Clinical validity and utility of genetic risk scores in prostate cancer

Brian T Helfand, James Kearns, Carly Conran, Jianfeng Xu

Current issues related to prostate cancer (PCa) clinical care (e.g., over-screening, over-diagnosis, and over-treatment of nonaggressive PCa) call for risk assessment tools that can be combined with family history (FH) to stratify disease risk among men in the general population. Since 2007, genome-wide association studies (GWASs) have identified more than 100 SNPs associated with PCa susceptibility. In this review, we discuss (1) the validity of these PCa risk-associated SNPs, individually and collectively; (2) the various methods used for measuring the cumulative effect of multiple SNPs, including genetic risk score (GRS); (3) the adequate number of SNPs needed for risk assessment; (4) reclassification of risk based on evolving numbers of SNPs used to calculate genetic risk, (5) risk assessment for men from various racial groups, and (6) the clinical utility of genetic risk assessment. In conclusion, data available to date support the clinical validity of PCa risk-associated SNPs and GRS in risk assessment among men with or without FH. PCa risk-associated SNPs are not intended for diagnostic use; rather, they should be used the same way as FH. Combining GRS and FH can significantly improve the performance of risk assessment. Improved risk assessment may have important clinical utility in targeted PCa testing. However, clinical trials are urgently needed to evaluate this clinical utility as well as the acceptance of GRS by patients and physicians.

Asian Journal of Andrology (2016) 18, 509–514; doi: 10.4103/1008-682X.182981; published online: 14 June 2016

Keywords: genetic risk score; prostate cancer; pyramid model; single nucleotide polymorphisms

INTRODUCTION

Together with advancements in treatments for both localized and advanced disease, there is strong evidence that serum prostate-specific antigen (PSA) screening has played a significant role in decreasing prostate cancer (PCa) mortality over the last several decades. However, widespread implementation of PSA screening has also contributed to a stage- and grade-migration, and to the detection and subsequent potential over-treatment of low-risk, indolent disease. There has been great controversy regarding the interpretation of USA- and European-based screening trials that were aimed at evaluating the impact of PSA screening on PCa mortality. While it is likely that PSA screening has saved thousands of men’s lives, it does so at the expense of high rates of potential over-treatment of indolent tumors that may never go on to harm or cause mortality for an individual. The question, as well as the likely solution to this major dilemma, should not be whether to screen, but how to apply PCa screening more intelligently and judiciously. What is needed to overcome this issue of over-screening is a way to identify those most at risk of developing PCa and aggressive disease. Currently, most authoritative groups, including the American Cancer Society (ACS), American Urologic Association (AUA) and National Comprehensive Network (NCCN), recommend early and selective PCa screening in men with a positive family history (FH) of PCa.

At present, FH of PCa is the most commonly used method to determine whether a man is at increased risk of developing the disease. It has been reported that men with an FH of PCa are at approximately 1.5- to 2.5-fold heightened risk for PCa if they have an affected relative. However, <10% of men in the general population are estimated to have a FH of the disease. Furthermore, the clinical use of FH to guide PCa risk assessment is problematic, as most men do not know their complete FH and FH status is unstable over time. For men who are confident in their knowledge of their family’s PCa history, FH assigns the same risk to all family members with the same degree of relation to an affected family member; despite the fact that it has been shown that these family members do not, in fact, have the same underlying genetic susceptibility.

In addition to FH to FH for PCa risk assessment, novel, more personalized methods are called for in order to determine which men in the general population harbor increased disease risk.

Recent, unprecedented genomic findings offer a promising solution to this current inability to stratify PCa risk among men. Specifically, many genome-wide association studies (GWASs) that have compared the germline genotypes of men with PCa to controls without known disease have identified single nucleotide polymorphisms (SNPs) that are associated with increased susceptibility to PCa (referred to as PCa risk-associated SNPs). There are now more than 100 PCa risk-associated SNPs that have been described, and it has been estimated that these genetic loci increase the estimated proportion of the familial risk to 33%.

