

# Risk prediction and population screening for breast, ovarian and prostate cancers

**Individual disease-associated variants may confer only modest disease risk, but the combined effect of all known associated SNPs on risk can be substantial. Using a combined genetic risk score, these studies show that profiling individuals in the highest and lowest relative risk groups has potential for the development of population-based risk screening and stratification programs. However, further studies are needed to incorporate non-genetic susceptibility factors in these models, as well as validation of these estimates and risk prediction models on additional empirical data sets.**

**Orli Bahcall**

## **Risk estimates in COGS**

In addition to the discovery of new susceptibility loci, the findings reported in the current collection of COGS papers can be applied to derive estimates for disease risk based on an individual's genetic profile. Although individual associated SNPs may confer only modest disease risk, the combined effect of all known associated SNPs on risk can be substantial. Under the assumption of a simple multiplicative model, the collection of known susceptibility SNPs for any given trait or disease can be combined into a risk profile that can be used for risk prediction.

The examples noted here use a polygenic risk score (PRS) to estimate the relative risks of disease, and a summary of the estimated relative risks by different levels of the PRS for breast, prostate and ovarian cancers is shown in **Table 1**. In these reported estimates, the predictive power is higher for prostate cancer than for breast cancer, which reflects the fact that the known associated SNP effect sizes for prostate cancer are greater and account for a larger percentage of the familial relative risk. The lower predictive estimate for ovarian cancer is primarily a reflection of the lower incidence of the disease and the need for the analysis of substantially larger numbers of cases, especially if there is significant variability across histological subtypes of the disease.

## **Large-scale genotyping identifies 41 new loci associated with breast cancer risk**

Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J. *et al.*

doi:10.1038/ng.2563

On the basis of the current set of loci and assuming that all loci combine multiplicatively, the currently known loci now define a genetic profile for which 5% of the female population has a risk that is ~2.3-fold higher than the population average, and for which 1% of the population has a risk that is ~3-fold higher. However, the large excess of significant associations among the SNPs selected from the GWAS suggests that many more susceptibility loci exist that have not met our threshold for genome-wide-significant association in this study and that these explain a similar fraction of the heritability as the currently known loci. The observation, made by comparing effect sizes in the iCOGS stage with those in the GWAS, that a very large number of loci, perhaps several thousand, contribute to polygenic susceptibility to breast cancer is consistent with results from GWAS in other complex disorders such as schizophrenia, using a different analytical approach<sup>60</sup>. Incorporating these loci into risk models should substantially improve disease prediction, even if not all loci can be identified individually. Moreover, fine-scale mapping of the identified regions may uncover more of the missing heritability, either through identifying a more strongly associated variant (as found for the *CCND1* locus; see French *et al.*<sup>61</sup>) or by identifying additional signals (exemplified for the *TERT* region in Bojesen *et al.*<sup>62</sup>). Genetic profiling using these

common susceptibility loci in combination with rarer high-risk loci and other risk factors may provide a rational basis for targeted breast cancer prevention.

#### **Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array**

Eeles, R.A., Al Olama, A.A., Benlloch, S., Saunders, E.J., Leongamornlert, D.A. *et al.*

