Received: 18 July 2024 | Revised: 15 November 2024 | Accepted: 21 November 2024

DOI: 10.1002/ijc.35279

RESEARCH ARTICLE

Cancer Epidemiology

IJC INTERNATIONAL JOURNAL of CANCER | UICC

Etiology of prostate cancer with the *TMPRSS2:ERG* fusion: A systematic review of risk factors

Colleen B. McGrath[1] | Alaina H. Shreves[1,2,3] | Megan R. Shanahan[1] | Hannah E. Guard[1] | Manelisi V. Nhliziyo[1] | Claire H. Pernar[1] | Kathryn L. Penney[1,4] | Tamara L. Lotan[5] | Michelangelo Fiorentino[1,6] | Lorelei A. Mucci[1,7] | Konrad H. Stopsack[1,8,9]

[1]Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA

[2]Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

[3]Nuffield Department of Population Health, University of Oxford, Oxford, UK

[4]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

[5]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[6]Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

[7]Discovery Sciences, American Cancer Society, Atlanta, GA, USA

[8]Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[9]Department of Epidemiological Methods and Etiological Research, Leibniz Institute for Prevention Research and Epidemiology—BIPS and Faculty of Human and Health Sciences, University of Bremen, Bremen, Germany

Correspondence
Konrad H. Stopsack, Leibniz Institute for Prevention Research and Epidemiology—BIPS, Achterstr. 30, 28359 Bremen, Germany.
Email: stopsack@leibniz-bips.de

Funding information
Prostate Cancer Foundation; National Cancer Institute, Grant/Award Numbers: P01 CA228696, P50 CA090381, R01 CA136578, R37 CA227190, R37 CA275914, T32 CA009001, U01 CA167552

Abstract

The most common somatic alteration in primary prostate cancer is the *TMPRSS2:ERG* gene fusion, which may be caused or promoted by distinct etiologic factors. The objective of this systematic review was to assess epidemiologic evidence on etiologic factors for prostate cancer by tumor *TMPRSS2:ERG* fusion status in human populations. Of 3071 publications identified, 19 cohort or case–control studies from six distinct study populations were included in this systematic review. Etiologic factors included germline genetic variants, circulating hormones, and dietary and lifestyle factors. Taller height, higher total and free testosterone levels, and fewer trinucleotide repeats in *AR* were possibly associated with higher risk of *TMPRSS2:ERG*-positive prostate cancer. Excess body weight, greater vigorous physical activity, higher lycopene intake, and the use of calcium channel blockers were associated with lower risk of *TMPRSS2:ERG*-positive prostate cancer. Diabetes and family history of prostate cancer were associated with both *TMPRSS2:ERG*-positive and *TMPRSS2:ERG*-negative prostate cancer. Prostate cancer germline variants had suggestive differential

Colleen B. McGrath and Alaina H. Shreves shared first authorship. Lorelei A. Mucci and Konrad H. Stopsack shared last authorship.

associations with *TMPRSS2:ERG*-positive or *TMPRSS2:ERG*-negative prostate cancer. However, results were based on few distinct study populations and generally had low precision, underscoring the need for replication. In conclusion, prostate cancer with *TMPRSS2:ERG* fusion is an etiologically distinct subtype that may be, in part, preventable by addressing modifiable and hormonally acting etiologic factors that align with the established mechanistic role of *TMPRSS2:ERG* in androgen, insulin, antioxidant, and growth factor pathways.

KEYWORDS
etiology, gene fusion, prostate cancer, systematic review, *TMPRSS2:ERG*

What's New?

The most common somatic alterations in primary prostate cancer are gene fusions involving *TMPRSS2* and *ERG*. *TMPRSS2:ERG*-positive cancers may be etiologically distinct from cancers without this fusion. This study sought to better understand etiologic factors underlying *TMPRSS2:ERG* fusion in prostate cancer via systematic review of epidemiologic studies. Of the etiologic factors for tumors with the *TMPRSS2:ERG* fusion, most were not associated with *TMPRSS2:ERG*-negative prostate cancer risk. Observed differences in relation between etiologic factors and *TMPRSS2:ERG*-positive prostate cancer require further investigation but offer novel insight into diverse etiologies of prostate cancer subtypes and mechanisms of carcinogenesis.

1 | INTRODUCTION

The most common somatic alterations in primary prostate cancer are gene fusions between *TMPRSS2* and members of the oncogenic E26 transformation specific (ETS) transcription factor family, most commonly *ERG*.[1–4] As many as 140,000 individuals may be diagnosed with *TMPRSS2:ERG*-positive prostate cancer annually among the 300,000 newly diagnosed patients in the United States.[5–7]

TMPRSS2:ERG and other ETS fusions are typically clonal events that occur early in carcinogenesis. Etiologic factors may act by initiating the gene fusion or creating a milieu conducive to the promotion or progression of tumors with such fusions.[2,3,8] In a cell line, the fusion results under irradiation when androgens induce chromosomal proximity of these two genes.[9] By bringing the oncogene *ERG* under transcriptional control of androgen-regulated *TMPRSS2*, *TMPRSS2:ERG* fusions have a profound impact on the tumor transcriptome and metabolome, with hundreds of transcripts and dozens of signaling pathways reproducibly differentially expressed between tumors with *TMPRSS2:ERG* fusions and those without ETS fusions.[10–14] Rewired amino acid and fatty acid metabolism are particularly notable, including differences in expression of the insulin-like growth factor 1 (IGF-1) and insulin receptors.[14–17]

Considering these notable differences, *TMPRSS2:ERG*-positive prostate cancers may have etiologies that are distinct from tumors lacking the fusion. An understanding of etiology according to molecular subtypes, as well as any heterogeneity in risk factor associations, could inform the development of prevention strategies. This systematic review aimed to synthesize evidence on etiologic factors for *TMPRSS2:ERG*-positive prostate cancers in human populations and elucidate the extent to which this molecular subtype is etiologically distinct from cancers without the fusion.

2 | MATERIALS AND METHODS

The protocol was registered at PROSPERO on August 15, 2019 (#133254). A search of Ovid MEDLINE was done from 1946 through October 21, 2020, as well as Embase and Web of Science from inception through October 21, 2020. An updated search was performed on March 15, 2023. The search strategy was developed together with a reference librarian to identify studies broadly related to the *TMPRSS2:ERG* gene fusion and prostate cancer. Search terms included, "TMPRSS2", "TMPRSS2:ERG", "Transmembrane protease serine 2", "T2ERG", "T2E", "ETS-related gene", "ETS gene fusions", "ERG positive" coupled with "prostatic cancer", "prostate cancer", "prostatic neoplasia", "prostate neoplasia", among others (Table S1).

Abstract and title screening based on pre-determined criteria considered all studies without date or language restriction (e.g., papers, book chapters, conference abstracts). Records were excluded if: (1) they were about non-human species, (2) *TMPRSS2:ERG* status was not confirmed by established biomarker methods, (3) the study did not assess an association between a potential etiologic factor and risk of prostate cancer with or without the *TMPRSS2:ERG* fusion, or (4) the study was restricted to patients with a prostate cancer diagnosis (case-only design). Titles and abstracts were independently screened using Covidence software by a combination of two reviewers (MVN, CHP, LAM, CBM, MRS) and conflicts were resolved by discussion with a third reviewer. All retained records underwent full-text screening by a combination of two reviewers (CBM, CHP, LAM, MRS).

