JNCI J Natl Cancer Inst (2021) 113(6): djaa154

doi: 10.1093/jnci/djaa154
First published online October 3, 2020
Article

Recommended Definitions of Aggressive Prostate Cancer for Etiologic Epidemiologic Research

Lauren M. Hurwitz, PhD,[1,*] Ilir Agalliu, MD, ScD,[2] Demetrius Albanes, MD,[1] Kathryn Hughes Barry, PhD,[3,4] Sonja I. Berndt, PharmD, PhD,[1] Qiuyin Cai, MD, PhD,[5] Chu Chen, PhD,[6] Iona Cheng, PhD,[7] Jeanine M. Genkinger, PhD,[8] Graham G. Giles, PhD,[9,10,11] Jiaqi Huang, PhD,[1] Corinne E. Joshu, PhD,[12] Tim J. Key, DPhil,[13] Synnove Knutsen, MD, PhD,[14] Stella Koutros, PhD,[1] Hilde Langseth, PhD,[15,16] Sherly X. Li, PhD,[9,10,17] Robert J. MacInnis, PhD,[9,10] Sarah C. Markt, ScD,[18] Kathryn L. Penney, ScD,[19,20] Aurora Perez-Cornago, PhD,[13] Thomas E. Rohan, MBBS, PhD,[2] Stephanie A. Smith-Warner, PhD,[20,21] Meir J. Stampfer, MD, DrPH,[20] Konrad H. Stopsack, MD,[22] Catherine M. Tangen, DrPH,[23] Ruth C. Travis, DPhil,[13] Stephanie J. Weinstein, PhD,[1] Lang Wu, PhD,[24] Eric J. Jacobs, PhD,[25,†] Lorelei A. Mucci, ScD,[20,†] Elizabeth A. Platz, ScD,[12,†] Michael B. Cook, PhD[1,†]; on behalf of the Prostate Cancer Cohort Consortium (PC3) Working Group

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA; [2]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA; [3]Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA; [4]Program in Oncology, University of Maryland Marlene and Stewart Greenebaum Comprehensive Cancer Center, Baltimore, MD, USA; [5]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA; [6]Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [7]Department of Epidemiology and Biostatistics, Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA, USA; [8]Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA; [9]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia; [10]Centre for Epidemiology and Biostatistics, University of Melbourne, Parkville, VIC, Australia; [11]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia; [12]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; [13]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK; [14]School of Public Health, Loma Linda University, Loma Linda, CA, USA; [15]Department of Research, Cancer Registry of Norway, Oslo, Norway; [16]Department of Epidemiology and Biostatistics, Imperial College London, London, UK; [17]Medical Research Council Epidemiology Unit, University of Cambridge, Cambridge, UK; [18]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA; [19]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital/Harvard Medical School, Boston, MA, USA; [20]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [21]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA; [22]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA; [23]SWOG Statistics and Data Management Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [24]Cancer Epidemiology Division, Population Sciences in the Pacific Program, University of Hawaii Cancer Center, University of Hawaii at Manoa, Honolulu, HI, USA and [25]Behavioral and Epidemiology Research Group, American Cancer Society, Atlanta, GA, USA

†These authors share senior authorship.

*Correspondence to: Lauren M. Hurwitz, PhD, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Dr, Rockville, MD 20850 (e-mail: lauren.hurwitz@nih.gov).

Abstract

Background: In the era of widespread prostate-specific antigen testing, it is important to focus etiologic research on the outcome of aggressive prostate cancer, but studies have defined this outcome differently. We aimed to develop an evidence-based consensus definition of aggressive prostate cancer using clinical features at diagnosis for etiologic epidemiologic research. **Methods:** Among prostate cancer cases diagnosed in 2007 in the National Cancer Institute's Surveillance, Epidemiology, and End Results-18 database with follow-up through 2017, we compared the performance of categorizations of aggressive prostate cancer in discriminating fatal prostate cancer within 10 years of diagnosis, placing the most emphasis on sensitivity and positive predictive value (PPV). **Results:** In our case population (n = 55 900), 3073 men died of prostate cancer within 10 years. Among 12 definitions that included TNM staging and Gleason score, sensitivities ranged from 0.64 to 0.89 and PPVs ranged from 0.09 to 0.23. We propose defining aggressive prostate cancer as diagnosis of category T4 or N1 or M1 or Gleason score of 8 or greater prostate cancer, because this definition had one of the higher PPVs (0.23, 95% confidence interval = 0.22 to 0.24) and reasonable sensitivity (0.66, 95% confidence interval = 0.64 to 0.67) for prostate cancer death within 10 years. Results were similar across sensitivity analyses. **Conclusions:** We recommend that etiologic epidemiologic studies

Received: June 5, 2020; **Revised:** August 7, 2020; **Accepted:** September 15, 2020

of prostate cancer report results for this definition of aggressive prostate cancer. We also recommend that studies separately report results for advanced category (T4 or N1 or M1), high-grade (Gleason score ≥8), and fatal prostate cancer. Use of this comprehensive set of endpoints will facilitate comparison of results from different studies and help elucidate prostate cancer etiology.

Since the advent of prostate cancer screening with the prostate-specific antigen (PSA) test, the epidemiology of this malignancy has changed dramatically. Prostate cancer mortality has declined in several countries (1); in the United States, for example, mortality declined by 51% between 1993 and 2016 (2). Concurrently, many countries have also experienced sharp increases in prostate cancer incidence because of the diagnosis of asymptomatic disease that had previously gone undetected (3,4). A large proportion of screen-detected prostate cancers are now thought to be clinically indolent and overdiagnosed (5,6), because they are unlikely to progress or cause harm to a man during his lifetime. However, some prostate cancers have greater potential to be lethal. This clinical heterogeneity has posed a persistent challenge to researchers seeking to understand etiologic differences between indolent and aggressive prostate cancers (7) and to clinicians seeking to optimize treatment benefits and minimize treatment harms (8). Though clinicians may use risk stratification tools, nomograms, and genomic classifiers to identify cases that are likely to be aggressive, these tools are not uniformly available in population-based studies, and a variety of outcome definitions based on clinical parameters are typically used instead.