SNPS AND CUMULATIVE EFFECT

The initial PCa risk-associated SNPs that were identified in 2007 were associated with a relatively high ability to estimate an individual's
Genetic risk scores of prostate cancer
BT Helfand et al

susceptibility to PCa with individual odds ratios (ORs) ranging between 1.20 and 1.79. Several studies have been published evaluating the combined predictive power of these SNPs into various genetic scores. In 2008, Zheng and colleagues published a study of the combined effect of five SNPs, each in different genetic regions. They described that the five SNPs plus FH accounted for 46% of the PCa cases in the Swedish men in their study, with an OR of 9.46 for developing PCa in men with a FH and all five SNPs. Cumulatively, these five SNPs could increase a man's risk by 4–5 fold. This significance was later confirmed in a US population. By 2009, 28 additional PCa risk-associated SNPs were identified by subsequent GWASs, with individual estimated ORs ranging from 1.05 to 1.35. Other studies evaluated a cumulative performance of these 33 SNPs (5 original + 28 novel SNPs) in discriminating prostate biopsy outcomes, and found their performance, measured by the area under the receiver operating characteristic curve (AUC), to be significantly higher than family history (0.59 and 0.52, respectively). Later, a total of 76 PCa risk-associated SNPs were cataloged by 2013. Most recently, 23 additional PCa risk-associated SNPs were discovered by evaluating more than 10 million SNPs across the genome among 43,303 PCa cases and 43,737 controls from the international PRACTICAL Consortium.

**METHODOLOGIES FOR ASSESSMENT OF GENETIC RISK**

As mentioned above, there are many potential flaws with the use of FH data to assess PCa risk. In fact, several studies have demonstrated that FH information is a relatively poor indicator of PCa risk as measured by area under the receiver operating curve (AUC = 0.52). PCa risk-associated SNPs are currently the best measure of the genetic component of disease susceptibility in the general population. It has been demonstrated that, together, they may be able to provide useful information to better identify the men who are at risk of developing the disease.

Several methodologies that incorporate multiple PCa risk variants have been developed as a clinical tool to more accurately assess PCa risk including: (1) a simple count of risk-associated alleles (referred to as risk allele count, RAC); (2) weighted risk allele count (wRAC, also referred to as polygenic risk score, or PRS), in which risk alleles are weighted by odds ratios (ORs) and (3) a population-standardized genetic risk score (GRS), in which alleles are weighted by the OR and frequency in the population. A recent study compared the performances of the three different methodologies using a large, prospective dataset from the Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trial. To assess whether the cumulative effect measured by each method was independently associated with PCa and with high-grade disease, multivariate analyses adjusting for age, family history, and other clinical variables were performed when testing the association of each method with disease. The authors used AUC to discriminate PCa and high-grade disease and compared the values with the AUC of FH. The AUC for discriminating PCa using FH alone was 0.53 (95% CI: 0.49–0.56) and for discriminating high-grade cancer was 0.54 (95% CI: 0.48–0.59). In comparison, the AUCs for discriminating PCa were 0.60 (95% CI: 0.57–0.63), 0.62 (95% CI: 0.59–0.64), and 0.62 (95% CI: 0.59–0.64) for the RAC, PRS and GRS methods, respectively. Similarly, the AUCs for discriminating high-grade tumors were 0.57 (95% CI: 0.51–0.63), 0.60 (95% CI: 0.54–0.65), and 0.60 (95% CI: 0.54–0.65) for the RAC, PRS, and GRS methods, respectively. No statistically significant difference was found between the three SNP-based methodologies, although the AUC estimates of GRS and PRS were higher than RAC. All three methodologies performed better than family history.