Data extraction using a Covidence form was performed independently by a combination of two reviewers (CBM, CHP, AHS, MRS), and a consensus was reached after comparing data that were extracted manually for all included manuscripts. The data extracted included study characteristics (e.g., design/type, population), risk factor measurement, confounders, outcome ascertainment, tumor subtyping assay methods, and estimates of association.

Because simple scoring systems cannot adequately capture bias,[18,19] each study was reviewed for information bias, selection bias, confounding (with predefined confounders age, race, and prostate-specific antigen [PSA] screening), statistical modeling aspects, and consideration of effect modification.

2.1 | Statistical analysis

Numeric results from included original articles were extracted and displayed in tables. No additional statistical analyses were conducted, with the exception of inverting estimates from one study to match the direction reported in the second study examining this same risk factor.[20,21]

3 | RESULTS

3.1 | Systematic review and study characteristics

A total of 3504 articles were identified on systematic literature search, of which 3071 were unique. After title and abstract screening, 90 records were retrieved for full-text screening (Figure 1). All included studies had a full manuscript available.

Of the 19 studies that met inclusion criteria,[20–38] eight were cohort studies and 11 were case–control studies (Table 1). Of the case–control studies, five were nested in longitudinal studies, three sampled controls from population registers, and one used cumulative incidence sampling from a randomized control trial. Study populations overlapped between several of the included studies. Eight studies were in the Health Professionals Follow-up Study (HPFS), a prospective cohort of 51,529 men working in health professions in the United States aged 40–75 years at enrollment.[39,40] Five additional studies included both the HPFS and the Physicians' Health Study (PHS), initially a randomized-controlled trial of aspirin and micronutrients among 22,071 men working as physicians aged 40–84 years at enrollment.[41] In these two cohorts, 46%–51% of tumors were *TMPRSS2:ERG*-positive by immunohistochemistry. Three additional studies were based in two population-based case–control studies among Black and White men in King County, Washington. Cases were identified via the Seattle-Puget Sound Surveillance, Epidemiology, and End Results (SEER) cancer registry, of which 49%–52% were *TMPRSS2:ERG*-positive by fluorescence in situ hybridization (FISH). Controls were recruited via random digit dialing with frequency matching by age and calendar year. The remaining three populations were from case–control studies in the Prostate Cancer Prevention Trial (PCPT, 50% *TMPRSS:ERG*-positive by FISH), at the University of Ulm, Germany (68% *TMPRSS2:ERG*-positive by FISH), and in the Collaborative Oncological Gene–environment Study (COGS) Consortium from five study centers in Finland, Germany, the United Kingdom, the United States, and Portugal (54% *TMPRSS2:ERG*-positive by FISH or quantitative real-time polymerase chain reaction).

3.2 | Bias assessment

In terms of information bias, exposure assessment was generally adequate for the included lifestyle exposures, with use of repeated measures to reduce measurement error in the prospective studies. While inherently imperfect, reliability and validity for several lifestyle factors and circulating hormonal biomarkers have been formally assessed. The type of assay used to assess *TMPRSS2:ERG* status varied across studies; reliability and validity were generally considered high given the inherently high signal-to-noise ratio of ERG overexpression[14] and the rarity of other 5′ fusion partners (e.g., *SLC45A3*, *NDRG1*). In terms of selection bias, all studies were restricted to surgically treated patients with available tissue from prostatectomy or transurethral resection of the prostate, except one case–control study, which used biopsy tissue.[20] The use of inverse-probability weights to address selective missingness of tissue in two cohort studies[26,28] did not materially change results, suggesting limited impact of potential selection bias for that cohort and etiologic factor. In terms of confounding control, studies controlled for age, race, and PSA screening by multivariable adjustment or restriction of study populations, except for studies examining germline genetic factors and one study[25] that did not control for PSA screening. Additional confounders related to specific exposures were also adjusted for across studies. Modeling typically used logistic regression or Cox models for competing events, depending on the study design; sparse data presumably affected estimates of some studies.

Studies were restricted to the United States, with predominantly White men included (93% or more), as well as select European countries. The HPFS and PHS cohorts further restricted by educational attainment, which could impact generalizability but inherently controls for confounding by many socioeconomic factors.

3.3 | Body size, diabetes mellitus, and physical activity

Body size as an etiologic factor was assessed in two independent studies (Table 2), a cohort study in the HPFS cohort and a case–control study in King County, Washington.[23,26] The study in the HPFS cohort found an association for higher body mass index (BMI) assessed early in life but after puberty (rate ratio [RR] per 5 kg/m^2 0.87, 95% confidence interval [CI] 0.73–1.03) and current BMI (RR per 5 kg/m^2 0.86, 95% CI 0.74–1.00) with lower rates of *TMPRSS2:ERG*-positive prostate cancer.[26] Similarly, the study conducted in the King County, Washington population reported BMI in