From an etiology and prevention perspective, there is a strong need to accurately identify aggressive prostate cancers and study them as a distinct outcome. Risk factors for aggressive prostate cancer, defined based on disease stage, grade, or long-term outcomes, often differ from risk factors for early-stage or total prostate cancer (9,10). For example, common cancer risk factors such as obesity and cigarette smoking have been positively associated with fatal prostate cancer but not total prostate cancer risk (11,12). Risk factors for total prostate cancer have also differed for cases diagnosed in the pre-PSA era, during which most prostate cancers were clinically detectable (eg, palpable or symptomatic), and for cases diagnosed in the era since PSA testing became widespread, during which asymptomatic prostate cancers have predominated (9). These findings imply that indolent and aggressive prostate cancers have distinct risk factor profiles and that separate analyses of indolent and aggressive prostate cancer may be necessary to elucidate prostate cancer etiology.

Despite this need, there is currently no standardized approach for defining the outcome of aggressive prostate cancer for use in etiologic epidemiologic research. Epidemiologic studies often use various combinations of clinical parameters to define aggressive prostate cancer, including stage, Gleason score, and diagnostic PSA value, because these data fields are most commonly available in population-based studies. However, the exact outcome definition varies from study to study, limiting the ability to compare study results and combine results in meta-analyses. There have been efforts to pool cohort study data and harmonize outcomes across studies, but these post hoc efforts rely on individual studies to have collected the data needed to consistently define the outcome, and various pooling efforts have not always used the same outcome definitions (13–16).

The goal of this article is to propose a widely applicable consensus outcome definition for aggressive prostate cancer for use in etiologic epidemiologic research, developed based on the hard endpoint of prostate cancer death. By promoting use of a common, evidence-based definition, we hope to increase comparability across studies to improve understanding of prostate cancer etiology and inform strategies for prevention.

Methods

Study Population

We used data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program. SEER collects incidence and survival data from population-based cancer registries from several states and metropolitan regions across the United States. Information on all deaths occurring in the United States is obtained annually from the National Center for Health Statistics; the most recent SEER submission includes deaths through 2017. Our analytic population comprised men in the SEER-18 database (17) who were diagnosed with prostate cancer (primary site C619) in 2007, the most recent year with 10 years of follow-up in the PSA era to assess long-term outcomes. Autopsy-only or death certificate–only cases and cases missing age were excluded. We further excluded cases missing American Joint Committee on Cancer (AJCC) clinical T category or primary or secondary Gleason pattern (n = 6493 cases excluded, 10.4% of cases otherwise eligible). Cases missing AJCC N or M category (ie, Nx or Mx) were assumed to be N0 or M0, respectively.

Definitions of Aggressive Prostate Cancer

Within our analytic dataset of prostate cancer cases, we examined the performance of various categorizations of clinical category (≥T3, T4, N1, and M1) and Gleason score (≥3 + 4, ≥4 + 3, and ≥8, corresponding to grade groups ≥2, ≥3, and ≥4, respectively). To develop a definition that is widely applicable for epidemiologic research, we focused on clinical factors that are commonly collected in prospective cohort studies, as determined via a brief, email-based survey that we sent to cohorts participating in the NCI Cohort Consortium (Supplementary Table 1, Supplementary Figure 1, available online) (18). This consortium includes cohorts from 4 continents and is representative of studies that we encourage to report etiologic results using our definition. Only approximately one-half of all cohorts surveyed collected information on diagnostic PSA value, and of the cohorts with PSA data, many had high proportions of missingness. As a result, we elected to leave PSA out of the definitions tested in our primary analysis.

The AJCC 6th edition was used for clinical staging. Pathologic TNM staging was used, when available, in a sensitivity analysis. In all analyses, Gleason score was recorded as either clinical grade or pathologic grade, because separate variables for clinical and pathologic Gleason patterns were not available in SEER before 2010.

Table 1. Performance of the definitions for discriminating fatal prostate cancer within 10 years of diagnosis among prostate cancer cases diagnosed in 2007 in SEER-18

Definition of aggressive prostate cancer	Cases that meet the definition		Sensitivity (95% CI)	PPV (95% CI)
	No. (%)[a]	No. Fatal within 10 years		
Based on clinical stage only				
≥T3 or N1 or M1	3123 (5.9)	1158	0.38 (0.36 to 0.39)	0.37 (0.35 to 0.39)
T4 or N1 or M1	2127 (3.8)	1013	0.33 (0.31 to 0.35)	0.48 (0.45 to 0.50)
N1 or M1	1937 (3.5)	956	0.31 (0.29 to 0.33)	0.49 (0.47 to 0.52)
M1	1316 (2.4)	821	0.27 (0.25 to 0.28)	0.62 (0.60 to 0.65)
Based on Gleason score[b] only				
≥3 + 4	29 906 (53.5)	2703	0.88 (0.87 to 0.89)	0.09 (0.09 to 0.09)
≥4 + 3	13 984 (25.0)	2228	0.73 (0.71 to 0.74)	0.16 (0.15 to 0.17)
≥8	8095 (14.5)	1800	0.59 (0.57 to 0.60)	0.22 (0.21 to 0.23)
Based on combinations of clinical stage and Gleason score[b]				
≥T3 or N1 or M1 or GS >3 + 4	30 159 (54.0)	2727	0.89 (0.88 to 0.90)	0.09 (0.09 to 0.09)
T4 or N1 or M1 or GS >3 + 4	30 051 (53.8)	2724	0.89 (0.87 to 0.90)	0.09 (0.09 to 0.09)
N1 or M1 or GS >3 + 4	30 027 (53.7)	2722	0.89 (0.87 to 0.90)	0.09 (0.09 to 0.09)
M1 or GS >3 + 4	29 981 (53.6)	2718	0.88 (0.87 to 0.90)	0.09 (0.09 to 0.09)
≥T3 or N1 or M1 or GS >4 + 3	14 789 (26.5)	2358	0.77 (0.75 to 0.78)	0.16 (0.15 to 0.17)
T4 or N1 or M1 or GS >4 + 3	14 422 (25.8)	2337	0.76 (0.74 to 0.78)	0.16 (0.16 to 0.17)
N1 or M1 or GS >4 + 3	14 371 (25.7)	2329	0.76 (0.74 to 0.77)	0.16 (0.16 to 0.17)
M1 or GS >4 + 3	14 210 (25.4)	2321	0.76 (0.74 to 0.77)	0.16 (0.16 to 0.17)
≥T3 or N1 or M1 or GS >8	9351 (16.7)	2060	0.67 (0.65 to 0.69)	0.22 (0.21 to 0.23)
T4 or N1 or M1 or GS >8	8805 (15.8)	2013	0.66 (0.64 to 0.67)	0.23 (0.22 to 0.24)
N1 or M1 or GS >8	8738 (15.6)	2001	0.65 (0.63 to 0.67)	0.23 (0.22 to 0.24)
M1 or GS >8	8469 (15.2)	1975	0.64 (0.63 to 0.66)	0.23 (0.22 to 0.24)

