a causal genetic factor known to influence the outcome – and a negative control that is known not to influence the outcome – can help to interpret the causality of the association observed between the exposure of interest and the outcome. Instrumental variable does not apply only on the outcome but also on the marker used for stratification of the outcome before to use it as marker for selection.

Instruments define factors such as genetic variants (single nucleotide polymorphism) or other condition chosen because they have a causal effect on their target. For instance, Instrument-1 could be a genetic factor, which drives the suspected risk factor. If you look for the impact of obesity on cancer, you can use SNP associated with the susceptibility to obesity; it is the instrumental variable in Mendelian randomization design. Instrument-2 is a recognized causal factor such as genetic factor, which drives the outcome (e.g. genetic susceptibility gene for PCa). Instrument-3 is also a 'positive' control but acts as a confounding factor. It is selected because it is known as independent of the outcome (risk of cancer) and known to change the accuracy of the marker test. If the exposition changes the accuracy of the test used to define Controls versus Cases, the red pathway is activated and introduces a bias; the direction of the association can change with the stratification based on of cases versus controls the marker test of recruitment. The use of Instrument-1, which is upstream to the exposition, does not correct the bias. Instrument-2 is a 'causal control' because its causal effect is independent of the marker; the direction of the association cannot change even if the modalities to select cases versus controls change. Instrument-3 is a 'non causal control' because the direction of the association can change with the stratification based on of cases versus controls the marker test of recruitment.

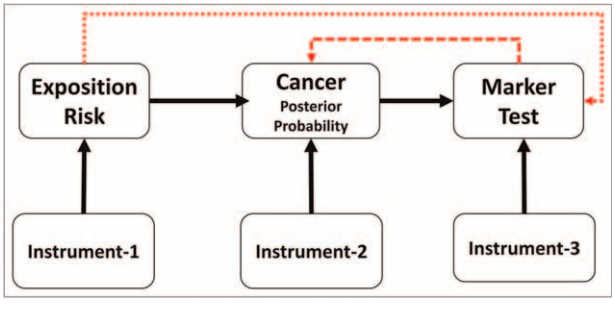

FIGURE 3. 'Instruments' definition.

CASE STUDY

Family history, BMI, height and prostate volume and association with prostate cancer diagnosis

To illustrate the difficulties in the design, analysis and interpretation of causal effects from case–control studies or routine databases, we processed data from the PROGENE cohort [20] according to different modalities to select the control groups based on PSA diagnostic procedures (948 PCas; 694 controls defined by a low PSA <4 ng/ml; 689 controls defined by a negative prostate biopsy justified by a PSA >4 ng/ml) (Figure 1-Supplement, http://links.lww.com/COU/A42). Family history of PCa (defined as > 1 relatives, e.g. brother, father with PCa) and more recently genetic variants have been shown to impact on the prevalence of PCa. Genetic variants could be considered as causal risk factors for PCa. So, we evaluated effect of the 8q24 variant – rs6983267 – on PCa risk [21]. Prostate volume is widely known to be correlated positively with PSA blood levels and to not be a causal factor for PCa. The relationship between BMI (BMI: weight/height^2) and PCa risk or aggressiveness has been the subject of contradictory reports across over 1000 publications over the last 10 years [22,23]. Recent results using Mendelian randomization approaches suggest that height increases risk of PCa, whereas BMI decreases this risk [21]. On the contrary, a systematic review, taking in account the impact of BMI on PSA, does not support an association between BMI and risk of PCa but confirms strong evidence of an inverse association between BMI and PSA [24]. We introduce the genetic variant rs17782313 (MCAR) associated with susceptibility to obesity [25]. Data were computed using logistic regression (Table No. 1-supplement,

