

# Public health genomics and personalized prevention: lessons from the COGS project

■ N. Pashayan<sup>1</sup>, A. Hall<sup>2</sup>, S. Chowdhury<sup>2</sup>, T. Dent<sup>2</sup>, P. D. P. Pharoah<sup>3</sup> & H. Burton<sup>2</sup>

From the <sup>1</sup>Department of Applied Health Research, University College London, London; <sup>2</sup>PHG Foundation, Cambridge; and <sup>3</sup>Departments of Oncology and of Public Health and Primary Care, University of Cambridge, Cambridge, UK

**Abstract.** Pashayan N, Hall A, Chowdhury S, Dent T, Pharoah PDP, Burton H (University College London, London; PHG Foundation, Cambridge; University of Cambridge, Cambridge, UK). Public health genomics and personalized prevention: lessons from the COGS project. (Review). *J Intern Med* 2013; **274**: 451–456.

Using the principles of public health genomics, we examined the opportunities and challenges of implementing personalized prevention programmes for cancer at the population level. Our model-based estimates indicate that polygenic risk stratification can potentially improve the effective-

ness and cost-effectiveness of screening programmes. However, compared with ‘one-size-fits-all’ screening programmes, personalized screening adds further layers of complexity to the organization of screening services and raises ethical, legal and social challenges. Before polygenic inheritance is translated into population screening strategy, evidence from empirical research and engagement with and education of the public and the health professionals are needed.

**Keywords:** ethical, legal and social issues, personalized prevention, polygenic risk stratification, public health genomics, screening.

## Introduction

The success of the Collaborative Oncological Gene-environment Study (COGS) [1] in identifying multiple new genetic variants associated with the risk of breast, prostate and ovarian cancers promises a new era in personalized prevention. By stratifying the population into several groups according to genetic risk alone or combined with traditional disease risk factors (such as age and family history), standard public health interventions could be applied differentially to each population stratum with potentially more efficient outcome [2, 3]. However, the challenge remains how to translate genomic knowledge into health improvement programmes and disease prevention interventions at the population level. Responding to such a challenge represents the core work of public health genomics.

Public health genomics is a relatively new discipline that emerged in 1997 when small units were established independently on both sides of the Atlantic – the Office of Genetics and Disease Prevention at the Centers for Disease Control and Prevention in Atlanta, USA, and the Public Health

Genetics Unit in Cambridge, UK. Subsequently, following a multidisciplinary expert meeting supported by the Rockefeller Foundation at Bellagio, Italy, in 2005, public health genomics was formally defined as ‘the responsible and effective translation of genome-based knowledge and technologies for the benefit of population health’, and the broad content and methodology of this discipline was specified [4, 5]. It was stated at the Bellagio workshop that improving population health would require an integrated and multidisciplinary approach, taking account of emerging genome-based science and technologies and integrating such knowledge with the humanities, social sciences and population sciences, particularly epidemiology. An important part of the definition of public health genomics was that interventions arising from genomic research should not be introduced into clinical or public health practice prematurely without the support of a solid evidence base. The element of ‘responsibility’ in the definition demonstrated the importance of the societal context in which new technologies will be applied and the collateral effects on individuals and populations. Finally, the definition encompassed the view that implementation of new genetic applications would require an active process, involving multidisciplinary analysis, public dialogue and involvement, engagement with a wide

The copyright line for this article has been changed since first published on 16 OCT 2013.

range of stakeholders, informing public policy, developing new programmes and services, strategic planning and evaluating health services, and training and educating health professionals, policy makers and the public.

In this review, we illustrate the potential utility of polygenic risk stratification in prevention programmes at a population level using the case of population-based screening for breast and prostate cancers [6]. Using the principles of public health genomics, we examine the challenges of implementing personalized screening programmes. These analyses were carried out as part of the Work Package 7 of the COGS project funded by the Seventh Framework Programme of the European Commission.

#### Polygenic risk stratification

The risk alleles identified by genome-wide association studies individually confer a modest increase in risk of disease (usually per-allele relative risk of less than 1.5). Consequently, the predictive utility of a genetic test based on a single risk allele is poor. Even using a combination of multiple alleles, the clinical utility of a polygenic test in predicting future disease for the individual will be limited [7]. This is because most individuals are at only slightly increased or decreased risk, and the difference is modest even for the small number at the extremes of the distribution. However, discrimination (i.e. determining which individuals will or will not experience disease) is not the only measure of clinical utility of a risk prediction model. In disease prevention, the aim is to stratify risk rather than to discriminate events [8].