[a]Percent of cases that meet the definition, out of the 55 900 cases from SEER-18 included in the analysis. CI = confidence interval; GS = Gleason score; PPV = positive predictive value.
[b]If using the new grade groups (GG) instead of Gleason score, GS 3 + 4 = GG 2, GS 4 + 3 = GG 3, GS 8 = GG 4.

Comparison Outcome

In the absence of an established "gold standard" definition for aggressive prostate cancer, we chose to validate definitions against the outcome of fatal prostate cancer within 10 years of diagnosis, an outcome that is clearly clinically relevant. Fatal prostate cancer was defined as death with prostate cancer listed as the underlying cause according to the death certificate. The 10-year window for assessing fatal prostate cancer was based in part on the observation that more than 70% of all prostate cancer deaths are captured within 10 years of diagnosis (17,19). Although the choice of any time window is somewhat arbitrary, 10-year disease risk estimates are familiar to many clinicians, and we believe 10 years is a reasonable time frame to consider given that prostate cancer is typically diagnosed late in life.

Statistical Analysis

For each clinical factor and for combinations of clinical factors, we calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value for fatal prostate cancer within 10 years of diagnosis. We also calculated the area under the receiver operating characteristic curve (AUC). We made an a priori decision to prioritize sensitivity and PPV when selecting the final definition, because we wanted a definition that could both capture a high proportion of all cases with life-threatening potential (sensitivity) and maximize the likelihood that cases meeting the definition were truly life-threatening (PPV). Ultimately, we restricted our selection to definitions with reasonable sensitivity, which we defined as sensitivity greater than or equal to 0.60, but prioritized PPV over sensitivity because we wanted to identify a case group relatively highly enriched for cases with life-threatening disease. We also carried out analyses stratified by age at diagnosis (<60, 60-69, ≥70 years) and by race and ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, non-Hispanic Asian or Pacific Islander) to ensure that definitions performed similarly across these subpopulations.

We conducted sensitivity analyses to examine the impact of using pathologic stage, when available, instead of clinical stage, adding diagnostic PSA values to the definitions (using cutoff points of >10 ng/mL and >20 ng/mL), varying the time interval for assessing fatal prostate cancer, and using multiple imputation with chained equations to retain cases missing clinical T category or Gleason score. Analyses were conducted in SAS 9.4, Stata 15, and R 3.5.1.

Results

Our study population included 55 900 men diagnosed with prostate cancer in SEER-18 in 2007. The study population was 72.3% non-Hispanic White, 13.7% non-Hispanic Black, 7.8% Hispanic, 4.6% non-Hispanic Asian or Pacific Islander, and 1.6% other or unknown. At diagnosis, 24.6% of cases were younger than 60 years, 38.7% were 60-69 years, and 36.7% were older than 70 years. Among these cases, 3073 men (5.5%) died from prostate cancer within 10 years of diagnosis (Supplementary Table 2, available online).

Definitions of aggressive prostate cancer based solely on clinical stage had relatively high PPV (range = 0.37-0.62) but low sensitivity, ranging from 0.27 (95% confidence interval [CI] = 0.25 to 0.28) for M1 disease alone to 0.38 (95% CI = 0.36 to 0.39) for T3 or greater or N1 or M1 disease (Table 1). In contrast, definitions based solely on Gleason scores of 3+4 or greater or 4+3

or greater had higher sensitivity (0.88, 95% CI = 0.87 to 0.89; and 0.73, 95% CI = 0.71 to 0.74, respectively) but lower PPV (0.09, 95% CI = 0.09 to 0.09; and 0.16, 95% CI = 0.15 to 0.17, respectively). The definition based solely on a Gleason score of 8 or higher did achieve a higher PPV (0.22, 95% CI = 0.21 to 0.23) but at the cost of reduced sensitivity (0.59, 95% CI = 0.57 to 0.60). Other performance measures, including specificity, negative predictive value, and AUC, are shown in Supplementary Table 3 (available online).

When we examined definitions of aggressive prostate cancer based on combinations of clinical stage and Gleason score, sensitivities ranged from 0.64 to 0.89, whereas PPVs ranged from 0.09 to 0.23 (Table 1). Gleason score was the primary determinant of the sensitivity and PPV of each definition. Definitions including Gleason score of 3 or greater + 4 had the highest sensitivity, and definitions including Gleason score greater than or equal to 8 had lower sensitivities but achieved PPVs of 0.22 or greater. Of note, including pathologic stage instead of clinical stage, when available, did not generally alter the sensitivity or PPV of these definitions (Supplementary Table 4, available online). Adding high diagnostic total PSA as a criterion to the definitions increased the sensitivity but decreased the PPV (Supplementary Table 5, available online). As the time interval for identifying prostate cancer deaths increased from 2 to 10 years, the sensitivity of the definitions decreased while the PPV of the definitions increased, but the relative rankings of the definitions for all performance measures remained the same (Supplementary Figure 2, available online). In the analysis with multiple imputation, the sensitivity and PPV of all definitions improved (Supplementary Table 6, available online).

