EUROPEAN UROLOGY XXX (2017) XXX–XXX

available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Platinum Priority – Prostate Cancer
Editorial by XXX on pp. x–y of this issue

Validation of a Genomic Risk Classifier to Predict Prostate Cancer-specific Mortality in Men with Adverse Pathologic Features

R. Jeffrey Karnes[a], *Voleak Choeurng*[b], *Ashley E. Ross*[c], *Edward M. Schaeffer*[c,h], *Eric A. Klein*[d], *Stephen J. Freedland*[e,f], *Nicholas Erho*[b], *Kasra Yousefi*[b], *Mandeep Takhar*[b], *Elai Davicioni*[b], *Matthew R. Cooperberg*[g], *Bruce J. Trock*[c,*]

[a] *Department of Urology, Mayo Clinic College of Medicine, Rochester, MN, USA;* [b] *GenomeDx Biosciences, Vancouver, BC, Canada;* [c] *James Buchanan Brady Urological Institute and Department of Urology, Johns Hopkins School of Medicine, Baltimore, MD, USA;* [d] *Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland, OH, USA;* [e] *Surgery Section, Durham Veteran Affairs Medical Center, Durham, NC, USA;* [f] *Department of Surgery, Division of Urology, Cedars-Sinai Medical Center, Los Angeles, CA, USA;* [g] *Department of Urology and Epidemiology and Biostatistics, University of California San Francisco Comprehensive Cancer Center, San Francisco, CA, USA;* [h] *Department of Urology, Northwestern University School of Medicine, Chicago, IL, USA*

Article info

Article history:
Accepted March 24, 2017

Associate Editor:
Giacomo Novara

Keywords:
Radical prostatectomy
Adverse pathologic features
Prostate cancer-specific mortality
Genomic classifier
CAPRA-S

Abstract

Background: Risk of prostate cancer-specific mortality (PCSM) is highly variable for men with adverse pathologic features at radical prostatectomy (RP); a majority will die of other causes. Accurately stratifying PCSM risk can improve therapy decisions.
Objective: Validate the 22 gene Decipher genomic classifier (GC) to predict PCSM in men with adverse pathologic features after RP.
Design, setting, and participants: Men with adverse pathologic features: pT3, pN1, positive margins, or Gleason score >7 who underwent RP in 1987–2010 at Johns Hopkins, Cleveland Clinic, Mayo Clinic, and Durham Veteran's Affairs Hospital. We also analyzed subgroups at high risk (prostate-specific antigen >20 ng/ml, RP Gleason score 8–10, or stage $>$ pT3b), or very high risk of PCSM (biochemical recurrence in <2 yr [BCR2], or men who developed metastasis after RP [MET]).
Outcome measurements and statistical analysis: Logistic regression evaluated the association of GC with PCSM within 10 yr of RP (PCSM10), adjusted for the Cancer of the Prostate Risk Assessment Postsurgical Score (CAPRA-S). GC performance was evaluated with area under the receiver operating characteristic curve (AUC) and decision curves.
Results and limitations: Five hundred and sixty-one men (112 with PCSM10), median follow-up 13.0 yr (patients without PCSM10). For high GC score (>0.6) versus low-intermediate (≤ 0.6), the odds ratio for PCSM10 adjusted for CAPRA-S was 3.91 (95% confidence interval: 2.43–6.29), with AUC = 0.77, an increase of 0.04 compared with CAPRA-S. Subgroup odds ratios were 3.96, 3.06, and 1.95 for high risk, BCR2, or MET, respectively (all $p<0.05$), with AUCs 0.64–0.72. GC stratified cumulative PCSM10 incidence from 2.8% to 30%. Combined use of case-control and cohort data is a potential limitation.
Conclusions: In a large cohort with the longest follow-up to date, Decipher GC demonstrated clinically important prediction of PCSM at 10 yr, independent of CAPRA-S, in men with adverse pathologic features, BCR2, or MET after RP.
Patient summary: Decipher genomic classifier may improve treatment decision-making for men with adverse or high risk pathology after radical prostatectomy.

* Corresponding author. Department of Urology, Johns Hopkins School of Medicine, 600 North Wolfe Street, Park 211, Baltimore, MD 21287, USA. Tel. +1 410 614 9440; Fax: +1 410 502 7711.
E-mail address: btrock@jhmi.edu (B.J. Trock).

http://dx.doi.org/10.1016/j.eururo.2017.03.036
0302-2838/

1. Introduction

For men with adverse pathologic features at radical prostatectomy (RP), there is considerable variability in risk of subsequent recurrence, metastasis and death, and time intervals between these progression events can be long [1]. This variability complicates decisions for adjuvant, salvage, and metastatic treatment. In recent years, there has been a rapid increase in the number of treatments for advanced prostate cancer (PCa), with growing interest in sequencing these treatments earlier in the disease course [2,3]. Because these agents are not without side effects, and because early development of resistance could preclude more effective use at a later stage, it is important to target novel and aggressive treatment regimens to men at highest risk of PCa-specific mortality (PCSM).