doi:10.1038/ng.2560

With the identification of these new loci, 77 susceptibility loci for prostate cancer have now been identified. On the basis of an overall twofold familial relative risk for the first-degree relatives of prostate cancer cases and on the assumption that SNPs combine multiplicatively, the new loci reported here, together with those already known, explain approximately 30% of the familial risk of prostate cancer. Taking into consideration these SNPs and this risk model, the top 1% of men in the highest risk stratum have a 4.7-fold greater risk relative to the population average, and the top 10% of men have a 2.7-fold greater risk. For comparison, the former risk estimate is similar to that conferred by deleterious mutations in *BRCA2* (ref. 28), and such mutation carriers are undergoing targeted screening in trials, for example, in the IMPACT (Identification of Men with a genetic predisposition to ProstAte Cancer: Targeted screening in men at higher genetic risk and controls) Study (see URLs). The SNP-based prostate cancer risk profile now available should therefore be able to distinguish men at a clinically meaningful level of risk. To evaluate the combined effect of the loci associated with prostate cancer risk, we included 68 of the known loci in a logistic regression (59 which were on iCOGS and 9 for which a surrogate with  $r^2 > 0.76$  was available). The parameters from this model were used to generate polygenic risk scores (Online Methods). On the basis of these scores, the estimated risk for men in the top 1% of the risk distribution was 4.4-fold greater than the population average risk (**Supplementary Table 8**), very close to the theoretical estimate predicted under a simple polygenic model (4.7-fold). Furthermore, under a polygenic genetic risk model<sup>29</sup>, an unaffected man aged 50 who has a father with prostate cancer diagnosed at 60 years of age would have a predicted lifetime risk of prostate cancer from his family history alone of just over 20%. However, if family history is taken into consideration along with the explicit effects of all known common prostate cancer susceptibility alleles, this predicted risk would rise to just over 60% if he were in the top 1% of the known polygenic risk score distribution (A. Antoniou, personal communication). Such differences in predicted risks will be important for facilitating risk stratification in targeted screening and prevention programs.

#### **GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer**

Pharoah, P.D.P., Tsai, Y.-Y., Ramus, S.J., Phelan, C.M., Goode, E.L. *et al.*

doi:10.1038/ng.2564

However, the effects of the new susceptibility loci were modest, and together they explain less than 1% of the excess familial risk of EOC [epithelial ovarian cancer], with about 4% being explained by all known loci with common susceptibility alleles.

Assuming a log-additive model of interaction between loci, the currently known loci (**Table 2** and **Supplementary Table 3**) define a genetic risk profile with a combined variance for the log relative risk distribution of 0.057. Such a distribution has limited discriminatory ability: the estimated relative risks at the 5th and 95th percentiles are 0.63 and 1.48, respectively. However, on the basis of what is known about the architecture of genetic susceptibility for other cancers, it is probable that many more common susceptibility alleles exist.

#### **Genome-wide association study in *BRCA1* mutation carriers identifies novel loci associated with breast and ovarian cancer risk**

Couch, F.J. *et al.*

doi:10.1371/journal.pgen.1003212

### **Identification of a *BRCA2*-specific modifier locus at 6p24 related to breast cancer risk**

Gaudet, M.M., Kuchenbaecker, K.B., Vijai, J., Klein, R.J., Kirchhoff, T. *et al.*

doi:10.1371/journal.pgen.1003173

SNP profiling can also be used to predict cancer risk in *BRCA1* and *BRCA2* mutation carriers. Couch *et al.* and Gaudet *et al.* show that the majority of the breast cancer susceptibility SNPs modify breast cancer risk to a similar relative extent in *BRCA2* mutation carriers, whereas most of the SNPs for ER-negative breast cancer modify breast cancer risk in *BRCA1* mutation carriers to a similar relative extent, and most ovarian cancer SNPs modify ovarian cancer risk to a similar relative extent in *BRCA1* and *BRCA2* mutation carriers. However, because the absolute cancer risk in *BRCA1* and *BRCA2* mutation carriers is high, the absolute differences in the risk conferred by different SNP profiles are much greater than in the general population. Thus, Couch *et al.* estimate that the breast cancer risk for the 5% of *BRCA1* mutation carriers at the lowest risk, based on PRS, is 28–80%, compared with 81–100% for the carriers at the highest risk. Using a combined SNP profile, Gaudet *et al.* estimate that the 5% of the *BRCA2* mutation carriers at lowest risk have a 21–47% risk of developing breast cancer by the age of 80 years, compared to 83–100% for the 5% of carriers at the highest risk.

### **Further studies on risk estimates**

The estimates for risk prediction reported in the COGS publications and described here demonstrate the potential power of SNP profiling using a polygenic risk score. However, these predictions need to be validated empirically with direct estimates from additional case-control or, preferably, prospective cohort studies. Risk prediction methods can also be improved by incorporating information on family history, as well as other risk factors including reproductive history, hormone levels, mammographic density and environment. Evaluating the combined effects of these risk factors will be important in developing population risk stratification and screening programs.