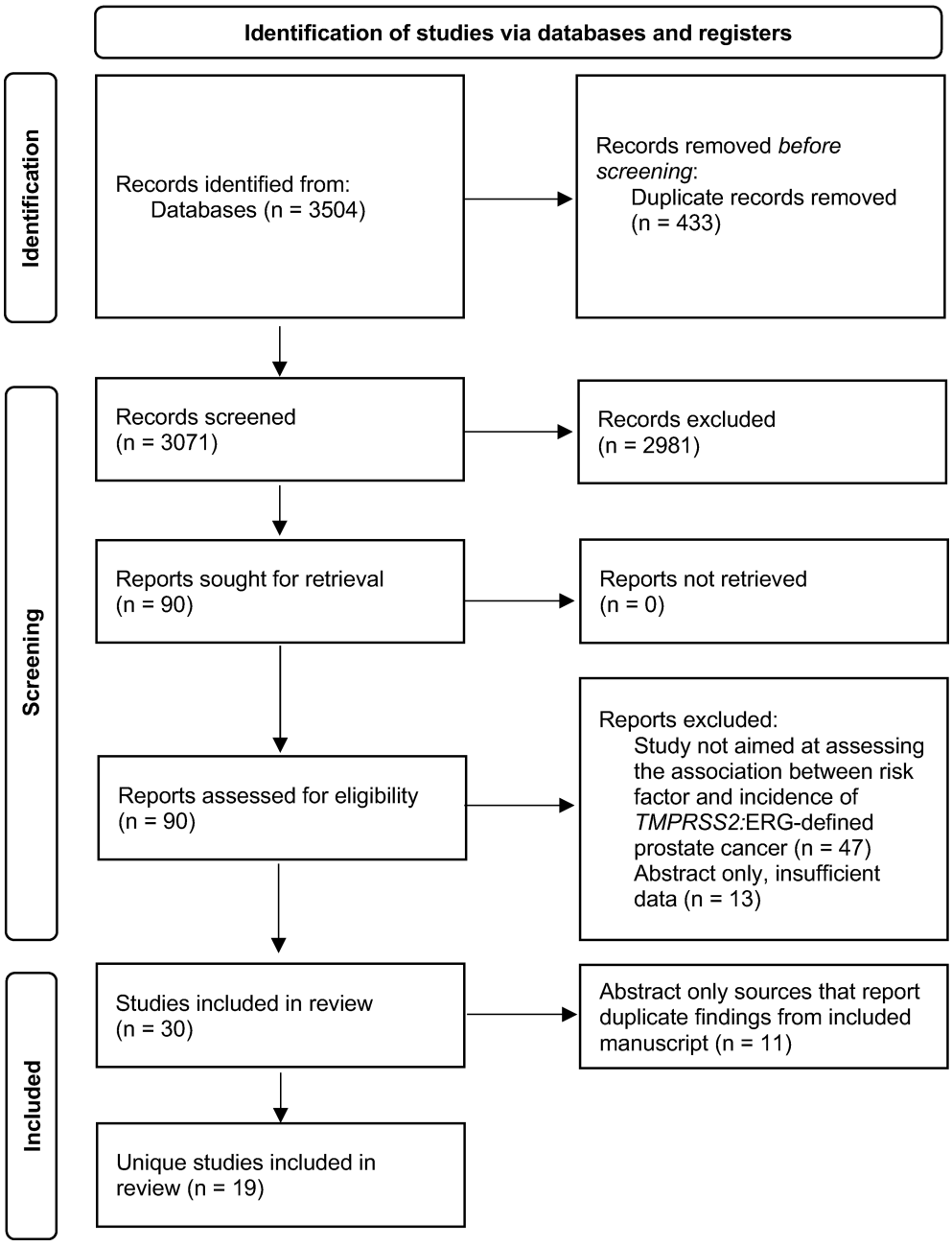

FIGURE 1 PRISMA flowchart of study selection using defined screening criteria.

early-life but after puberty (RR per 5 kg/m^2 0.83, 95% CI 0.68–1.03) and current BMI (RR per 5 kg/m^2 0.86, 95% CI 0.73–1.00) was associated with lower rates of *TMPRSS2:ERG*-positive prostate cancer.[23] In both studies, associations between early-life BMI and current BMI with *TMPRSS2:ERG*-negative prostate cancer were null. Similarly, greater waist circumference, a measure of central adiposity, was inversely associated with *TMPRSS2:ERG*-positive prostate cancer (RR per 5 inches 0.85, 95% CI 0.72–1.02), but not *TMPRSS2:ERG*-negative prostate cancer (RR per 5 inches 1.06, 95% CI 0.90–1.25) in one study.[26]

In the cohort study that examined adult height,[26] taller height was positively associated with *TMPRSS2:ERG*-positive prostate cancer, (RR per 5 inches 1.24, 95% CI 1.03–1.50) with evidence of a linear trend. Taller height was not associated with *TMPRSS2:ERG*-negative prostate cancer (RR per 5 inches 0.98, 95% CI 0.82–1.18).

Diabetes mellitus was associated with lower rates of prostate cancer in one study, without heterogeneity by *TMPRSS2:ERG* (Table 2).[24]

Total physical activity was not associated with either *TMPRSS2:ERG*-positive or *TMPRSS2:ERG*-negative prostate cancer (Table 2).[35] Vigorous physical activity was inversely associated with *TMPRSS2:ERG*-positive prostate cancer (RR 0.71, 95% CI 0.52–0.97 comparing quintile 5 to quintile 1), but not *TMPRSS2:ERG*-negative prostate cancer (RR 1.05, 95% CI 0.78–1.40).

3.4 | Family history and common genetic variation in the germline

The genetic susceptibility to *TMPRSS2:ERG*-positive and *TMPRSS2:ERG*-negative prostate cancer was assessed both using family history

TABLE 1 Characteristics of the included studies on risk factors for prostate cancer by *TMPRSS2:ERG* status, through March 2023.

First author, year	Design	Study Base	*N* ERG+	*N* ERG−	ERG Assay	Risk factor(s)
Graff, 2018	Prospective cohort	HPFS	439	474	IHC	Height, BMI, waist circumference
Egbers, 2015	Population-based case–control	King County, WA	295	268	FISH	BMI
Feng, 2020	Prospective cohort	HPFS	452	497	IHC	Diabetes mellitus
Pernar, 2019	Prospective cohort	HPFS	449	496	IHC	Physical activity
Khan, 2020	Prospective cohort	HPFS	325	371	IHC	Baldness
Hashim, 2020	Prospective cohort	HPFS	417	471	IHC	Family history of prostate cancer
Graff, 2016	Nested case-control[a]	HPFS and PHS	94	106	IHC	Sex hormones
Figg, 2014	Case-control[b]	PCPT	97	98	FISH	*AR* trinucleotide repeat length
Yoo, 2014	Nested case-control[a]	HPFS and PHS	147	144	IHC	*AR* trinucleotide repeat length
Luedeke, 2009	Case–control	University of Ulm	32	15	FISH	13 variants in DNA repair genes
Luedeke, 2016	Case-control	COGS	296	256	FISH, qPCR	27 prostate cancer risk variants
Penney, 2016	Nested case-control[b]	HPFS and PHS	227	260	IHC	39 prostate cancer risk variants
Allott, 2020	Prospective cohort	HPFS	417	471	IHC	Statin use
Stopsack, 2018	Prospective cohort	HPFS	439	473	IHC	Aspirin use
Wright, 2016	Population-based case–control	King County, WA	171	175	FISH	Aspirin use, NSAID use
Geybels, 2017	Population-based case-control	King County, WA	295	268	FISH	Calcium channel blocker use
Graff, 2016	Prospective cohort	HPFS	426	458	IHC	Lycopene, tomato sauce
Graff, 2017	Nested case-control[b]	HPFS and PHS	179	191	IHC	Antioxidants
Feng, 2021	Nested case-control[b]	HPFS and PHS	128	149	IHC	Metabolites

Abbreviations: BMI, body mass index; COGS, Collaborative Oncological Gene–environment Study; HPFS, Health Professionals Follow-up Study; IHC, immunohistochemistry; FISH, fluorescence in-situ hybridization; qPCR, quantitative real-time polymerase chain reaction; NSAID, non-steroidal anti-inflammatory drug; PCPT, Prostate Cancer Prevention Trial; PHS, Physicians' Health Study.

[a]Nested in a prospective cohort study.

[b]With cumulative incidence sampling from a randomized-controlled trial.

TABLE 2 Summary of associations between risk factors including body size, diabetes mellitus, physical activity and prostate cancer by *TMPRSS2:ERG* status.

