

Clinical risk-stratification for prostate cancer: Where are we, and where do we need to go?

Matthew R. Cooperberg, MD, MPH

Departments of Urology and Epidemiology & Biostatistics, UCSF Helen Diller Family Comprehensive Cancer Centre, University of California, San Francisco, CA, United States

Cite as: *Can Urol Assoc J* 2017;11(3-4):101-2. <http://dx.doi.org/10.5489/cuaj.4520>

See related article on page 94.

Risk-stratification for prostate cancer, both at the time of diagnosis and at subsequent decision points, is absolutely critical to guide appropriate treatment decision-making. The venerable risk classification published by D'Amico et al in 1998,¹ and closely related classifications adopted by the American Urological Association (AUA), the National Comprehensive Cancer Network (NCCN), and the Genitourinary Radiation Oncologists of Canada (GUROC), are no longer sufficient for this task.

These classification systems do not distinguish between Gleason 3 + 4 or 4 + 3, overweigh clinical T stage,² do not account for extent of biopsy involvement, and most importantly do not account for multiple adverse risk parameters. For example, a man with a Gleason 4 + 3, prostate-specific antigen (PSA) 19 tumour in eight of 12 cores is grouped in the same intermediate risk category as one with a Gleason 3 + 4, PSA 4 tumour in one of 12 cores.

The Prostate Cancer Risk Stratification (ProCaRS) risk groups offer improved granularity within the D'Amico/NCCN intermediate- and high-risk groups, but still reflect an ad hoc, rather than mathematically derived model, and do not perform well as linear predictors.

Risk models derived from multivariable Cox regression or other survival analysis—which can be represented using nomograms, lookup tables, or other scoring systems—offer much more granular risk-stratification. Since the publication of the original preoperative nomogram by Kattan et al³ a few months before the D'Amico classification, hundreds of risk models have been published, with many expressed as nomograms. While nomograms offer superior accuracy over risk-groupings, they can be cumbersome to calculate, requiring either a multistep paper tool or computer software. The regression formula underlying each nomogram is rarely published, complicating efforts to validate them or use them in subsequent research. As a result, most are never validated.

When validation studies are performed, they frequently find that in a new population, a given nomogram may offer discrimination, as indicated by the c-index, comparable to the original study, but that calibration often suffers, with the nomogram consistently over or underestimating the risk of the endpoint in the validation population.

The Cancer of the Prostate Risk Assessment (CAPRA) score was developed to combine the superior accuracy of a multivariable model with the simplicity of a scoring system requiring no software or calculation beyond rudimentary arithmetic. It has been validated in over a dozen independent studies,⁴ including five radiation therapy cohorts (two of which were other Canadian cohorts^{5,6}). CAPRA is the only score validated to predict both biochemical and clinical endpoints, including cancer-specific mortality following both surgical and radiation modalities, with a c-index of 0.80.⁷

In the study by Tiberi et al, the CAPRA score was compared to the NCCN, GUROC, and ProCaRS risk groups. CAPRA and the ProCaRS risk groups generally outperformed NCCN or GUROC, and CAPRA outperformed all the other scores in some analyses—although all the differences were fairly subtle in these analyses. The relatively low c-index calculations in this study reflect, in part, the somewhat restricted risk range in both cohorts. A substantial majority had lower-risk disease (CAPRA ≤ 3), and few had high-risk disease. Reflecting this distribution, the event rate was low in both cohorts, further limiting analysis of risk model performance. Finally, followup was quite short in the one of the databases assessed (Centre Hospitalier de l'Université de Montréal [CHUM]), with only 3.7 years followup and outcomes by the Phoenix definition only interpretable at 1.7 years.⁸

Multiple prior studies have also shown CAPRA to outperform the AUA and NCCN risk groups.⁴ The ProCaRS risk groups appear to perform nearly as well as CAPRA in the ProCaRS cohorts, but they are not as easy to calculate and do not offer any other obvious advantages. These risk groups also do not substratify the D'Amico/AUA low-risk group. The ProCaRS nomograms validated in this study, although not compared to the other risk systems, nor to existing radia-

tion nomograms,⁹ showed good discrimination, but poor calibration. The nomograms weigh PSA quite heavily and do not differentiate Gleason 3 + 4 vs. 4 + 3 disease. They are also as difficult to calculate for large numbers of men, as are many other nomograms.