When choosing a consensus definition, among the definitions with sensitivity greater than or equal to 0.60, we first narrowed our choices to definitions 9-12 (ie, those incorporating Gleason score ≥8), because these definitions had the highest PPVs. These definitions had the added advantage of not needing to distinguish Gleason 3 + 4 from Gleason 4 + 3 disease, which is important given that many existing prospective cohort studies have not collected, and are not collecting, information on primary and secondary Gleason pattern (Supplementary Table 1, Supplementary Figure 1, available online). Of these 4 definitions, definition 10 (category T4 or N1 or M1 or Gleason ≥8) had one of the highest PPVs (0.23, 95% CI = 0.22 to 0.24) and maintained a reasonably high sensitivity (0.66, 95% CI = 0.64 to 0.67) for discriminating fatal prostate cancer within 10 years of diagnosis. Definitions 11 and 12 had similar PPVs (0.23, 95% CI = 0.22 to 0.24 for both) but slightly lower sensitivities (0.65, 95% CI = 0.63 to 0.67; and 0.64, 95% CI = 0.63 to 0.66, respectively). Definition 9 had a higher sensitivity (0.67, 95% CI = 0.65 to 0.69) at the cost of a lower PPV (0.22, 95% CI = 0.21 to 0.23) but did not perform as well when pathologic stage was used, when available, instead of clinical stage (PPV = 0.18, 95% CI = 0.17 to 0.19; Supplementary Table 4, available online). Definition 10 performed similarly across subgroups based on race and ethnicity (Table 2) and age at diagnosis (Table 3). As a result, we selected definition 10 (category T4 or N1 or M1 or Gleason ≥8) as the final recommended definition for aggressive prostate cancer for etiologic epidemiologic research.

Discussion

Based on our analysis within SEER-18, in the setting of etiologic epidemiologic research, we recommend defining the outcome of aggressive prostate cancer as cancers that are clinical or pathologic category T4 or N1 or M1 or Gleason score greater than or equal to 8 (or if using the new grade group system instead of Gleason score, grade group ≥4). This definition had a good balance of sensitivity and PPV for capturing prostate cancers that were fatal within 10 years of diagnosis. This definition also performed similarly across subgroups defined by diagnostic age and race and ethnicity and uses data fields that are commonly collected in cohort studies. We recommend that epidemiologic studies report results for this definition of prostate cancer to enrich for aggressive disease and improve comparability of study results.

For this analysis, we tested potential definitions against the outcome of fatal prostate cancer. Fatal prostate cancer can, itself, be assessed as an outcome in prospective cohort studies and may be considered the current gold-standard outcome for studying prostate cancer that is undoubtedly aggressive. We recommend that studies continue to publish results for fatal prostate cancer, when possible. However, there are limitations to studying fatal prostate cancer in etiologic research that our definition seeks to overcome. Ascertaining enough fatal prostate cancers to provide sufficient power for a study requires long-term follow-up, which in some studies may be prohibitive in terms of time and cost. The outcome of fatal prostate cancer also does not capture aggressive cancers that are not fatal, potentially because of successful early diagnosis and treatment or competing causes of death. A strength of our definition is that it relies solely on disease characteristics measurable at diagnosis, allowing it to capture prostate cancers with life-threatening potential regardless of whether they ultimately prove fatal. The study of aggressive prostate cancer can thus complement research of fatal prostate cancer and elucidate the etiology of this distinct case group.

We also considered incorporating fatal prostate cancer into our definition (ie, defining aggressive prostate cancer as category T4 or N1 or M1 or Gleason score ≥8 or fatal). Adding in the fatal prostate cancers would help to reduce false negatives and capture the 34.5% of prostate cancer deaths that were otherwise missed by our final recommended definition. However, we ultimately chose against this decision because we wanted our definition to be standardized, both across population-based studies and across individuals within a study with differing lengths of follow-up for mortality. We were also concerned about potential bias that may arise if both diagnostic criteria and follow-up information are used to define aggressive prostate cancer at diagnosis. If an exposure is associated with overall mortality, for example, exposed case patients might be more likely to die of competing causes of death and less likely to die of prostate cancer and thus less likely to be retroactively classified as having aggressive disease at diagnosis, regardless of whether the exposure influences risk of aggressive prostate cancer. We hope that future work will empirically compare the advantages and disadvantages of incorporating postdiagnostic data into definitions of aggressive prostate cancer. For these reasons, though, we recommend defining aggressive prostate cancer based on diagnostic criteria only and analyzing fatal prostate cancer as a separate outcome.

Several approaches exist for identifying potentially aggressive prostate cancers in the clinical setting. Clinicians use risk stratification tools such as D'Amico risk groups, risk scores, or nomograms to assess prognosis and guide treatment decisions for their patients. These tools often incorporate stage, grade, and PSA values, as well as additional clinical parameters such as the percentage of biopsy cores that are positive and the percentage of cancer within each core, and have been shown to

Table 2. Performance of the definitions for discriminating fatal prostate cancer within 10 years of diagnosis by race and ethnicity among prostate cancer cases diagnosed in 2007 in SEER-18[a]