Based on these findings, the authors concluded that all three methodologies successfully capture the cumulative effect of PCa risk SNPs and are sound predictors of PCa risk. However, there are potential benefits to GRS over the other methodologies. GRS is the only method that assesses PCa risk using both the odds ratio and allele frequency of each SNP. As such, the GRS incorporates the effect size of each SNP, and, more importantly, compares the total to the average population risk. A GRS value equal to 1.0 represents the average population risk, a value <1.0 represents lower risk and a value >1.0 represents increased PCa risk. This unique property of the GRS calculation is not affected by the number of PCa risk SNPs that are incorporated. In comparison, RAC and PRS will change depending on the number of risk-SNPs that are included in the calculation, which makes it difficult to interpret the results for an individual patient.

**OPTIMAL NUMBER OF PROSTATE CANCER RISK SNPS TO INCORPORATE**

As mentioned above, the first PCa risk-associated SNPs identified were associated with a relatively high ability to estimate an individual’s disease susceptibility with individual estimated odd ratios ranging from 1.20 to 1.79. With increasing sample size and better genotyping and sequencing capabilities, more PCa risk-associated SNPs are expected to be identified. For example, 23 novel PCa risk-associated SNPs were discovered after evaluating 211,155 SNPs across the genome among 25,074 PCa cases and 24,272 controls from the International PRACTICAL Consortium. The effect size of these SNPs on PCa risk was relatively small compared to most prior SNPs, with ORs in the range of 1.06–1.15. This observation could be expected, as stronger PCa risk-associated SNPs were more likely to have been detected in prior studies with smaller sample sizes. However, they still contribute to the risk.

To determine whether all PCa risk-associated SNPs are required to calculate a GRS, a genetic study of 667 patients of Chinese ancestry was performed that compared the performance of GRS values that included a variable number of PCa risk-associated SNPs. Based on their results, the authors suggested that there was little value to adding additional newly discovered SNPs into a model to predict the outcome, a plateau effect. However, a recent study raised concerns about the possibility of risk reclassification when different numbers of SNPs were utilized in the calculation of a GRS. To address this concern, we calculated GRS for men enrolled in the REDUCE trial using various numbers of PCa risk-associated SNPs (unpublished data). There were four sequential sets of PCa risk-associated SNPs (17, 34, 51, and 68 SNPs) that were evaluated and used to calculate GRS values. GRS values calculated using each of the four SNP sets were significantly associated with the risk for PCa, and all had a better performance than family history. Although there was variability in GRS values from these four sequential sets of risk-associated SNPs, they were highly correlated. Using a cutoff value to define higher PCa risk (e.g., GRS ≥1.5), reclassification of risk categories using increasing numbers of SNPs was observed. However, multiple GRS values from evolving SNP sets is not a limitation for GRS; rather, having multiple GRS values for each individual actually provides a valuable tool for refining risk for all subjects. Risk reclassification effectively captures men with GRS values in a gray zone (near 1.5) that are at intermediate risk, and men who have consistently lower (<1.5) or higher (≥1.5) GRS values from multiple SNP sets are further assured of their low or high genetic risk, respectively. Based on these results, it appears that (1) GRS values from all currently available risk-associated SNPs should be used to stratify PCa risk and (2) newly discovered SNPs should be used to calculate.
new GRS values, and a combination of new and previous GRS values should be considered together to further refine risk.

**CLINICAL VALIDITY AND UTILITY OF GENETIC RISK ASSESSMENT**

For genetic risk assessment to be useful, it has to meet several criteria. For example, genetic risk assessment must provide additional information that cannot be easily captured using currently available tools, and it must be applicable to all men. In addition, it has to out-perform the current standard of PCa risk assessment including advanced age, family history, and race.