### **Public health implications from COGS and potential for risk stratification and screening**

Burton, H., Chowdhury, S., Dent, T., Hall, A., Pashayan, N. *et al.*

doi:10.1038/ng.2582

The potential usefulness of population risk stratification for the prevention of breast cancer and prostate cancer was estimated by WP7 using detailed comparisons of risk stratification with current mammographic screening programs in the UK and a hypothetical screening strategy for prostate cancer based on age alone<sup>13</sup>. For each cancer, the 10-year absolute risk of being diagnosed with disease was estimated, taking into account age and polygenic risk profile, using all known susceptibility variants, including new variants identified in the COGS program (a total of 67 susceptibility variants for breast cancer and 72 variants for prostate cancer). The number of individuals eligible for screening and the number of cases potentially detectable by screening were estimated in a population undergoing screening on the basis of age alone in comparison to a population undergoing personalized screening. For breast cancer, using the current UK National Health Service (NHS) breast cancer screening program as a comparator, it was found that, compared with existing age-based screening (ages 47–73 years), stratified screening of women in a wider age range (ages 35–79 years) at the same 10-year absolute risk (2.5%) would be expected to result in 24% fewer women being eligible for screening while potentially detecting 3% fewer cases through screening. Similarly, with prostate cancer, in a hypothetical screening strategy comparing risk-stratified screening for men aged 45–79 years with screening of men from age 55 (10-year

**Table 1 | Estimates of the relative risks for breast, ovarian and prostate cancers based on PRS and the combined set of currently known susceptibility SNPs for each cancer.**

| Cancer   | Standard deviation of PRS | Area under ROC <sup>b</sup> | Approximate relative risk <sup>a</sup> |           |            |            |
|----------|---------------------------|-----------------------------|--|-----------|------------|------------|
|          |                           |                             | Lowest 1%                              | Lowest 5% | Highest 5% | Highest 1% |
| Breast   | 0.471                     | 63%                         | 0.25                                   | 0.34      | 2.4        | 3.2        |
| Ovarian  | 0.259                     | 57%                         | 0.49                                   | 0.57      | 1.7        | 1.9        |
| Prostate | 0.656                     | 68%                         | 0.14                                   | 0.21      | 3.2        | 4.7        |

<sup>a</sup>Estimated relative risk compared to the population mean for the specified PRS quantile (for example, lowest 1% refers to the 1% of the population with the lowest PRS estimates). <sup>b</sup>Estimated area under the receiver operator characteristic (ROC) curve for a predictive test based on PRS.

absolute risk of prostate cancer of 2%), 19% fewer men would be eligible for screening at a cost of 4% fewer cases potentially detected by screening.

The advantages of such stratified programs would be increased opportunities to detect cancers in individuals of younger ages, in whom the cancers tend to behave more aggressively<sup>14,15</sup>, and reduced risk of false positives, with reduction of harm due to unnecessary biopsies and invasive treatments. However, to estimate the true benefits of risk-stratified screening, it will be necessary to understand whether and how tumor subtypes, screening test sensitivity, the natural history of cancers and the probability of over-diagnosis (the diagnosis of indolent cancers that may never have manifested clinically) vary by polygenic profile.

This COGS WP7 discussion on the potential effectiveness of risk stratification was based entirely on mathematical modeling. The eventual implementation of stratified prevention programs will require the development of risk prediction tools and empirical as well as modeled evidence of the effectiveness of risk-stratified interventions. Thus, it will be necessary to evaluate risk models using empirical data sets. Demonstration of the benefits of stratified screening, particularly in the reduction of cancer-specific mortality, would ideally be achieved through a randomized trial, but this may not always be feasible. Alternatively, it may be reasonable to implement stratified screening on the basis of the best available evidence, including validated model findings, and then have pragmatic service evaluation.