Risk factor(s)	Study	ERG-positive RR (95% CI)	ERG-negative RR (95% CI)	*p*-heterogeneity
Body mass index (BMI)				
Current, per 5 kg/m^2	Egbers, 2015	0.86 (0.73 to 1.00)	0.93 (0.80 to 1.09)	NR
Current, per 5 kg/m^2	Graff, 2018	0.86 (0.74 to 1.00)	1.03 (0.90 to 1.18)	0.07
At age 18–29 years, per 5 kg/m^2	Egbers, 2015	0.83 (0.68 to 1.03)	0.96 (0.78 to 1.19)	NR
At age 21 years, per 5 kg/m^2	Graff, 2018	0.87 (0.73 to 1.03)	0.96 (0.81 to 1.12)	0.43
Maximum, per 5 kg/m^2	Egbers, 2015	0.87 (0.75 to 1.01)	0.95 (0.82 to 1.10)	NR
Height				
Per 5 inches	Graff, 2018	1.24 (1.03 to 1.50)	0.98 (0.82 to 1.18)	0.07
Waist circumference				
Per 5 inches	Graff, 2018	0.85 (0.72 to 1.02)	1.06 (0.90 to 1.25)	0.07
Diabetes mellitus				
Present (vs. absent)	Feng, 2020	0.72 (0.46 to 1.12)	0.63 (0.42 to 0.95)	0.67
Physical activity				
Total activity, quintile 5 (vs. quintile 1)	Pernar, 2019	1.13 (0.82 to 1.57)	1.04 (0.78 to 1.38)	0.20
Vigorous activity, quintile 5 (vs. quintile 1)	Pernar, 2019	0.71 (0.52 to 0.97)	1.05 (0.78 to 1.40)	0.50

Abbreviations: CI, confidence interval; NR, not reported; RR, rate ratio (hazard ratio or odds ratio that estimates a rate ratio).

of prostate cancer as well as germline variants. Family history of prostate cancer was assessed in a single cohort study (Table 3).[30] A positive family history of prostate cancer was associated with higher rates of both *TMPRSS2:ERG*-positive and *TMPRSS2:ERG*-negative prostate cancer, with similar results for family history in the father, a brother, or both.[30]

Germline genetic single nucleotide polymorphisms (SNPs) at prostate cancer risk loci were assessed in two studies (Table S2).[32,34] These studies assessed a partially overlapping set of some of the earliest-identified prostate cancer risk SNPs, some of which had stronger associations with *TMPRSS2:ERG*-positive prostate cancer than *TMPRSS2:ERG*-negative prostate cancer, or vice versa. A subset of risk variants assessed in both studies was replicated by meeting the null-hypothesis significance testing threshold in both individual studies.[32,34] A third, earlier study assessed a distinct set of germline variants in DNA repair genes.[33]

3.5 | Sex hormones and androgen receptor signaling

Sex hormones measured in pre-diagnostic blood were evaluated in one case–control study nested in a prospective cohort (Table 3).[29] Etiologic heterogeneity was seen with free testosterone levels, which were positively associated with *TMPRSS2:ERG*-positive prostate cancer (odds ratio [OR] per 1 standard deviation [SD] 1.37, 95% CI 1.05–1.77), but not with *TMPRSS2:ERG*-negative cancer (OR per 1 SD 1.09, 95% CI 0.86–1.38). Results for other sex hormones were null for both subtypes, and confidence intervals were wide. Baldness patterns, a trait influenced by androgens, were not associated with either *TMPRSS2:ERG*-positive or *TMPRSS2:ERG*-negative prostate cancer (Table 3).[31]

Genetic variation in *AR* includes a trinucleotide repeat that influences androgen receptor-driven transcription. Shorter trinucleotide repeat length, leading to higher androgen receptor (AR) activity, was positively associated with *TMPRSS2:ERG*-positive prostate cancer in one study (OR per 1 fewer repeat 1.07, 95% CI 1.00–1.14),[21] while a second study had compatible results but closer to null (OR per 1 fewer repeat 1.03, 95% CI 0.96–1.10; Table 3).[20] The association with *TMPRSS2:ERG*-negative cancer was null in both studies.

3.6 | Medications

Commonly used medications were assessed in four studies (Table 4).[22,25,36,37] In the case–control studies from King County, Washington, use of calcium channel blockers (RR 0.38, 95% CI 0.19–0.78),[25] aspirin (RR 0.63, 95% CI 0.43–0.93),[37] and non-aspirin non-steroidal anti-inflammatory drugs (NSAIDs; RR 0.65, 95% CI 0.37–1.14)[37] were associated with lower rates of *TMPRSS2:ERG*-positive prostate cancer, but not associated with *TMPRSS2:ERG*-negative cancer. Null associations, albeit with limited precision, were observed for

TABLE 3 Summary of associations between risk factors involving hormonal factors and heritability, including baldness, circulating sex hormone levels, CAG repeat length, family history of prostate cancer, and prostate cancer by *TMPRSS2:ERG* status.

Risk factor(s)	Study	ERG-positive HR (95% CI)	ERG-negative HR (95% CI)	*p*-heterogeneity
Family history of prostate cancer	Hashim, 2020			
In brother or father (vs. none)		1.49 (1.13 to 1.95)	2.15 (1.71 to 2.70)	0.04
In father (vs. none in father)		1.30 (0.96 to 1.76)	2.09 (1.64 to 2.66)	0.01
In brother (vs. none in brother)		1.96 (1.16 to 3.29)	1.71 (1.03 to 2.86)	0.72
Baldness	Khan, 2020			
Frontal baldness (vs. no baldness)		1.22 (0.94 to 1.58)	1.03 (0.80 to 1.32)	0.38
Vertex baldness (vs. no baldness)		0.89 (0.68 to 1.17)	1.00 (0.78 to 1.27)	
		OR (95% CI)	**OR (95% CI)**	
Sex hormones, per 1 standard deviation	Graff, 2016			
Total testosterone		1.20 (0.95 to 1.52)	1.12 (0.90 to 1.38)	0.63
Free testosterone		1.37 (1.05 to 1.77)	1.09 (0.86 to 1.38)	0.17
Dihydrotestosterone		1.04 (0.80 to 1.35)	1.03 (0.81 to 1.30)	0.93
Androstanediol glucuronide		1.11 (0.90 to 1.37)	1.22 (0.99 to 1.49)	0.51
Estradiol		1.07 (0.84 to 1.35)	0.95 (0.78 to 1.16)	0.44
Sex hormone-binding globulin		1.12 (0.88 to 1.44)	1.14 (0.91 to 1.42)	0.94
Germline variation in the androgen receptor				
Per 1 fewer CAG repeat in *AR*	Figg, 2014	1.03 (0.96 to 1.10)	0.96 (0.90 to 1.03)	NR
Per 1 fewer CAG repeat in *AR*	Yoo, 2014	1.07 (1.00 to 1.14)	0.99 (0.93 to 1.05)	NR

Abbreviations: CAG, cytosine-adenine-guanine; CI, confidence interval; HR, hazard ratio; NR, not reported; OR, odds ratio.

TABLE 4 Summary of associations between medication use and prostate cancer by *TMPRSS2:ERG* status.

Risk factor(s)	Study	ERG-positive RR (95% CI)	ERG-negative RR (95% CI)	*p*-heterogeneity
Statin use				
Current use (vs. non-use)	Allott, 2020	0.99 (0.74 to 1.33)	1.01 (0.78 to 1.31)	NR
Aspirin and non-aspirin NSAID use				
Current non-aspirin NSAID use (vs. non-use)	Wright, 2016	0.65 (0.37 to 1.14)	1.04 (0.64 to 1.67)	NR
Current aspirin use (vs. non-use)	Wright, 2016	0.63 (0.43 to 0.93)	0.99 (0.69 to 1.42)	NR
Current aspirin use (vs. non-use)	Stopsack, 2018	1.02 (0.85 to 1.23)	1.09 (0.91 to 1.30)	0.63
Acetaminophen use				
Current use (vs. non-use)	Wright, 2016	0.93 (0.43 to 1.98)	1.12 (0.58 to 2.18)	NR
Calcium channel blocker use				
Ever use (vs. never use)	Geybels, 2017	0.38 (0.19 to 0.78)	1.05 (0.65 to 1.69)	NR

Abbreviations: CI, confidence interval; NR, not reported; NSAID, non-steroidal anti-inflammatory drug; RR, rate ratio (hazard ratio or odds ratio that estimates a rate ratio).