Many research studies have consistently supported the use of risk-associated SNPs to more precisely estimate PCa risk among men in the general population. For example, in the largest PCa genomics study in Caucasian men conducted in 2014, 100 PCa risk-associated SNPs were genotyped in 43,303 prostate cancer cases, and 43,737 controls, and a PRS was calculated for each patient. Results from this study showed that the top 10% of men in the highest risk stratum have an approximately 3-fold increased risk of developing PCa compared to the population average, and the top 1% of men have an approximately 6-fold increase in relative risk in comparison with the population average. In a similar study, the Prostate, Lung, Colorectal, Ovarian (PLCO) trial, 33 risk-associated SNPs were genotyped in 1017 cases affected by prostate cancer and 1277 controls. Results from this study showed that for GRS score quartiles, PCa detection rates significantly increased by quartile: 43.2%, 47.8%, 58.8%, and 69.4% in the first, second, third, and fourth quartiles, respectively ($P < 0.001$). Taken together, these studies support that SNP-based risk assessment provides additional risk stratification that is not currently captured by available clinical information. This is an important concept in an era of personalized medicine because it provides an exceptionally individualized quantitative risk assessment that can guide both physicians and patients in clinical care.

Due to the availability of samples for analyses, the vast majority of genetic studies in PCa conducted to date have mostly involved Caucasian men. Many research studies have consistently supported the use of risk-associated SNPs to more precisely estimate PCa risk among men in the general population. In light of the fact that PCa and the lethal disease affects certain races such as African Americans at disproportionate rates, this has generally been viewed as a major study limitation. However, some PCa genetic studies have been conducted in men of other ancestries including African-American, Latino, Japanese, Asian-Indian, and Chinese. Many of these studies conducted in non-Caucasian ancestries have identified unique PCa risk SNPs. Despite these potential race-specific differences, a recent study examining 82 PCa risk variants in ~4800 PCa cases and ~4700 controls of African ancestry found that 83% of variants have directionally consistent effect estimates, suggesting that the majority of GWAS-identified loci harbor risk alleles that are common and shared across populations. Based on this concept, SNPs can be used to estimate disease susceptibility in men of all races. In support of this, Hoffmann and colleagues used risk profile scoring to assess the predictive value of 105 known PCa risk SNPs. The risk score that was calculated was highly statistically significant for all major ethnic groups evaluated including Caucasians ($P \times 10^{-213}$); Latino ($P \times 10^{-56}$); East Asian ($P \times 10^{-3}$); and African-Americans ($P \times 10^{-15}$). When comparing how these scores were associated with PCA detection in each ethnic population, it was apparent that the genetic scores performed better in Caucasians and Latino groups than African-Americans and East Asians. The results of these studies demonstrate that there are SNP associations with diseases that differ between men of different ancestry. Furthermore, in a recent study by Na et al. findings demonstrated the improved ability of GRSs that use race-specific SNPs to predict PCa compared to GRSs calculated from SNP sets that include SNPs that are not implicated in PCa for an individual's race. However, regardless of ancestry, SNPs can be successfully incorporated into genetic scores to estimate PCa risk for men.

**INCORPORATION INTO SCREENING ALGORITHMS**

Guidelines that include genetic risk assessment for most diseases, including PCa, are in their infancies. However, it is important to begin to consider how GRS could be incorporated into screening algorithms may improve risk stratification and focus screening on the populations most at risk of developing PCa and lethal disease. The Stockholm 3 (STHLM3) study was a prospective, population-based, paired, screen-positive, diagnostic study of men without prostate cancer aged 50–69 years randomly invited from the Swedish Population Register. The study evaluated the utility of genetic risk assessment in combination with other serum biomarkers (PSA, free PSA, etc.) for an individual’s race.
PSA, intact PSA, hK2, MSMB, MIC1), and clinical variables (age, FH, previous prostate biopsy, prostate examination), for identifying men with PCa. The genetic score was determined using a panel of 232 PCa risk-associated SNPs (including the 100 SNPs and additional SNPs that were marginally associated with increased disease risk). Multivariate analyses demonstrated that the genetic risk assessment provided additional and independent value in identifying men at risk of harboring PCa.