What then, should define the current state-of-the-art for risk-stratification? It should be increasingly clear that the AUA/NCCN/GUROC risk groups do not. Many nomograms are available, some of which have been well-validated, but few of which are actually used in clinical practice or contemporary research. The CAPRA score is now over 10 years old, has been extensively validated, and is one of the only instruments relevant for men who have not yet made a treatment decision. As Tiberi et al note, CAPRA has also been used as a baseline “gold standard” clinical risk-stratification to which emerging genomic tests have been added to demonstrate their independent prognostic value. This point is one worth stressing: any putative novel biomarker or imaging test must be proved to offer improvement over a true multivariable model. Simply substratifying the standard AUA or NCCN risk groups is insufficient; this can be done with no cost or additional testing using the CAPRA score or any number of other multivariable tools.

Given how many options are already available,¹⁰ development and promulgation of additional risk-stratifications based on standard clinical variables at this point is somewhat hard to justify. What Tiberi et al have not clarified, in short, is: even if the ProCaRS risk groups perform nearly as well as CAPRA, why should they be preferred? In 2017, we do not necessarily need more risk assessments; what we do need is a small number of truly multivariable tools to be adopted consistently in both clinical practice and research, across multiple treatment modalities, so that we can both improve decision-making for men diagnosed with prostate cancer and accelerate research progress toward truly personalized management.

Competing interests: The author reports no competing personal or financial interests.

References

1. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998;280:969-74. <https://doi.org/10.1001/jama.280.11.969>
2. Reese AC, Cooperberg MR, Carroll PR. Minimal impact of clinical stage on prostate cancer prognosis among contemporary patients with clinically localized disease. *J Urol* 2010;184:114-9. <https://doi.org/10.1016/j.juro.2010.03.025>
3. Kattan MW, Eastham JA, Stapleton AM, et al. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst* 1998;90:766-71. <https://doi.org/10.1093/jnci/90.10.766>
4. Brajtburg JS, Leapman MS, Cooperberg MR. The CAPRA score at 10 years: Contemporary perspectives and analysis of supporting studies. *Eur Urol* September 2016. [Epub ahead of print]. <https://doi.org/10.1016/j.eururo.2016.08.065>
5. Delouya G, Krishnan V, Bahary J-P, et al. Analysis of the Cancer of the Prostate Risk Assessment (CAPRA) to predict for biochemical failure after external beam radiotherapy or prostate seed brachytherapy. *Urology* 2014;84:629-33. <https://doi.org/10.1016/j.urolgy.2014.05.032>
6. Krishnan V, Delouya G, Bahary J-P, et al. The Cancer of the Prostate Risk Assessment (CAPRA) score predicts biochemical recurrence in intermediate-risk prostate cancer treated with external beam radiotherapy (EBRT) dose escalation or low-dose rate (LDR) brachytherapy. *BJU Int* 2014;114:865-71. <https://doi.org/10.1111/bju.12587>
7. Cooperberg MR, Hinotsu S, Namiki M, et al. Risk assessment among prostate cancer patients receiving primary androgen-deprivation therapy. *J Clin Oncol* 2009;27:4306-13. <https://doi.org/10.1200/JCO.2008.21.5228>
8. Roach M, Hanks G, Thames H, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: Recommendations of the RTOG-ASTRO Phoenix Consensus Conference. *Int J Radiat Oncol Biol Phys* 2006;65:965-74. <https://doi.org/10.1016/j.ijrobp.2006.04.029>
9. Kattan MW, Zelefsky MJ, Kupelian PA, et al. Pretreatment nomogram that predicts five-year probability of metastasis following three-dimensional conformal radiation therapy for localized prostate cancer. *J Clin Oncol* 2003;21:4568-71. <https://doi.org/10.1200/JCO.2003.05.046>
10. Shariat SF, Karakiewicz PI, Roehrborn CG, et al. An updated catalog of prostate cancer predictive tools. *Cancer* 2008;113:3075-99. <https://doi.org/10.1002/cncr.23908>

Correspondence: Dr. Matthew R. Cooperberg, Departments of Urology and Epidemiology & Biostatistics UCSF Helen Diller Family Comprehensive Cancer Centre, University of California, San Francisco, CA, United States; mcooperberg@urology.ucsf.edu