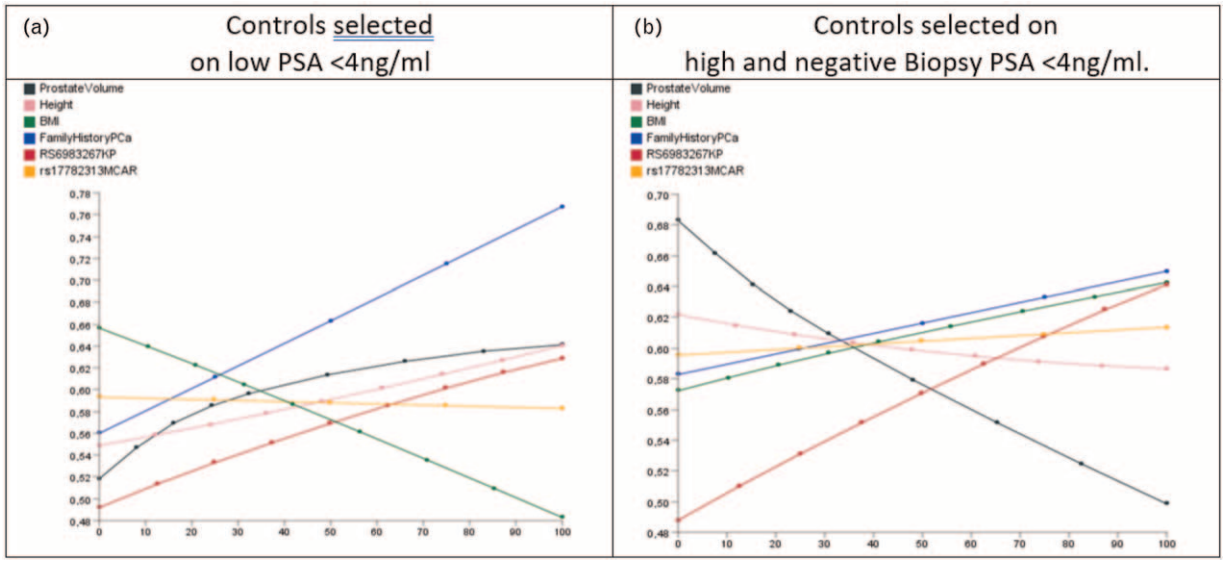

FIGURE 4. (a,b) Representation of direct effects.

http://links.lww.com/COU/A42) (XLSTAT 1.4 Lumivero 2023, France) to provide an odds ratio and Tree Augmented Naive Bayes Learning method to provide Bayes factors (Table No. 2-supplement, http://links.lww.com/COU/A42) and direct effects [26] (BayesiaLab 10.2, France). To understand how probability of the outcome variable is dependent on the variability of factors (input variables: Figure 2-supplement, http://links.lww.com/COU/A42), direct effects analysis (Table No. 3-supplement, http://links.lww.com/COU/A42) was performed. The dichotomization of continuous variables was based on the usual clinical ranges. BMI was dichotomised as less than 25% to more than 30% and prostate volume as less than 30 ml to more than 60 ml. They are tools for measuring the responsiveness of one variable to changes in another, analysing both linear and nonlinear dependencies. Direct effects are calculated from the percentage change (d) in risk divided by the percentage change in each input variable according to the formula: Direct effect $D_e x = dy/dx$ [26].

Results show (Supplement) that the direction of the associations with the outcome (PCa) changed, for both the frequentist and Bayesian inference (Fig. 4a,b) calculation methods, for factors that influence PSA levels, such as prostate volume, height and BMI. The direction of associations did not change for recognized causal factors: family history and the GG genotype of the Pca susceptibility variant rs6983267. Pearson analysis shows that prostate volume correlates positively with PSA level (LogPSA) (Table No. 4, http://links.lww.com/COU/A42- and Figure 4 supplement, http://links.lww.com/COU/A42). The genetic variant (genotype CC) of susceptibility to obesity rs17782313-MCAR changes according to the BMI direction ($P = 0.03$). Curves show the direction of the direct effects of the variables on the response variable (PCa risk).

CONCLUSION

Epidemiological evidence on the cause of PCa has been provided by observational studies. However, the impact of various biases on causal inferences from these observational studies could be considerable, and lead to inappropriate decision making on the conduct of randomized controlled trials, and policy and clinical guidelines. Representative population surveys or sampling strategies that avoid the problems of selection bias related to PSA stratification are required to provide more reliable evidence. We show that we can overcome the limitations induced by various selection biases by understanding and exploiting their mechanisms. The use of a double modality of recruitment based on marker stratification can point out real or spurious causal risk factors. In addition, the integration of multiple validated genetic instruments allows more robust causal interpretation of the etiological role of the factors studied.

Acknowledgements

None.

Financial support and sponsorship

O.C. is funded by SCOR Corporate Foundation for Science

R.M.M. is a National Institute for Health Research Senior Investigator (NIHR202411). R.M.M. is supported

by a Cancer Research UK 25 (C18281/A29019) programme grant (the Integrative Cancer Epidemiology Programme). R.M.M. is also supported by the NIHR Bristol Biomedical Research Centre, which is funded by the NIHR (BRC-1215-20011) and is a partnership between University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. R.M.M. is affiliated with the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC_UU_00011/1, MC_UU_00011/3, MC_UU_00011/6 and MC_UU_00011/4) and the University of Bristol. Department of Health and Social Care disclaimer: The views expressed are those of the author (s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ■ of special interest
- ■■ of outstanding interest

