STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	**Tumor subtype defines distinct pathways of molecular and clinical progression in primary prostate cancer**	1
		Abstract: We explore the molecular and clinical progression of different genomic subtypes of PCa using distinct tumor lineage models based on human genomic and transcriptomic data. We develop novel transcriptional classifiers, and define "early" and "late" categories of molecular subclasses from 8,158 PCa patients. Molecular subclasses are correlated with clinical outcomes and pathologic characteristics using Kaplan-Meier and logistic regression analyses. We identify *PTEN* and *CHD1* alterations as subtype-specific late progression events specifically in *ERG*+ and *SPOP* mutant tumors respectively, and two distinct progression models consisting of *ERG*/*PTEN* and *SPOP*/*CHD1* with shared early tumorigenesis but distinct pathways towards progression. We find that within *ERG*+ and *SPOP* mutant subtypes, late events are associated with worse prognosis. Importantly, the clinical and pathologic features associated with distinct late events at radical prostatectomy are strikingly different, *PTEN* deletions are associated with increased locoregional stage, while *CHD1* deletions are only associated with increased grade, despite equivalent metastatic potential. These findings suggest a paradigm in which specific subtypes of prostate cancer follow distinct pathways of progression, at both the molecular and clinical levels. Therefore, the interpretation of common clinical parameters such as locoregional tumor stage may be influenced by the underlying tumor lineage, and potentially influence management decisions.	3
Introduction			
Background/rationale	2	Prostate cancer (PCa) is a clinically and molecularly heterogeneous disease. Emerging next-generation DNA and RNA sequencing data point toward different molecular subclasses of prostate cancer, defined by underlying genomic alterations. However, how these early alterations influence subsequent molecular events and the course of the disease over its long natural history remains unclear. We previously established a framework using RNA-based model to classify tumor subtype from transcriptional data, allowing the interrogation of cohorts with the long follow-up necessary to define clinical outcomes	4
Objectives	3	To understand molecular and clinical progression in *ERG*+ and *SPOP* mutant subtypes of PCa.	4
Methods			
Study design	4	We established distinct tumor lineage models of PCa progression, by defining early and late progression events within specific subtypes, and investigating their unique and shared transcriptional alterations and signaling pathways. We developed transcriptional classifiers to categorize subtype-specific early and late states, and applied these to a retrospective cohort including 1,626 patient samples and a prospective cohort including 6,532 samples using microarray-based gene expression data from a clinically available prognostic assay.	4
Setting	5	NA	
Participants	6	NA	
Variables	7	Patient outcomes including biochemical recurrence (BCR), metastasis (MET) and prostate cancer specific mortality (PCSM), and clinical variables, including age, race, preoperative PSA, Gleason score, lymph node invasion (LNI), surgical margin status (SMS), extracapsular extension (ECE), and seminal vesicle invasion (SVI)	16

Data sources/ measurement	8*	NA	
Bias	9	NA	
Study size	10	Retrospective (n=1,626) and prospective (n=6,532) cohorts were derived from the Decipher GRID registry (NCT02609269).	12
Quantitative variables	11	NA	
Statistical methods	12	Kaplan-Meier and logistic regression analyses Logistic regression analyses NA NA NA	16
Results			
Participants	13*	Retrospective (n=1,626) and prospective (n=6,532) cohorts were derived from the Decipher GRID registry (NCT02609269). NA NA	12
Descriptive data	14*	NA NA NA	
Outcome data	15*	Retrospective (n=1,626) and prospective (n=6,532) cohorts	12

Main results	16	We found worse metastasis (MET) free survival in both *CHD1*^{del} and *PTEN*^{del} tumors compared to the early state within each subtype (*SPOP*^{mut} and *PTEN*^{wt}). Of note, "early" states of each subtype had similar favorable prognosis, while both "late" states showed similar unfavorable prognosis. Endpoints of biochemical recurrence (BCR) free survival and prostate cancer specific mortality (PCSM) free survival rates followed similar patterns, consistent with previous findings. We found tumors with predicted *PTEN* deletion were more likely to harbor adverse pathological features at radical prostatectomy: lymph node invasion, extracapsular extension, seminal vesicle invasion, and higher Gleason score in both retrospective and prospective cohorts, consistent with pathologic features of late progression events. Strikingly, however, tumors with predicted *CHD1* deletion were only associated with higher Gleason score but no other adverse clinical features. NA NA	9
Other analyses	17	We found specific subtypes of PCa are associated with subsequent molecular changes; tumors with *ERG* fusions later may acquire *PTEN* deletions, while *SPOP* mutant tumors may progress with *CHD1* deletion. By comparing the transcriptional pathways between these two tumor lineages, we identified similar enriched functions from the "normal" to "early" states, but divergent signatures from the "early" to "late" states, in multiple localized prostate cancer cohorts. These analyses credential two distinct transcription-based tumor lineage progression models consisting of *ERG*/*PTEN* and *SPOP*/*CHD1*, with shared early tumorigenesis but distinct pathways towards progression.	5
Discussion			
Key results	18	We identify *PTEN* and *CHD1* alterations as subtype-specific late progression events specifically in *ERG*+ and *SPOP* mutant tumors respectively, and two distinct progression models consisting of *ERG*/*PTEN* and *SPOP*/*CHD1* with shared early tumorigenesis but distinct pathways towards progression. We find that within *ERG*+ and *SPOP* mutant subtypes, late events are associated with worse prognosis. Importantly, the clinical and pathologic features associated with distinct late events at radical prostatectomy are strikingly different, *PTEN* deletions are associated with increased locoregional stage, while *CHD1* deletions are only associated with increased grade, despite equivalent metastatic potential.	10
Limitations	19	Current genomic and clinical data are derived from bulk tumor sample and limited by intratumor heterogeneity. Molecular and clinical progression for distinct subtypes need to be further investigated at the single cell level.	12
Interpretation	20	In conclusion, we established mutually exclusive tumor lineage models of PCa progression: *ERG*/*PTEN* and *SPOP*/*CHD1*. Using transcriptional classifiers to categorize progressive events, we predict lineage and progression status from a large population of human patients, and find that molecularly defined late progression events are associated with worse clinical outcome, but may be associated with distinct clinical pathways toward metastasis. More broadly, these data suggest a paradigm in which specific subtypes of prostate cancer follow distinct molecular pathways of tumor progression, and the interpretation of common risk stratification parameters such as locoregional tumor staging may be influenced by the underlying tumor lineage and degree of molecular progression.	12
Generalisability	21	Whether tumor lineages and molecular subclasses will add clinical value to current risk stratification tools remains unclear, and need to be prospectively tested in future clinical studies. However, these data do provide compelling rationale to consider molecular subclass in future clinical trial designs.	11
Other information			
Funding	22	This work was supported by: NCI (P50CA211024 and R01CA233650, C.E.B.), a Urology Care Foundation Rising Star in Urology Research Award (C.E.B.), Damon Runyon Cancer	16

	Research Foundation MetLife Foundation Family Clinical Investigator Award (C.E.B.), Prostate Cancer Foundation (C.E.B., D.L., M.A.A.), Department of Defense Postdoctoral Fellowship, United States (W81XWH-17-1-0137, M.A.A.), and Foundation AIRC (IG 19221 F.D.).	

16

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.