TABLE 5 Summary of associations between pre-diagnostic circulating blood levels or dietary intakes of antioxidants and prostate cancer by *TMPRSS2:ERG* status.

Risk factor(s)	Study	ERG-positive OR (95% CI)	ERG-negative OR (95% CI)	*p*-heterogeneity
Blood levels, quartile 4 (vs. quartile 1)	Graff, 2017			
Alpha-carotene		0.99 (0.62 to 1.56)	0.81 (0.53 to 1.26)	0.35
Beta-carotene		0.67 (0.39 to 1.15)	0.94 (0.56 to 1.59)	0.73
Alpha-tocopherol		0.76 (0.48 to 1.20)	1.03 (0.66 to 1.61)	0.13
Gamma-tocopherol		1.16 (0.75 to 1.78)	1.25 (0.80 to 1.96)	0.58
Beta-cryptoxanthin		1.08 (0.69 to 1.68)	1.26 (0.82 to 1.92)	0.89
Lutein		0.62 (0.38 to 1.01)	0.82 (0.53 to 1.27)	0.63
Lycopene		0.86 (0.54 to 1.39)	1.00 (0.65 to 1.53)	0.95
Retinol		0.92 (0.51 to 1.66)	1.16 (0.65 to 2.07)	0.40
Selenium		1.59 (0.78 to 3.25)	1.56 (0.74 to 3.31)	0.87
		HR (95% CI)	**HR (95% CI)**	
Dietary Lycopene, cumulative average intake	Graff, 2016			
Quintile 5 (vs. quintile 1)		0.65 (0.47 to 0.89)	0.80 (0.59 to 1.09)	0.79
Tomato sauce, cumulative average intake	Graff, 2016			
≥2 servings/week (vs. <1 serving/month)		0.54 (0.37 to 0.81)	0.96 (0.62 to 1.50)	0.04

Abbreviations: CI, confidence interval; HR, hazard ratio; OR, odds ratio.

acetaminophen.[37] In the HPFS cohort, null results were observed for aspirin and *TMPRSS2:ERG*-positive prostate cancer (RR 1.02, 95% CI 0.85–1.23).[36] Use of cholesterol-lowering statin medications was not associated with *TMPRSS2:ERG*-positive or *TMPRSS2:ERG*-negative prostate cancer.[22]

3.7 | Dietary factors and metabolites

Two studies prospectively investigated dietary or circulating levels of antioxidants in prediagnostic blood samples (Table 5).[27,28] One study found inverse associations between higher dietary intakes of tomato sauce (hazard ratio [HR] 0.54, 95% CI 0.37–0.81) and the antioxidant lycopene (HR 0.65, 95% CI 0.47–0.89) with *TMPRSS2:ERG*-positive prostate cancer, comparing extreme categories of intake.[28] In the same study population, results for pre-diagnostic circulating blood levels of lycopene and *TMPRSS2:ERG*-positive cancer were inconclusive (HR 0.86, 95% CI 0.54–1.39, comparing quartile 4 to quartile 1),[27] as were results for associations with *TMPRSS2:ERG*-negative prostate cancer. Associations for the antioxidants α-carotene, α-tocopherol, β-cryptoxanthin, γ-tocopherol, lutein, retinol, and selenium were also null with wide confidence intervals for both *TMPRSS2:ERG*-positive and *TMPRSS2:ERG*-negative cancer.[27]

Pre-diagnostic circulating blood metabolites levels were assessed in one study (data not shown).[38] Using enrichment analysis, three lipid

classes—sphingomyelins, ceramides, and phosphatidylethanolamines—tended to be associated with *TMPRSS2:ERG*-positive prostate cancer, and none with *TMPRSS2:ERG*-negative prostate cancer, but confidence intervals were wide.[38]

4 | DISCUSSION

This systematic review summarizes evidence for heterogeneity in the etiology of *TMPRSS2:ERG*-positive and *TMPRSS2:ERG*-negative prostate cancer. The rationale for etiologic heterogeneity between tumor molecular subtypes is established in other cancers, such as breast cancer, where tumors have distinct risk factor profiles depending on molecular subtype.[42] In prostate cancer, the 19 studies in this systematic review provide evidence for etiologic factors specifically for *TMPRSS2:ERG*-positive tumors and often null associations with *TMPRSS2:ERG*-negative tumors, except germline genetic factors and family history. Etiologic factors for *TMPRSS2:ERG*-positive prostate cancer were assessed via germline variants, circulating blood levels, and questionnaire-based information. Those with at least suggestive evidence include body composition, sex hormones, vigorous physical activity, certain antioxidants, and commonly used medications.

Etiologic heterogeneity by *TMPRSS2:ERG* fusion status is motivated by the profound tumor molecular differences induced by the gene fusion.[2,10,13,14] These differences, among others, include the insulin and IGF-1 signaling pathway. Higher IGF-1 blood levels are consistent etiologic factors for overall and aggressive prostate cancer.[43,44] Likewise, dietary and lifestyle factors associated with hyperinsulinemia are associated with prostate cancer risk.[45] Notably, *TMPRSS2:ERG* directly regulates IGF-1 receptor in vitro,[17] and ETS/ERG-positive tumors have higher expression of IGF1R and insulin receptor than *TMPRSS2:ERG*-negative tumors.[15–17,46] Among the etiologic factors in this systematic review, adult height in part reflects IGF-1 levels during adolescence, which may partially explain the positive association between height and *TMPRSS2:ERG*-positive prostate cancer, but not with *TMPRSS2:ERG*-negative cancer.[26] While also limited in precision, associations of circulating lipid metabolites and risk of *TMPRSS2:ERG*-positive prostate cancer are consistent with other lines of evidence on metabolic influences on *TMPRSS2:ERG*-positive tumors.[11–14]

Sex hormones and *AR* signaling are a second set of mechanisms motivating several etiologic factors in the included studies, given that the fusion brings ETS genes under control by the androgen-regulated gene *TMPRSS2*,[47] and in light of potential estrogen sensitivity of *TMPRSS2:ERG*.[10] The studies in this systematic review on prediagnostic circulating levels of free testosterone and trinucleotide repeat length in *AR* partially support this connection.[20,21] In established tumors, however, differences in androgen receptor signaling by ETS status are more complex.[13]

The hypothesis that oxidative stress promotes formation of the gene fusion[48] motivated studies on the antioxidant lycopene. While corresponding dietary intakes appeared to be protective factors for *TMPRSS2:ERG*-positive prostate cancer in one study,[28] results for plasma antioxidant levels in another small study were null with wide confidence intervals.[27] Inflammation and oxidative stress may also be influenced by commonly used anti-inflammatory medications, including aspirin and NSAIDs, but results of the included studies[36,37] are heterogeneous.