Based on the overall potential clinical utility of genetic risk assessment, we propose a method by which GRS values can be incorporated with other clinical variables (e.g., FH data) to determine the age of initiation and frequency of PSA testing for men (Figure 3). For the purposes of this hypothetical screening algorithm, a GRS value of ≥1.5 is defined as a cutoff for increased genetic risk. This was because in many clinical studies (e.g., REDUCE trial), a positive family history was associated with an OR of 1.5-fold increased risk of PCa and <1.5 is defined as lower genetic risk. For men at relatively decreased risk of developing PCa (i.e., <40 years of age, GRS <1.5 and/or with negative FH), standard care should be followed. Men at higher risk of developing PCa (i.e., GRS ≥1.5 or a positive FH) should undergo PSA testing, also in accordance with many standard recommendations. The uniqueness of this cancer screening plan model is the addition of genomic information to FH information for a more comprehensive assessment of PCa risk, which allows for more personalized PCa screening.

CONCLUSIONS
Similar to FH, GRS is not intended for diagnostic use (i.e., identifying which men currently harbor PCa). Rather, GRS is a tool that has been consistently shown to aid in assessing an individual patient's propensity for developing PCa. In many ways, it should be considered a quantitative FH. However, it is much more than FH as evidenced by the fact that the positive predictive value of GRS is significantly stronger than FH. GRS values are particularly informative for all men without a FH of PCa, which is the case for the vast majority of individuals in the general population, most of whom possess, at best, incomplete knowledge of their FH. In addition, it is also important to determine GRS values for men with a FH to better assess their increased or potentially decreased risk of PCa. GRS can also be applied to all men without regard to ancestry, although GRS values that use race-specific SNPs may have even greater clinical utility.

Based on their performance in research studies, GRS values potentially have important clinical utility in an era of personalized medicine. By identifying individuals at heightened genetic risk based on GRS, clinicians can more effectively assess a patient's unique risk of developing PCa and thereby recommend personalized, “smarter” screening strategies. For those with increased genetic risk based on their GRS and FH, screening can be recommended earlier and more frequently. Similarly, for those at decreased genetic risk (e.g., low...
GENETIC RISK SCORES OF PROSTATE CANCER

BT Helfand et al

Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat Genet* 2013; 45: 385–91.

27 Ren S, Xu J, Zhou T, Jiang H, Chen H, et al. Plateau effect of prostate cancer risk-associated SNPs in discriminating prostate biopsy outcomes. *Prostate* 2013; 73: 1824–35.

28 Krieger J, Barfield R, Green RC, Kraft P. Reclassification of genetic-based risk predictions as GWAS data accumulate. *Genome Med* 2016; 8: 20.

29 Akamatsumi S, Takahashi A, Takata R, Kubo M, Inoue T, et al. Reproducibility, performance, and clinical utility of a genetic risk prediction model for prostate cancer in Japanese. *PLoS One* 2012; 7: e46454.

30 Gronberg H, Adofsson J, Aly M, Nordstrom T, Wiklund P, et al. Prostate cancer screening in men aged 50-69 years (STHLM3): a prospective population-based diagnostic study. *Lancet Oncol* 2015; 16: 1667–76.

31 Jiang H, Liu F, Wang Z, Na R, Zhang L, et al. Prediction of prostate cancer from prostate biopsy in Chinese men using a genetic score derived from 24 prostate cancer risk-associated SNPs. *Prostate* 2013; 73: 1651–9.

32 Kader AK, Sun J, Isaacs SD, Wiley KE, Yan G, et al. Individual and cumulative effect of prostate cancer risk-associated variants on clinicopathologic variables in 5,895 prostate cancer patients. *Prostate* 2009; 69: 1195–205.

33 Kears JT, Roehl KA, Cooper PR, Catalona WJ, Helfand BT. Validation of the iCOGS prostate cancer risk loci and associations with aggressive pathology in independent surgical and active surveillance cohorts. *J Clin Oncol.* [Under Review].

34 Klein EA, Cooperberg MR, Magi-Galluzzi C, Simko JP, Falzarano SM, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur Urol* 2014; 66: 550–60.

35 Nordstrom T, Aly M, Ekulund M, Egevad L, Gronberg H. A genetic score can identify men at high risk for prostate cancer among men with prostate-specific antigen of 1-3 ng/ml. *Eur Urol* 2014; 65: 1184–90.