According to the multiplicative model, the polygenic risk in the population at birth follows the normal distribution when relative risk is plotted on a logarithmic scale. The normal distribution is defined by its mean and variance or spread [9]. The variance of the distribution of polygenic risk is calculated from the risk allele frequencies and per-allele relative risk [2, 10]. The distribution of relative risk amongst cases is also log-normal with the same variance as the population, but with a larger mean; cases are on average at a higher prior risk of developing cancer than the population [9]. Given the variance and mean of the distribution of the log-normal relative risk, the percentile rank associated with a given polygenic relative risk (or age-conditional absolute risk) threshold in the population or in cases can be calculated [6].

#### Utility of polygenic risk stratification for personalized screening

We modelled the number of individuals eligible for screening and the number of cases potentially detectable by screening in a population undergoing screening based on age alone, as compared to a population undergoing personalized screening based on the 10-year absolute risk of being diagnosed with breast or prostate cancer. We calculated the conditional absolute risk taking into account age and polygenic risk profile. We set the risk threshold equivalent to the threshold for eligibility in the age-based screening programme [6].

In the case of breast cancer, we estimated the polygenic risk based on 67 common genetic susceptibility variants [11], which confer a polygenic variance of 0.28 and explain approximately 14% of the genetic component of breast cancer risk. Compared with the UK National Health Service Breast Screening Programme, which offers screening to women aged 47 to 73 years (10-year absolute risk of being diagnosed with breast cancer of 2.5% or greater), personalized screening of women aged 35 to 79 years at the same risk threshold would result in 24% fewer women being eligible for screening whilst potentially detecting 3% fewer cases.

In the UK, there is no national screening programme for prostate cancer, but a similar approach could be applied if there were. We estimated the polygenic risk of prostate cancer based on 72 common susceptibility variants [12], which confer a polygenic variance of 0.44 and explain approximately 30% of the genetic component of prostate cancer risk. Compared with a hypothetical screening strategy based on age alone in which men are eligible for screening from age 55 to 79 years (10-year absolute risk of being diagnosed with prostate cancer of 2% or greater), personalized screening of men aged 45 to 79 years at the same risk threshold would result in 19% fewer men being eligible for screening at a cost of 4% fewer potentially screen-detectable cases.

The efficiency of personalized screening will improve as more susceptibility variants are known. Based on the currently known susceptibility variants for breast cancer, 43% of 35- to 79-year-old women would be eligible for screening with 71% of cases being potentially screen detectable. In a hypothetical best-case-scenario analysis, assuming all possible susceptibility variants for breast

cancer were known (predicted polygenic variance of 1.44 [9, 13]), 28% would be eligible for screening whilst potentially detecting 76% of cases. For prostate cancer, based on the currently known susceptibility variants, 51% of 45- to 79-year-old men would be eligible for screening whilst potentially detecting 92% of cases. In the best-case-scenario (predicted polygenic variance of 1.58), 34% would be eligible whilst potentially detecting 89% of cases (Table 1).

Figure 1 shows how stratification based on the absolute risk that is dependent on age and polygenic risk profile reclassifies individuals into different risk groups, whereby some cancers that would have been detected under age-based screening will not be detected by personalized screening and vice versa. Such reclassification may improve the benefits and reduce some of the harms associated with screening. A personalized screening programme would enable the detection of cancer in younger individuals who are at high risk. Prostate and breast cancers detected in younger individuals tend to behave more aggressively [14, 15]. So detecting