Several etiologic factors are likely reflective of multiple mechanisms. For example, adult height reflects both IGF-1 and androgen signaling.[49] Adiposity alters multiple signaling pathways. While adiposity-induced hyperinsulinemia increases free or bioactive IGF-1, adiposity also decreases androgen levels including free testosterone.[50] Higher levels of free testosterone are related to higher risk of aggressive prostate cancer.[51] Two studies in independent populations consistently reported an inverse association between higher BMI and *TMPRSS2:ERG*-positive prostate cancer and null associations for *TMPRSS2:ERG*-negative prostate cancer.[23,26] These findings suggest that the established inverse association of adiposity and overall prostate cancer risk[52] could be driven by fewer *TMPRSS2:ERG*-positive tumors in a low-androgen environment; however, additional data on tumors with non-*ERG* ETS fusions and on the etiology of non-ETS tumors are needed. Physical activity reduces IGF-1 and insulin,[50] and the study in this review suggested an inverse association of vigorous physical activity and *TMPRSS2:ERG*-positive prostate cancer.[35] These associations may suggest that effects of physical activity on *TMPRSS2:ERG*-positive prostate cancer are not primarily mediated via body weight.

Finally, beyond individual mechanisms, several included studies evaluated etiologic heterogeneity by tumor subtype based on germline genetics. Prostate cancer is a highly heritable cancer, with 58% of variation in incidence explained by germline genetic factors (95% CI 52–63),[53] and the omnibus measure of family history was associated with both *TMPRSS2:ERG*-positive and *TMPRSS2:ERG*-negative prostate cancer.[30] In contrast, studies evaluating individual SNPs suggested that at least some underlying germline variants predispose specifically to *TMPRSS2:ERG*-positive prostate cancer but not *TMPRSS2:ERG*-negative prostate cancer, and vice versa. Similar findings have been obtained in case-only studies, not included in this review,[54,55] and in a study published after completion of the systematic search.[56] However, these studies still had limited precision, likely leaving a substantial amount of germline predisposition to subtypes undetected, and there is a need for larger genome-wide association studies of prostate cancer subtypes.

This systematic review focused on etiologic factors for incidence of prostate cancer subtypes defined by *TMPRSS2:ERG* status. Studies that evaluated the prevalence of risk factors between the subtypes among prostate cancer cases were not eligible for inclusion. Such studies can, however, be informative about etiologic heterogeneity, for example, by implying common[54,55] and rare germline genetic variation in the etiology of *TMPRSS2:ERG*-positive tumors.[57] We further did not consider age and race as potential etiologic factors, as neither have a causal interpretation, and instead assessed whether studies controlled for them as confounders. The prevalence of *TMPRSS2:ERG*-positive tumors among diagnosed cases differs by race and is lower

among Black men than White men.[6] In addition, *TMPRSS2:ERG*-positive tumors tend to be diagnosed at a slightly younger age than *TMPRSS2:ERG*-negative tumors.[5,58] Identifying the underlying etiologic factors for tumor subtypes is necessary to understand which causes manifest in such differences by age and race.

A key limitation of the evidence in this systematic review is limited sample sizes in the underlying original studies, resulting in low precision of subtype-specific associations and weak formal statistical evidence for heterogeneity between subtype-specific associations. Further, most of the etiologic factors were assessed in one or two separate and demographically homogenous study populations. It is critical that findings be replicated with larger sample sizes, that studies include more non-White men to assess generalizability, and that additional etiologic factors be considered.

Tumors with the *TMPRSS2:ERG* fusion are transcriptomically similar to those with other ETS fusions,[14] which account for one in seven prostate tumors.[52] Studies should thus disaggregate *TMPRSS2:ERG*-negative tumors into those with non-*ERG* ETS fusions and tumors without any ETS fusions, and further subtype the latter. A plausible but unstudied hypothesis is that tumors with non-*ERG* ETS fusions have a similar etiology to those with fusions involving *ERG*. Strikingly, none of the included studies reported etiologic factors for *TMPRSS2:ERG*-negative tumors beyond certain germline risk variants and the etiologic factors that did not have heterogeneity by ERG status. This absence of specific risk factor associations for *TMPRSS2:ERG*-negative tumors may reflect the fact that the term "ERG-negative" summarizes a heterogeneous collection of tumor subtypes defined by non-*ERG* ETS fusions, *SPOP* or *FOXA1* mutations,[4] as well as other, more rare drivers with presumably different etiologies. Expecting that "ERG-negative prostate cancer" has a specific etiology would be akin to considering all subtypes of prostate cancer as a single entity. A question for further investigation related to the current review is the extent to which tumors with the *TMPRSS2:ERG* fusion and other prostate cancers differ by factors such as PSA kinetics or imaging features, which might help further tailor prostate cancer screening.

5 | CONCLUSIONS

Etiologic factors for *TMPRSS2:ERG*-positive prostate cancer, as summarized in this review, largely align with biological characteristics of this tumor subtype. If confirmed, at least a subset of prostate cancer is amenable to prevention approaches, as many of the etiologic factors are modifiable. More broadly, the preliminary evidence summarized here may motivate further research into the etiologies of different subtypes of prostate cancer to uncover subtype-specific mechanisms of carcinogenesis.

AUTHOR CONTRIBUTIONS

Colleen B. McGrath: Data curation; formal analysis; writing – original draft; writing – review and editing. **Alaina H. Shreves:** Data curation; formal analysis; writing – original draft; writing – review and editing. **Megan R. Shanahan:** Data curation; formal analysis; writing – review and editing; project administration. **Hannah E. Guard:** Formal analysis; writing – review and editing; project administration. **Manelisi V. Nhliziyo:** Conceptualization; methodology; data curation; writing – review and editing; project administration. **Claire H. Pernar:** Conceptualization; methodology; data curation; project administration; writing – review and editing. **Kathryn L. Penney:** Formal analysis; writing – review and editing; funding acquisition. **Tamara L. Lotan:** Writing – review and editing. **Michelangelo Fiorentino:** Writing – review and editing. **Lorelei A. Mucci:** Conceptualization; methodology; data curation; formal analysis; writing – original draft; writing – review and editing; funding acquisition; supervision. **Konrad H. Stopsack:** Formal analysis; writing – original draft; writing – review and editing; funding acquisition; supervision.

ACKNOWLEDGMENTS

The authors thank Paul Bain, PhD, MLIS, Reference and Instruction Librarian at the Countway Library of Harvard Medical School, for assistance in specifying search terms and performing the search. Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

This study was supported by the National Cancer Institute (grant numbers R01 CA136578, R37 CA275914, R37 CA227190, U01 CA167552, P01 CA228696, P50 CA090381, and T32 CA009001). Alaina H. Shreves is supported by the National Institutes of Health's Intramural Research Program. Kathryn L. Penney, Tamara L. Lotan, Lorelei A. Mucci, and Konrad H. Stopsack are Prostate Cancer Foundation Young Investigators. The funders did not play a role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

CONFLICT OF INTEREST STATEMENT

Lorelei A. Mucci receives research funding from AstraZeneca to Harvard University and is on the Scientific Advisory Board and holds equity in Convergent Therapeutics; none of these activities are related to this review. Tamara L. Lotan reports other support from AstraZeneca, grants from AIRA Matrix, grants from Artera, and non-financial support from MDxHealth/Exact Biosciences outsides of the submitted work. The other authors declare no relevant conflicts of interest.

DATA AVAILABILITY STATEMENT

The underlying data in this systematic review are publicly available in the published literature and are fully reported in the main tables and the supplement. Further information is available from the corresponding author upon request.