1. Mottet N, Bellmunt J, Bolla M, *et al.,* EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, diagnosis, and local treatment with curative intent. Eur Urol 2017; 71:618–629.
2. Catalona WJ. Clinical utility of measurements of free and total prostate-specific antigen (PSA): a review. Prostate Suppl 1996; 7:64–69.
3. Eldred-Evans D, Burak P, Connor MJ, *et al.* Population-based prostate cancer screening with magnetic resonance imaging or ultrasonography: the IP1-PROSTAGRAM Study. JAMA Oncol 2021; 7:395–402.
4. Vallières E, Mésidor M, Roy-Gagnon MH, *et al.* General and abdominal obesity trajectories across adulthood, and risk of prostate cancer: results from the PROtEuS study, Montreal, Canada. Cancer Causes Control 2021; 32:653–665.
5. Cornu JN, Cancel-Tassin G, Cox DG, *et al.* Impact of body mass index, age, prostate volume, and genetic polymorphisms on prostate-specific antigen levels in a control population. Eur Urol 2016; 70:6–8.
6. Henderson RJ, Eastham JA, Culkin DJ, *et al.* Prostate-specific antigen (PSA) and PSA density: racial differences in men without prostate cancer. J Natl Cancer Inst 1997; 89:134–138.
7. Seo DH, Yoon S, Choi JH, *et al.* The correlation between body mass index and routine parameters in men over fifty. World J Mens Health 2017; 35: 178–185.
8. Parikh H, Wang Z, Pettigrew KA, *et al.* Fine mapping the KLK3 locus on chromosome 19q13.33 associated with prostate cancer susceptibility and PSA levels. Hum Genet 2011; 129:675–685.
9. Nieder C, Norum J, Geinitz H. Impact of common medications on serum total prostate-specific antigen levels and risk group assignment in patients with prostate cancer. Anticancer Res 2011; 31:1735–1739.
10. Wong LI, Labrecque MP, Ibuki N, *et al.* p,p′-Dichlorodiphenyltrichloroethane (p,p′-DDT) and p,p′-dichlorodiphenyldichloroethylene (p,p′-DDE) repress prostate specific antigen levels in human prostate cancer cell lines. Chem Biol Interact 2015; 230:40–49.
11. Webb MPK, Sidebotham D. Bayes' formula: a powerful but counterintuitive tool for medical decision-making. BJA Educ 2020; 20:208–213.
12. Bell KJ, Del Mar C, Wright G, *et al.* Prevalence of incidental prostate cancer: a systematic review of autopsy studies. Int J Cancer 2015; 137:1749–1757.
13. Vallancien G, Prapotnich D, Veillon B, *et al.* Systematic prostatic biopsies in 100 men with no suspicion of cancer on digital rectal examination. J Urol 1991; 146:1308–1312.
14. Markozannes G, Kanellopoulou A, Dimopoulou O, *et al.* Systematic review of Mendelian randomization studies on risk of cancer. BMC Med 2022; 20:41.
15. Griffith GJ, Morris TT, Tudball MJ, *et al.* Collider bias undermines our understanding of COVID-19 disease risk and severity. Nat Commun 2020; 11:5749.

■ The article explains Berkson's selection bias and its persistence in recent epidemiological articles.

16. Gofrit ON, Zorn KC, Steinberg GD, *et al.* The Will Rogers phenomenon in urological oncology. J Urol 2008; 179:28–33.
17. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. N Engl J Med 1985; 312:1604–1608.
18. Gilbert R, Martin RM, Donovan J, *et al.* Misclassification of outcome in case-control studies: methods for sensitivity analysis. Stat Methods Med Res 2016; 25:2377–2393.
19. Thompson IM, Chi C, Ankerst DP, Goodman PJ, *et al.* Effect of finasteride on the sensitivity of PSA for detecting prostate cancer. J Natl Cancer Inst 2006; 98:1128–1133.
20. Valeri A, Berthon P, Fournier G, *et al.* The PROGENE study, the French project of genetic analysis of familial prostatic cancer: recruitment and analysis. Prog Urol 1996; 6:226–235.
21. Cussenot O, Azzouzi AR, Bantsimba-Malanda G, *et al.* Effect of genetic variability within 8q24 on aggressiveness patterns at diagnosis and familial status of prostate cancer. Clin Cancer Res 2008; 14:5635–5639.
22. Rivera-Izquierdo M, Pérez de Rojas J, Martínez-Ruiz V, *et al.* Obesity as a risk factor for prostate cancer mortality: a systematic review and dose-response meta-analysis of 280,199 patients. Cancers (Basel) 2021; 13:4169.
23. Harrison S, Tilling K, Turner EL, *et al.* Systematic review and meta-analysis of the associations between body mass index, prostate cancer, advanced prostate cancer, and prostate-specific antigen. Cancer Causes Control 2020; 31:431–449.
24. Davies NM, Gaunt TR, Lewis SJ, *et al.,* PRACTICAL consortium, Lathrop M, Smith GD, Martin RM. The effects of height and BMI on prostate cancer incidence and mortality: a Mendelian randomization study in 20,848 cases and 20,214 controls from the PRACTICAL consortium. Cancer Causes Control 2015; 26:1603–1616.
25. Yu K, Li L, Zhang L, Guo L, *et al.* Association between MC4R rs17782313 genotype and obesity: a meta-analysis. Gene 2020; 733:144372.
26. Conrady S, Jouffe L. Bayesian networks and BayesiaLab: a practical introduction for researchers. 1st ed. (19 June 2015), 382 pages. https://www.academia.edu/19627726/Bayesian_Networks_and_BayesiaLab_A_Practical_Introduction_for_Researchers