36 Sun J, Chang BL, Isaacs SD, Wiley KE, Wiklund F, et al. Cumulative effect of five genetic variants on prostate cancer risk in multiple study populations. *Prostate* 2008; 68: 1257–62.

37 Sun J, Na R, Hsu FC, Zheng SL, Wiklund F, et al. Genetic score is an objective and better measurement of inherited risk of prostate cancer than family history. *Eur Urol* 2013; 63: 588–97.

38 Xu J, Sun J, Kader AK, Lindstrom S, Wiklund F, et al. Estimation of absolute risk for prostate cancer using genetic markers and family history. *Prostate* 2009; 69: 1565–72.

39 Zheng J, Liu F, Lin X, Wang X, Ding Q, et al. Predictive performance of prostate cancer risk in Chinese men using 33 reported prostate cancer risk-associated SNPs. *Prostate* 2012; 72: 577–83.

40 Miyakawa S, Wiklund F, Xu J, Isaacs WB, Ekulund M, et al. Polygenic risk score improves prostate cancer risk prediction: results from the Stockholm-1 cohort study. *Eur Urol* 2011; 60: 21–8.

41 Koenel LN, Altshuler D, Henderson BE. The multiethnic cohort study: exploring genetic, lifestyle, and cancer risk. *Nat Rev Cancer* 2004; 4: 519–27.

42 Catalona WJ, Bailey-Wilson JE, Camp NJ, Chanock SJ, Cooney KA, et al. National cancer institute prostate cancer genetics workshop. *Carcinogenesis* 2011; 32: 3416–23.

43 Etzioni R, Goh E, Guy J, Wolmark N, Tandon A, et al. Potential impact of adding genetic risk to clinical parameters in predicting prostate biopsy outcomes in men following an initial negative biopsy: findings from the REDUCE trial. *Eur Urol* 2012; 62: 953–61.

44 Eeles RA, Oldham RM, Benlloch S, Antoniou AC, Giles GG, et al. The genetic epidemiology of prostate cancer and its clinical implications. *Nat Rev Clin Oncol* 2014; 11: 19–31.

45 Conran CA, Na R, Chen H, Jiang D, Liu Y, et al. Population-standardized genetic risk score: the SNP-based method of choice for inherited risk assessment of prostate cancer. *Asian J Androl* 2016; 18: 520–4.

46 Amin Al Olama A, Benlloch S, Antoniou AC, Giles GG, et al. Polygenic risk analysis of prostate cancer in a Chinese hospital-based biopsy cohort. *Asian J Androl* 2015; 17: 789–95.

47 Szulkin R, Whittington T, Ekulund M, Aly M, Eeles RA, et al. Prediction of individual genetic risk to prostate cancer using a polygenic score. *Prostate* 2015; 75: 1467–74.

48 Eeles RA, Olama AA, Benlloch S, Saunders JE, Leongamornlert DA, et al.
risk-associated SNPs from genome-wide association studies of various racial populations in Chinese men. Prostate 2013; 73: 1623–35.

54 Haiman CA, Chen GK, Blot WJ, Strom SS, Berndt SI, et al. Characterizing genetic risk at known prostate cancer susceptibility loci in African Americans. PLoS Genet 2011; 7: e1001387.

55 Wang M, Takahashi A, Liu F, Ye D, Ding Q, et al. Large-scale association analysis in Asians identifies new susceptibility loci for prostate cancer. Nat Commun 2015; 6: 8469.

56 Hoffmann TJ, Van Den Eeden SK, Sakoda LC, Jorgenson E, Habel LA, et al. A large multiethnic genome-wide association study of prostate cancer identifies novel risk variants and substantial ethnic differences. Cancer Discov 2015; 5: 878–91.

57 Virlogeux V, Graff RE, Hoffmann TJ, Witte JS. Replication and heritability of prostate cancer risk variants: impact of population-specific factors. Cancer Epidemiol Biomarkers Prev 2015; 24: 938–43.