ORCID

Colleen B. McGrath https://orcid.org/0000-0003-4542-2955
Alaina H. Shreves https://orcid.org/0000-0002-0127-4391
Konrad H. Stopsack https://orcid.org/0000-0002-0722-1311

REFERENCES

1. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*. 2005; 310(5748):644-648. doi:10.1126/science.1117679

2. Nicholas TR, Strittmatter BG, Hollenhorst PC. Oncogenic ETS factors in prostate cancer. *Adv Exp Med Biol*. 2019;1210:409-436. doi:10.1007/978-3-030-32656-2_18
3. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. *Nat Rev Cancer*. 2017;17(6):337-351. doi:10.1038/nrc.2017.20
4. The Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell*. 2015;163(4):1011-1025. doi:10.1016/j.cell.2015.10.025
5. Pettersson A, Graff RE, Bauer SR, et al. The TMPRSS2:ERG rearrangement, ERG expression, and prostate cancer outcomes: a cohort study and meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2012;21(9):1497-1509. doi:10.1158/1055-9965.EPI-12-0042
6. Zhou CK, Young D, Yeboah ED, et al. TMPRSS2:ERG gene fusions in prostate cancer of west African men and a meta-analysis of racial differences. *Am J Epidemiol*. 2017;186(12):1352-1361. doi:10.1093/aje/kwx235
7. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024;74(1):12-49. doi:10.3322/caac.21820
8. Findlay VJ, LaRue AC, Turner DP, Watson PM, Watson DK. Understanding the role of ETS-mediated gene regulation in complex biological processes. *Adv Cancer Res*. 2013;119:1-61. doi:10.1016/B978-0-12-407190-2.00001-0
9. Mani RS, Tomlins SA, Callahan K, et al. Induced chromosomal proximity and gene fusions in prostate cancer. *Science*. 2009;326(5957):1230. doi:10.1126/science.1178124
10. Setlur SR, Mertz KD, Hoshida Y, et al. Estrogen-dependent signaling in a molecularly distinct subclass of aggressive prostate cancer. *J Natl Cancer Inst*. 2008;100(11):815-825. doi:10.1093/jnci/djn150
11. Meller S, Meyer HA, Bethan B, et al. Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer. *Oncotarget*. 2015;7(2):1421-1438. doi:10.18632/oncotarget.6370
12. Hansen AF, Sandsmark E, Rye MB, et al. Presence of TMPRSS2-ERG is associated with alterations of the metabolic profile in human prostate cancer. *Oncotarget*. 2016;7(27):42071-42085. doi:10.18632/oncotarget.9817
13. Berglund AE, Rounbehler RJ, Gerke T, et al. Distinct transcriptional repertoire of the androgen receptor in ETS fusion-negative prostate cancer. *Prostate Cancer Prostatic Dis*. 2019;22(2):292-302. doi:10.1038/s41391-018-0103-4
14. Stopsack KH, Su XA, Vaselkiv JB, et al. Transcriptomes of prostate cancer with TMPRSS2:ERG and other ETS fusions. *Mol Cancer Res*. 2023;21(1):14-23. doi:10.1158/1541-7786.MCR-22-0446
15. Meisel Sharon S, Pozniak Y, Geiger T, Werner H. TMPRSS2-ERG fusion protein regulates insulin-like growth factor-1 receptor (IGF1R) gene expression in prostate cancer: involvement of transcription factor Sp1. *Oncotarget*. 2016;7(32):51375-51392. doi:10.18632/oncotarget.9837
16. Mancarella C, Casanova-Salas I, Calatrava A, et al. ERG deregulation induces IGF-1R expression in prostate cancer cells and affects sensitivity to anti-IGF-1R agents. *Oncotarget*. 2015;6(18):16611-16622. doi:10.18632/oncotarget.3425
17. Pettersson A, Lis RT, Meisner A, et al. Modification of the association between obesity and lethal prostate cancer by TMPRSS2:ERG. *J Natl Cancer Inst*. 2013;105(24):1881-1890. doi:10.1093/jnci/djt332
18. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25(9):603-605. doi:10.1007/s10654-010-9491-z
19. Maclure M, Schneeweiss S. Causation of bias: the episcope. *Epidemiology*. 2001;12(1):114-122. doi:10.1097/00001648-200101000-00019
20. Figg WD, Chau CH, Price DK, et al. Androgen receptor CAG repeat length and TMPRSS2:ETS prostate cancer risk: results from the prostate cancer prevention trial. *Urology*. 2014;84(1):127-131. doi:10.1016/j.urology.2014.03.015
21. Yoo S, Pettersson A, Jordahl KM, et al. Androgen receptor CAG repeat polymorphism and risk of TMPRSS2:ERG-positive prostate cancer. *Cancer Epidemiol Biomarkers Prev*. 2014;23(10):2027-2031. doi:10.1158/1055-9965.EPI-14-0020
22. Allott EH, Ebot EM, Stopsack KH, et al. Statin use is associated with lower risk of PTEN-null and lethal prostate cancer. *Clin Cancer Res*. 2020;26(5):1086-1093. doi:10.1158/1078-0432.CCR-19-2853
23. Egbers L, Luedeke M, Rinckleb A, et al. Obesity and prostate cancer risk according to tumor TMPRSS2:ERG gene fusion status. *Am J Epidemiol*. 2015;181(9):706-713. doi:10.1093/aje/kwu344
24. Feng X, Song M, Preston MA, et al. The association of diabetes with risk of prostate cancer defined by clinical and molecular features. *Br J Cancer*. 2020;123(4):657-665. doi:10.1038/s41416-020-0910-y
25. Geybels MS, McCloskey KD, Mills IG, Stanford JL. Calcium Channel blocker use and risk of prostate cancer by TMPRSS2:ERG gene fusion status. *Prostate*. 2017;77(3):282-290. doi:10.1002/pros.23267
26. Graff RE, Ahearn TU, Pettersson A, et al. Height, obesity, and the risk of TMPRSS2:ERG-defined prostate cancer. *Cancer Epidemiol Biomarkers Prev*. 2018;27(2):193-200. doi:10.1158/1055-9965.EPI-17-0547
27. Graff RE, Judson G, Ahearn TU, et al. Circulating antioxidant levels and risk of prostate cancer by TMPRSS2:ERG. *Prostate*. 2017;77(6):647-653. doi:10.1002/pros.23312
28. Graff RE, Pettersson A, Lis RT, et al. Dietary lycopene intake and risk of prostate cancer defined by ERG protein expression. *Am J Clin Nutr*. 2016;103(3):851-860. doi:10.3945/ajcn.115.118703
29. Graff RE, Meisner A, Ahearn TU, et al. Pre-diagnostic circulating sex hormone levels and risk of prostate cancer by ERG tumour protein expression. *Br J Cancer*. 2016;114(8):939-944. doi:10.1038/bjc.2016.61
30. Hashim D, Gonzalez-Feliciano AG, Ahearn TU, et al. Family history of prostate cancer and the incidence of ERG- and phosphatase and tensin homolog-defined prostate cancer. *Int J Cancer*. 2020;146(10):2694-2702. doi:10.1002/ijc.32577
31. Khan S, Caldwell J, Wilson KM, et al. Baldness and risk of prostate cancer in the health professionals follow-up study. *Cancer Epidemiol Biomarkers Prev*. 2020;29(6):1229-1236. doi:10.1158/1055-9965.EPI-19-1236
32. Luedeke M, Rinckleb AE, FitzGerald LM, et al. Prostate cancer risk regions at 8q24 and 17q24 are differentially associated with somatic TMPRSS2:ERG fusion status. *Hum Mol Genet*. 2016;25(24):5490-5499. doi:10.1093/hmg/ddw349
33. Luedeke M, Linnert CM, Hofer MD, et al. Predisposition for TMPRSS2-ERG fusion in prostate cancer by variants in DNA repair genes. *Cancer Epidemiol Biomarkers Prev*. 2009;18(11):3030-3035. doi:10.1158/1055-9965.EPI-09-0772
34. Penney KL, Pettersson A, Shui IM, et al. Association of prostate cancer risk variants with TMPRSS2:ERG status: evidence for distinct molecular subtypes. *Cancer Epidemiol Biomarkers Prev*. 2016;25(5):745-749. doi:10.1158/1055-9965.EPI-15-1078
35. Pernar CH, Ebot EM, Pettersson A, et al. A prospective study of the association between physical activity and risk of prostate cancer defined by clinical features and TMPRSS2:ERG. *Eur Urol*. 2019;76(1):33-40. doi:10.1016/j.eururo.2018.09.041
36. Stopsack KH, Gonzalez-Feliciano AG, Peisch SF, et al. A prospective study of aspirin use and prostate cancer risk by TMPRSS2:ERG status. *Cancer Epidemiol Biomarkers Prev*. 2018;27(10):1231-1233. doi:10.1158/1055-9965.EPI-18-0510
37. Wright JL, Chery L, Holt S, et al. Aspirin and NSAID use in association with molecular subtypes of prostate cancer defined by TMPRSS2:ERG fusion status. *Prostate Cancer Prostatic Dis*. 2016;19(1):53-56. doi:10.1038/pcan.2015.49