Definition	Cases that meet the definition		Sensitivity (95% CI)	PPV (95% CI)
	No. (%)	No. Fatal		
Non-Hispanic White (40 424 cases, 2150 deaths)				
≥T3 or N1 or M1 or GS >3 + 4	21 497 (53.2)	1913	0.89 (0.88 to 0.90)	0.09 (0.09 to 0.09)
T4 or N1 or M1 or GS >3 + 4	21 426 (53.0)	1910	0.89 (0.87 to 0.90)	0.09 (0.09 to 0.09)
N1 or M1 or GS >3 + 4	21 414 (53.0)	1909	0.89 (0.87 to 0.90)	0.09 (0.09 to 0.09)
M1 or GS >3 + 4	21 383 (52.9)	1907	0.89 (0.87 to 0.90)	0.09 (0.09 to 0.09)
≥T3 or N1 or M1 or GS >4 + 3	10 406 (25.7)	1646	0.77 (0.75 to 0.78)	0.16 (0.15 to 0.17)
T4 or N1 or M1 or GS >4 + 3	10 152 (25.1)	1630	0.76 (0.74 to 0.78)	0.16 (0.15 to 0.17)
N1 or M1 or GS >4 + 3	10 123 (25.0)	1624	0.76 (0.74 to 0.77)	0.16 (0.15 to 0.17)
M1 or GS >4 + 3	10 019 (24.8)	1618	0.75 (0.73 to 0.77)	0.16 (0.15 to 0.17)
≥T3 or N1 or M1 or GS >8	6507 (16.1)	1426	0.66 (0.64 to 0.68)	0.22 (0.21 to 0.23)
T4 or N1 or M1 or GS >8	6121 (15.1)	1393	0.65 (0.63 to 0.67)	0.23 (0.22 to 0.24)
N1 or M1 or GS >8	6081 (15.0)	1385	0.64 (0.62 to 0.66)	0.23 (0.22 to 0.24)
M1 or GS >8	5893 (14.6)	1364	0.63 (0.61 to 0.65)	0.23 (0.22 to 0.24)
Non-Hispanic Black (7672 cases, 492 deaths)				
≥T3 or N1 or M1 or GS >3 + 4	4423 (57.7)	439	0.89 (0.86 to 0.92)	0.10 (0.09 to 0.11)
T4 or N1 or M1 or GS >3 + 4	4407 (57.4)	439	0.89 (0.86 to 0.92)	0.10 (0.09 to 0.11)
N1 or M1 or GS >3 + 4	4400 (57.4)	438	0.89 (0.86 to 0.92)	0.10 (0.09 to 0.11)
M1 or GS >3 + 4	4394 (57.3)	436	0.89 (0.85 to 0.91)	0.10 (0.09 to 0.11)
≥T3 or N1 or M1 or GS >4 + 3	2153 (28.1)	382	0.78 (0.74 to 0.81)	0.18 (0.16 to 0.19)
T4 or N1 or M1 or GS >4 + 3	2094 (27.3)	380	0.77 (0.73 to 0.81)	0.18 (0.17 to 0.20)
N1 or M1 or GS >4 + 3	2081 (27.1)	378	0.77 (0.73 to 0.80)	0.18 (0.17 to 0.20)
M1 or GS >4 + 3	2058 (26.8)	376	0.76 (0.72 to 0.80)	0.18 (0.17 to 0.20)
≥T3 or N1 or M1 or GS >8	1368 (17.8)	339	0.69 (0.65 to 0.73)	0.25 (0.23 to 0.27)
T4 or N1 or M1 or GS >8	1288 (16.8)	331	0.67 (0.63 to 0.71)	0.26 (0.23 to 0.28)
N1 or M1 or GS >8	1274 (16.6)	328	0.67 (0.62 to 0.71)	0.26 (0.23 to 0.28)
M1 or GS >8	1242 (16.2)	325	0.66 (0.62 to 0.70)	0.26 (0.24 to 0.29)
Hispanic (4350 cases, 282 deaths)				
1, ≥T3 or N1 or M1 or GS >3 + 4	2289 (52.6)	243	0.86 (0.82 to 0.90)	0.11 (0.09 to 0.12)
T4 or N1 or M1 or GS >3 + 4	2277 (52.3)	243	0.86 (0.82 to 0.90)	0.11 (0.09 to 0.12)
N1 or M1 or GS >3 + 4	2274 (52.3)	243	0.86 (0.82 to 0.90)	0.11 (0.09 to 0.12)
M1 or GS >3 + 4	2268 (52.1)	243	0.86 (0.82 to 0.90)	0.11 (0.09 to 0.12)
≥T3 or N1 or M1 or GS >4 + 3	1216 (28.0)	212	0.75 (0.70 to 0.80)	0.17 (0.15 to 0.20)
T4 or N1 or M1 or GS >4 + 3	1183 (27.2)	211	0.75 (0.69 to 0.80)	0.18 (0.16 to 0.20)
N1 or M1 or GS >4 + 3	1177 (27.1)	211	0.75 (0.69 to 0.80)	0.18 (0.16 to 0.20)
M1 or GS >4 + 3	1152 (26.5)	211	0.75 (0.69 to 0.80)	0.18 (0.16 to 0.21)
≥T3 or N1 or M1 or GS >8	815 (18.7)	188	0.67 (0.61 to 0.72)	0.23 (0.20 to 0.26)
T4 or N1 or M1 or GS >8	767 (17.6)	185	0.66 (0.60 to 0.71)	0.24 (0.21 to 0.27)
N1 or M1 or GS >8	758 (17.4)	184	0.65 (0.59 to 0.71)	0.24 (0.21 to 0.27)
M1 or GS >8	722 (16.6)	182	0.65 (0.59 to 0.70)	0.25 (0.22 to 0.29)
Non-Hispanic Asian or Pacific Islander (2575 cases, 134 deaths)				
≥T3 or N1 or M1 or GS >3 + 4	1537 (59.7)	118	0.88 (0.81 to 0.93)	0.08 (0.06 to 0.09)
T4 or N1 or M1 or GS >3 + 4	1529 (59.4)	118	0.88 (0.81 to 0.93)	0.08 (0.06 to 0.09)
N1 or M1 or GS >3 + 4	1527 (59.3)	118	0.88 (0.81 to 0.93)	0.08 (0.06 to 0.09)
M1 or GS >3 + 4	1524 (59.2)	118	0.88 (0.81 to 0.93)	0.08 (0.06 to 0.09)
≥T3 or N1 or M1 or GS >4 + 3	849 (33.0)	109	0.81 (0.74 to 0.88)	0.13 (0.11 to 0.15)
T4 or N1 or M1 or GS >4 + 3	829 (32.2)	107	0.80 (0.72 to 0.86)	0.13 (0.11 to 0.15)
N1 or M1 or GS >4 + 3	827 (32.1)	107	0.80 (0.72 to 0.86)	0.13 (0.11 to 0.15)
M1 or GS >4 + 3	819 (31.8)	107	0.80 (0.72 to 0.86)	0.13 (0.11 to 0.16)
≥T3 or N1 or M1 or GS >8	574 (22.3)	101	0.75 (0.67 to 0.82)	0.18 (0.15 to 0.21)
T4 or N1 or M1 or GS >8	545 (21.2)	98	0.73 (0.65 to 0.80)	0.18 (0.15 to 0.21)
N1 or M1 or GS >8	542 (21.0)	98	0.73 (0.65 to 0.80)	0.18 (0.15 to 0.22)
M1 or GS >8	531 (20.6)	98	0.73 (0.65 to 0.80)	0.18 (0.15 to 0.22)