Currently, surgical pathology features are the primary means to identify men at highest risk of PCSM [4,5]. Despite the ability of these features to stratify PCSM risk, there remains considerable variability in outcomes when applied to individual men with PCa. To improve prognostic accuracy and clinical decision-making, several risk classifiers have been developed based on biomarker signatures alone or integrated with clinical features. One of the signatures most extensively tested and validated for predicting risk of metastasis in men at intermediate and high risk is the Decipher genomic classifier (GC), comprised of 22 gene expression markers derived from whole transcriptomic analysis of formalin fixed paraffin embedded tissue [6–9]. The GC generates a score from 0 to 1, with higher values associated with worse outcomes. Recently the GC was also shown to predict risk of PCSM in men with high risk based on preoperative prostate-specific antigen (PSA) levels or pathology [10]. Because that study was based on a relatively small number of men from a single institution, and with few PCSM events and short follow-up, we undertook a more extensive validation of the ability of the GC to predict PCSM.

2. Materials and methods

2.1. Patient cohort

The study population comprised four cohorts of PCa patients with adverse pathologic features, defined as RP Gleason score $\geq$7, RP stage pT3 or pN1, or positive surgical margins; per the study protocol patients with neoadjuvant therapy were excluded. The cohorts included 407 men from Mayo Clinic who underwent RP from 1987 to 2006, 355 from Johns Hopkins treated from 1992 to 2010, 179 from Cleveland Clinic treated from 1988 to 2008, and 113 from the Durham Veteran's Administration Medical Center from 1991 to 2010, totaling 1054 with all required analysis variables, among whom there were 141 confirmed PCa deaths. Patients from the Mayo Clinic cohort did not include men used to train the original GC [6]. Patient follow-up after RP was not standardized among the four institutions, but differences were minor. The primary outcome was PCSM within 10 yr of RP (PCSM10); patients who died of PCa $>$ 10 yr after RP were considered censored (n = 29), and patients alive with less than 10 yr of follow-up (n = 493) were excluded. This resulted in a total of 561 patients with 112 PCSM10 (79% of all PCSM). Institutional review boards at the participating institutions approved the research protocol.

2.2. Specimen processing

Specimen selection and processing have been described previously [7,11–13]. Following microarray quality control using the Affymetrix Power Tools packages (Thermo Fisher Scientific Inc., MA, USA) [14], probeset summarization and normalization was performed utilizing the single channel array normalization algorithm [15]. Information about obtaining Decipher for routine clinical practice is in the Supplementary data.

2.3. Statistical analysis

The GC was calculated as a numeric value ranging from 0 to 1, based on each patient's individual expression of the 22 genes integrated in a previously trained and validated signature [6]. The Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) score was likewise calculated by applying PSA, RP Gleason score, and RP stage values to a previously validated algorithm, producing a score ranging from 0 to 12 [16]. Characteristics of patients who were censored after 10 yr or had died from PCa within 10 yr were compared with univariate chi-square or Wilcoxon rank sum tests for categorical and continuous variables, respectively. Because the study cohorts incorporated case-control and cohort study designs, survival analysis was not appropriate. Although there are methods for adapting proportional hazards models to case-control data [17], they are not suitable when the data combine both case-control and cohort data. Therefore, we conservatively used a logistic regression approach with PCSM10 as the outcome.

The ability of the GC to improve upon prognostic information in clinical variables was evaluated in two ways. Unconditional logistic regression models were fit to individual clinical variables (PSA, RP Gleason score, RP stage) to generate a base model, then the GC was added to the base model. Alternatively, the GC was added to a model with CAPRA-S as a validated measure of postoperative risk. The latter approach may be a more realistic indication of GC performance because the base clinical model is derived from this dataset, hence subject to overfitting [18], whereas the GC and CAPRA-S were both trained and validated on datasets independent of the current data. Models were fit for: (1) all men with adverse pathologic features and PCSM10 defined, (2) men considered high risk (PSA $>$ 20 ng/ml or RP Gleason score 8–10 or stage pT3b or pN1) [10], and (3) men at very high risk of death due to biochemical recurrence within 2 yr (BCR2) [1] or metastasis (MET). For the overall and subgroup analyses the time frame for PCSM10 began with the date of RP. The univariate and adjusted effect of GC were measured by the odds ratio (OR) and 95% confidence interval for a 0.1 increase in the score (GC ranges from 0 to 1), or for GC high ($>$ 0.6) versus low-intermediate ($\leq$ 0.6). The bootstrap corrected area under the receiver operating characteristic curve (AUC) was used to determine incremental improvement in model performance by adding the GC to CAPRA-S or to the base model. Decision curve analysis was used to compare the net benefit associated with PCSM10 prediction using the base model, GC or CAPRA-S alone, and GC combined with either the base model or CAPRA-S [19]. Although bootstrapping was used for internal validation of the models, performance of the models must be regarded as best case scenario until externally validated [20]. Because the analysis combines data from four institutions, analyses that stratified by institution were also performed. These models gave nearly identical results to the unstratified models so only the latter are reported. Models were fit using R v3.2.0 (R Foundation for Statistical Computing, Vienna, Austria) and SAS v9.4 (SAS Institute, Cary, NC, USA).

3. Results

The analytic cohort consisted of 561 patients either censored alive after 10 yr of follow-up or died of PCa

Please cite this article in press as: Karnes RJ, et al. Validation of a Genomic Risk Classifier to Predict Prostate Cancer-specific Mortality in Men with Adverse Pathologic Features. Eur Urol (2017), http://dx.doi.org/10.1016/j.eururo.2017.03.036