38. Feng X, Zhou CK, Clish CB, et al. Association of pre-diagnostic blood metabolomics with prostate cancer defined by ERG or PTEN molecular subtypes. *Cancer Epidemiol Biomarkers Prev*. 2021;30(5):1000-1008. doi:10.1158/1055-9965.EPI-20-1363
39. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. A prospective cohort study of vasectomy and prostate cancer in US men. *JAMA*. 1993;269(7):873-877.
40. Rimm EB, Stampfer MJ, Colditz GA, Giovannucci E, Willett WC. Effectiveness of various mailing strategies among nonrespondents in a prospective cohort study. *Am J Epidemiol*. 1990;131(6):1068-1071. doi:10.1093/oxfordjournals.aje.a115598
41. Steering Committee of the Physicians' Health Study Research Group. Final report on the aspirin component of the ongoing Physicians' health study. *N Engl J Med*. 1989;321(3):129-135. doi:10.1056/NEJM198907203210301
42. Tamimi RM, Colditz GA, Hazra A, et al. Traditional breast cancer risk factors in relation to molecular subtypes of breast cancer. *Breast Cancer Res Treat*. 2012;131(1):159-167. doi:10.1007/s10549-011-1702-0
43. Pollak M. The insulin and insulin-like growth factor receptor family in neoplasia: an update. *Nat Rev Cancer*. 2012;12(3):159-169. doi:10.1038/nrc3215
44. Watts EL, Perez-Cornago A, Fensom GK, et al. Circulating insulin-like growth factors and risks of overall, aggressive and early-onset prostate cancer: a collaborative analysis of 20 prospective studies and mendelian randomization analysis. *Int J Epidemiol*. 2023;52(1):71-86. doi:10.1093/ije/dyac124
45. Fu BC, Tabung FK, Pernar CH, et al. Insulinemic and inflammatory dietary patterns and risk of prostate cancer. *Eur Urol*. 2021;79(3):405-412. doi:10.1016/j.eururo.2020.12.030
46. Toretsky JA, Kalebic T, Blakesley V, LeRoith D, Helman LJ. The insulin-like growth factor-I receptor is required for EWS/FLI-1 transformation of fibroblasts. *J Biol Chem*. 1997;272(49):30822-30827. doi:10.1074/jbc.272.49.30822
47. Lin B, Ferguson C, White JT, et al. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Res*. 1999;59(17):4180-4184.
48. Mani RS, Amin MA, Li X, et al. Inflammation induced oxidative stress mediates gene fusion formation in prostate cancer. *Cell Rep*. 2016;17(10):2620-2631. doi:10.1016/j.celrep.2016.11.019
49. Handelsman DJ, Yeap B, Flicker L, Martin S, Wittert GA, Ly LP. Age-specific population centiles for androgen status in men. *Eur J Endocrinol*. 2015;173(6):809-817. doi:10.1530/EJE-15-0380
50. Giovannucci E. A framework to understand diet, physical activity, body weight, and cancer risk. *Cancer Causes Control*. 2018;29(1):1-6. doi:10.1007/s10552-017-0975-y
51. Watts EL, Perez-Cornago A, Fensom GK, et al. Circulating free testosterone and risk of aggressive prostate cancer: prospective and mendelian randomisation analyses in international consortia. *Int J Cancer*. 2022;151(7):1033-1046. doi:10.1002/ijc.34116
52. Genkinger JM, Wu K, Wang M, et al. Measures of body fatness and height in early and mid-to-late adulthood and prostate cancer: risk and mortality in the pooling project of prospective studies of diet and cancer. *Ann Oncol*. 2020;31(1):103-114. doi:10.1016/j.annonc.2019.09.007
53. Hjelmborg JB, Scheike T, Holst K, et al. The heritability of prostate cancer in the Nordic twin study of cancer. *Cancer Epidemiol Biomarkers Prev*. 2014;23(11):2303-2310. doi:10.1158/1055-9965.EPI-13-0568
54. FitzGerald LM, Agalliu I, Johnson K, et al. Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: results from a population-based study of prostate cancer. *BMC Cancer*. 2008;8:230. doi:10.1186/1471-2407-8-230
55. Kohaar I, Li Q, Chen Y, et al. Association of germline genetic variants with TMPRSS2-ERG fusion status in prostate cancer. *Oncotarget*. 2020;11(15):1321-1333. doi:10.18632/oncotarget.27534
56. Ma C, Wang X, Dai JY, et al. Germline genetic variants associated with somatic TMPRSS2:ERG fusion status in prostate cancer: a genome-wide association study. *Cancer Epidemiol Biomarkers Prev*. 2023;32(10):1436-1443. doi:10.1158/1055-9965.EPI-23-0275
57. Nakazawa M, Fang M, Lotan TL, et al. Germline BRCA2, ATM and CHEK2 alterations shape somatic mutation landscapes in prostate cancer. *J Clin Oncol*. 2022;40(6_suppl):148. doi:10.1200/JCO.2022.40.6_suppl.148
58. Chalmers ZR, Burns MC, Ebot EM, et al. Early-onset metastatic and clinically advanced prostate cancer is a distinct clinical and molecular entity characterized by increased TMPRSS2-ERG fusions. *Prostate Cancer Prostatic Dis*. 2021;24(2):558-566. doi:10.1038/s41391-020-00314-z

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: McGrath CB, Shreves AH, Shanahan MR, et al. Etiology of prostate cancer with the *TMPRSS2:ERG* fusion: A systematic review of risk factors. *Int J Cancer*. 2024;1-11. doi:10.1002/ijc.35279