[a]CI = confidence interval; GS = Gleason score; PPV = positive predictive value.

perform well in discriminating fatal prostate cancer (AUCs ranging from 0.73 to 0.81) (20). Genomic classifiers and other biomarkers may also be used in clinical practice to aid risk stratification (21). For comparing therapies, clinicians have validated the outcome of metastasis-free survival as a surrogate endpoint for overall survival for patients with prostate cancer (22). However, although these tools and endpoints are clinically useful, they are generally not practical in population-based

Table 3. Performance of the definitions for discriminating fatal prostate cancer within 10 years of diagnosis, by age at diagnosis, among prostate cancer cases diagnosed in 2007 in SEER-18[a]

Definition	Cases that meet definition		Sensitivity (95% CI)	PPV (95% CI)
	No. (%)	No. Fatal		
<60 y (13 749 cases, 491 deaths)				
≥T3 or N1 or M1 or GS >3 + 4	6770 (49.2)	457	0.93 (0.9 to 0.95)	0.07 (0.07 to 0.07)
T4 or N1 or M1 or GS >3 + 4	6743 (49.0)	457	0.93 (0.9 to 0.95)	0.07 (0.06 to 0.07)
N1 or M1 or GS >3 + 4	6737 (49.0)	456	0.93 (0.9 to 0.95)	0.07 (0.06 to 0.07)
M1 or GS >3 + 4	6716 (48.8)	453	0.92 (0.9 to 0.94)	0.07 (0.06 to 0.07)
≥T3 or N1 or M1 or GS >4 + 3	2565 (18.7)	416	0.85 (0.81 to 0.88)	0.16 (0.15 to 0.18)
T4 or N1 or M1 or GS >4 + 3	2470 (18.0)	412	0.84 (0.8 to 0.87)	0.17 (0.15 to 0.18)
N1 or M1 or GS >4 + 3	2458 (17.9)	411	0.84 (0.8 to 0.87)	0.17 (0.15 to 0.18)
M1 or GS >4 + 3	2392 (17.4)	408	0.83 (0.79 to 0.86)	0.17 (0.16 to 0.19)
≥T3 or N1 or M1 or GS >8	1543 (11.2)	380	0.77 (0.73 to 0.81)	0.25 (0.22 to 0.27)
T4 or N1 or M1 or GS >8	1405 (10.2)	370	0.75 (0.71 to 0.79)	0.26 (0.24 to 0.29)
N1 or M1 or GS >8	1388 (10.1)	367	0.75 (0.71 to 0.79)	0.26 (0.24 to 0.29)
M1 or GS >8	1289 (9.4)	361	0.74 (0.69 to 0.77)	0.28 (0.26 to 0.31)
60–69 y (21 646 cases, 866 deaths)				
≥T3 or N1 or M1 or GS >3 + 4	11 313 (52.3)	757	0.87 (0.85 to 0.9)	0.07 (0.06 to 0.07)
T4 or N1 or M1 or GS >3 + 4	11 263 (52.0)	756	0.87 (0.85 to 0.89)	0.07 (0.06 to 0.07)
N1 or M1 or GS >3 + 4	11 255 (52.0)	755	0.87 (0.85 to 0.89)	0.07 (0.06 to 0.07)
M1 or GS >3 + 4	11 237 (51.9)	754	0.87 (0.85 to 0.89)	0.07 (0.06 to 0.07)
≥T3 or N1 or M1 or GS >4 + 3	5097 (23.5)	653	0.75 (0.72 to 0.78)	0.13 (0.12 to 0.14)
T4 or N1 or M1 or GS >4 + 3	4924 (22.7)	642	0.74 (0.71 to 0.77)	0.13 (0.12 to 0.14)
N1 or M1 or GS >4 + 3	4907 (22.7)	640	0.74 (0.71 to 0.77)	0.13 (0.12 to 0.14)
M1 or GS >4 + 3	4841 (22.4)	636	0.73 (0.7 to 0.76)	0.13 (0.12 to 0.14)
≥T3 or N1 or M1 or GS >8	3060 (14.1)	572	0.66 (0.63 to 0.69)	0.19 (0.17 to 0.2)
T4 or N1 or M1 or GS >8	2806 (13.0)	549	0.63 (0.6 to 0.67)	0.20 (0.18 to 0.21)
N1 or M1 or GS >8	2783 (12.9)	546	0.63 (0.6 to 0.66)	0.20 (0.18 to 0.21)
M1 or GS >8	2661 (12.3)	532	0.61 (0.58 to 0.65)	0.20 (0.18 to 0.22)
≥70 y (20 505 cases, 1716 deaths)				
≥T3 or N1 or M1 or GS >3 + 4	12 076 (58.9)	1513	0.88 (0.87 to 0.9)	0.13 (0.12 to 0.13)
T4 or N1 or M1 or GS >3 + 4	12 045 (58.7)	1511	0.88 (0.86 to 0.9)	0.13 (0.12 to 0.13)
N1 or M1 or GS >3 + 4	12 035 (58.7)	1511	0.88 (0.86 to 0.9)	0.13 (0.12 to 0.13)
M1 or GS >3 + 4	12 028 (58.7)	1511	0.88 (0.86 to 0.9)	0.13 (0.12 to 0.13)
≥T3 or N1 or M1 or GS >4 + 3	7127 (34.8)	1289	0.75 (0.73 to 0.77)	0.18 (0.17 to 0.19)
T4 or N1 or M1 or GS >4 + 3	7028 (34.3)	1283	0.75 (0.73 to 0.77)	0.18 (0.17 to 0.19)
N1 or M1 or GS >4 + 3	7006 (34.2)	1278	0.74 (0.72 to 0.77)	0.18 (0.17 to 0.19)
M1 or GS >4 + 3	6977 (34.0)	1277	0.74 (0.72 to 0.76)	0.18 (0.17 to 0.19)
≥T3 or N1 or M1 or GS >8	4748 (23.2)	1108	0.65 (0.62 to 0.67)	0.23 (0.22 to 0.25)
T4 or N1 or M1 or GS >8	4594 (22.4)	1094	0.64 (0.61 to 0.66)	0.24 (0.23 to 0.25)
N1 or M1 or GS >8	4567 (22.3)	1088	0.63 (0.61 to 0.66)	0.24 (0.23 to 0.25)
M1 or GS >8	4519 (22.0)	1082	0.63 (0.61 to 0.65)	0.24 (0.23 to 0.25)

[a]CI = confidence interval; GS = Gleason score; PPV = positive predictive value.

research because the necessary data inputs are not collected consistently within or across population-based studies. Our proposed outcome definition, intended for use in population-based etiologic research and not clinical practice, uses clinical data fields that are commonly ascertained and, as a result, should be widely applicable to both long-standing and new epidemiological studies.

The definition we are proposing is based on a methodical analysis of US SEER data from a year in which PSA testing to screen for prostate cancer was common. Factors that vary geographically and temporally, such as the background rate of PSA testing and the availability and guidelines for cancer treatment, may influence the performance of each definition for discriminating fatal prostate cancer within a 10-year window. Whether our definition performs similarly in regions of the world without widespread PSA testing and with different treatment regimens will need to be examined. Our definition may also perform differently for cases diagnosed before changes to the Gleason grading system in 2005, though the relative rankings of the definitions should remain the same.

We acknowledge that the preferred outcome for a study may depend on the specific question of interest as well as the study population. For example, some studies may want to focus on more moderately aggressive prostate cancers, in which case category T3 or Gleason score 7 cancers may also be grouped as aggressive. Other studies may be interested in risk factors specifically for advanced-stage or high-grade prostate cancer, which may represent etiologically distinct facets of aggressive disease (9). To facilitate meta-analyses of these outcomes as well and to test for etiologic heterogeneity in prostate cancer case groups defined by these clinical features, we recommend that studies additionally report separate results for advanced-

stage prostate cancer (category T4 or N1 or M1) and high-grade prostate cancer (Gleason score ≥8 or grade group ≥4). Existing and future cohorts should strive to collect complete data on prostate cancer stage and grade, as well as on primary and secondary Gleason pattern and diagnostic PSA value, to enable etiologic studies to tease apart risk factor associations by these clinical features and allow future proposed definitions that build on this work to have greater flexibility. In studies where complete data are unavailable, multiple imputation can be used to impute missing stage and grade and retain all cases in analyses; as in our SEER analysis (Supplementary Table 6, available online), when cases with missing data are more likely to have aggressive disease features, retaining these cases improves the performance of definitions of aggressive disease and can increase study power and precision in estimation.

We also recognize the limitations of our analytic approach. Though we benchmarked our definitions against the outcome of fatal prostate cancer within 10 years of diagnosis, this outcome may be considered an imperfect gold standard for aggressive prostate cancer. Some prostate cancer cases undoubtedly died from prostate cancer beyond the 10-year window, whereas others with disease aggressive enough to have caused death from prostate cancer within 10 years died instead from competing causes of death. In our analysis, 30.9% of cases meeting the recommended definition for aggressive prostate cancer died of other causes within 10 years of diagnosis; though these cases were considered false positives for fatal prostate cancer, they still may have had aggressive disease. Our selection criteria for choosing the final definition was, by nature, subjective, but we used a data-driven approach to inform our selection and sought to reach a consensus among a globally representative group of prostate cancer epidemiologists. In this study, we used performance metrics—specifically sensitivity and PPV—to assess, compare, and prioritize definitions within a population of men with prostate cancer to enrich for aggressive disease. We did not examine performance of the definitions, bias in estimation of effect, or optimal sample size with respect to etiologic studies conducted in populations of men at risk for prostate cancer, which are related but perhaps unattainable goals given the many assumptions and scenarios that would need to be considered.

We nonetheless recommend that all etiologic epidemiologic studies of prostate cancer report results for the outcome of aggressive prostate cancer defined as T4 or N1 or M1 or Gleason score greater than or equal to 8 (or grade group ≥4). To enable meta-analyses of other prostate cancer endpoints, we also recommend that etiologic epidemiologic studies report results for advanced stage prostate cancer (T4 or N1 or M1), high-grade prostate cancer (Gleason score ≥8 or grade group ≥4), and fatal prostate cancer. Even if not the primary study outcomes, results for these endpoints can be published as supplementary material to improve comparability and facilitate future meta-analyses. Ultimately, we hope that reporting associations for a common set of endpoints will provide standardization across the field, thereby accelerating advances in prostate cancer etiology and prevention.

Funding

This research was supported by the Intramural Research Program of the National Cancer Institute, National Institutes of Health, Department of Health and Human Services. Investigators were also supported by the following grants: U01-CA164975, U01-CA182883, U01-CA167552, P30-CA006973, P30-CA006516, P30-CA008748, U01-CA063673, UM1-CA167462, U01-CA167462, K07-CA230182 (NCI/NIH/HHS), W81XWH-18–1-0330 (DOD), C8221/A19170, C8221/A29017 (Cancer Research UK), a University of Hawaii Cancer Center Seed Grant, Prostate Cancer Foundation Young Investigator Awards, and Program Grants 209057 and 1074383 and an Enabling Grant 396414 from the Australian National Health and Medical Research Council.

Notes

Role of the funder: The funding source had no role in the design and conduct of the study; the collection, analysis, and interpretation of the data; the preparation of the manuscript; or the decision to submit the manuscript for publication.

Disclosures: The authors have no conflicts of interest to disclose.

Author contributions: Conceptualization: LMH, EJJ, LAM, EAP, MBC; Formal analysis: LMH; Investigation: LMH, IA, DA, KHB, SIB, QC, CC, IC, JMG, GGG, JH, CEJ, TJK, SK, SK, HL, SXL, RJM, SCM, KLP, APC, TER, SAS, MJS, KHS, CMT, RCT, SJW, LW, EJJ, LAM, EAP, MBC; Methodology: LMH, IA, DA, KHB, SIB, QC, CC, IC, JMG, GGG, JH, CEJ, TJK, SK, SK, HL, SXL, RJM, SCM, KLP, APC, TER, SAS, MJS, KHS, CMT, RCT, SJW, LW, EJJ, LAM, EAP, MBC; Visualization: LMH; Supervision: EJJ, LAM, EAP, MBC; Writing – original draft: LMH; Writing – review & editing: IA, DA, KHB, SIB, QC, CC, IC, JMG, GGG, JH, CEJ, TJK, SK, SK, HL, SXL, RJM, SCM, KLP, APC, TER, SAS, MJS, KHS, CMT, RCT, SJW, LW, EJJ, LAM, EAP, MBC.

Data Availability

The data that support the findings of this study are openly available from the SEER database with a signed SEER Research Data Use Agreement (https://seer.cancer.gov/data/access.html).

References

1. Torre LA, Siegel RL, Ward EM, et al. Global cancer incidence and mortality rates and trends–an update. *Cancer Epidemiol Biomarkers Prev*. 2016;25(1):16–27.
2. American Cancer Society. *Cancer Facts and Figures 2020*. Atlanta, GA: American Cancer Society; 2020.
3. Brawley OW. Prostate cancer epidemiology in the United States. *World J Urol*. 2012;30(2):195–200.
4. Pernar CH, Ebot EM, Wilson KM, et al. The epidemiology of prostate cancer. *Cold Spring Harb Perspect Med*. 2018;8(12):a030361.
5. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst*. 2009;101(6):374–383.
6. Albertsen PC. Observational studies and the natural history of screen-detected prostate cancer. *Curr Opin Urol*. 2015;25(3):232–237.
7. Platz EA, De Marzo AM, Giovannucci E. Prostate cancer association studies: pitfalls and solutions to cancer misclassification in the PSA era. *J Cell Biochem*. 2004;91(3):553–571.
8. Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986-2005. *J Natl Cancer Inst*. 2009;101(19):1325–1329.
9. Giovannucci E, Liu Y, Platz EA, et al. Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *Int J Cancer*. 2007;121(7):1571–1578.
10. Jahn JL, Giovannucci EL, Stampfer MJ. The high prevalence of undiagnosed prostate cancer at autopsy: implications for epidemiology and treatment of prostate cancer in the prostate-specific antigen-era. *Int J Cancer*. 2015;137(12):2795–2802.

11. World Cancer Research Fund International Continuous Update Project. Diet, nutrition, physical activity and prostate cancer. 2014. https://www.wcrf.org/sites/default/files/Prostate-Cancer-2014-Report.pdf. Accessed May 27, 2019.
12. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. *The Health Consequences of Smoking - 50 Years of Progress: A Report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention; 2014.
13. 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.
14. Kazmi N, Haycock P, Tsilidis K, et al. Appraising causal relationships of dietary, nutritional and physical-activity exposures with overall and aggressive prostate cancer: two-sample Mendelian-randomization study based on 79 148 prostate-cancer cases and 61 106 controls. *Int J Epidemiol*. 2020;49(2):587–596. [Inserted from online]
15. Key TJ, Appleby PN, Travis RC, et al. Endogenous hormones nutritional biomarkers. Carotenoids, retinol, tocopherols, and prostate cancer risk: pooled analysis of 15 studies. *Am J Clin Nutr*. 2015;102(5):1142–1157.
16. Seibert TM, Fan CC, Wang Y, et al. Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. *BMJ*. 2018;360:j5757.
17. Surveillance, Epidemiology, and End Results (SEER) Program. SEERStat Database: Incidence—SEER Research Data, 18 Registries, Nov 2019 Sub (2000-2017)—Linked To County Attributes - Time Dependent (1990-2017) Income/Rurality, 1969-2017 Counties, National Cancer Institute, DCCPS, Surveillance Research Program. www.seer.cancer.gov. Published April 2020, based on the November 2019 submission. Accessed April 27, 2020.
18. NCI Cohort Consortium. Membership of the NCI Cohort Consortium. https://epi.grants.cancer.gov/cohort-consortium/members/. Accessed May 20, 2020.
19. Kelly SP, Rosenberg PS, Anderson WF, et al. Trends in the incidence of fatal prostate cancer in the United States by race. *Eur Urol*. 2017;71(2):195–201.
20. Zelic R, Garmo H, Zugna D, et al. Predicting prostate cancer death with different pretreatment risk stratification tools: a head-to-head comparison in a nationwide cohort study. *Eur Urol*. 2020;77(2):180–188.
21. Cucchiara V, Cooperberg MR, Dall'Era M, et al. Genomic markers in prostate cancer decision making. *Eur Urol*. 2018;73(4):572–582.
22. Xie W, Regan MM, Buyse M, et al.; on behalf of the ICECaP Working Group. Metastasis-free survival is a strong surrogate of overall survival in localized prostate cancer. *J Clin Oncol*. 2017;35(27):3097–3104.