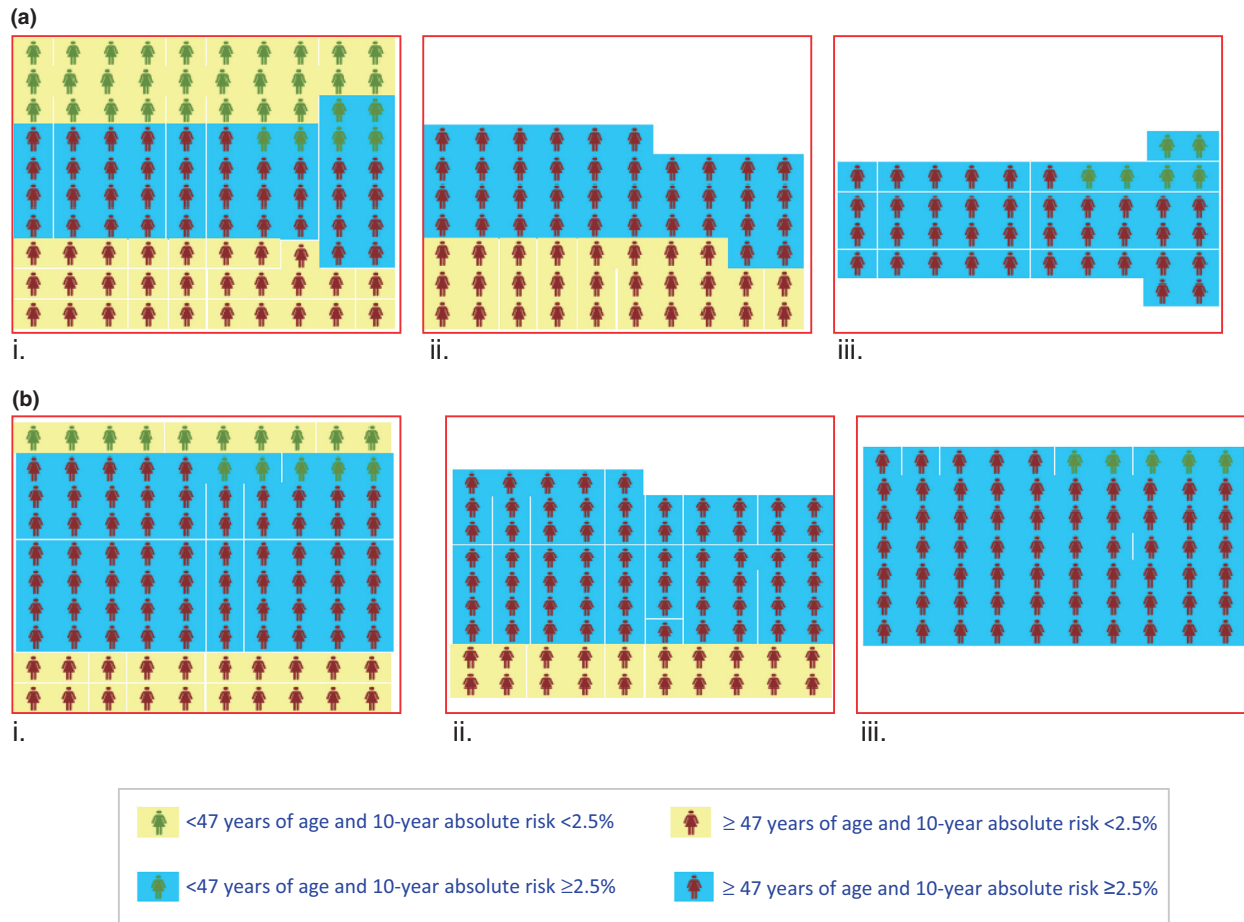
these cancers amongst younger individuals at high risk may improve their prognosis. If polygenic low risk is associated with indolent and possibly overdiagnosed cancer, then not detecting cancer amongst individuals at low risk may potentially reduce the harms associated with overdiagnosis and overtreatment.

Under an age-based screening strategy, all individuals within the age range will be eligible for screening, whereas individuals will be invited for screening at different ages based on their absolute risk level under a personalized screening strategy. For example, in the case of screening for prostate cancer, instead of inviting all 55- to 79-year-old men, 2%, 17% and 96% of men aged 47, 50 and 65 years, respectively, would be invited for screening. Thus, different individuals will undergo different numbers of screening episodes during their lifetime. A reduction in the number of individuals offered screening may reduce the number of false-positive screens, with a reduction in the harms associated with false-positive results and the benefit of saving further resources on diagnostic tests.

**Table 1** The likely percentage of the population eligible for screening and of the cases potentially detectable by screening, considering age-based and personalized screening for breast and prostate cancers with increasing numbers of known susceptibility variants

Screening strategy	Population eligible for screening (%)	Cases potentially detectable by screening (%)
	Age 35–79 years; <i>n</i> = 13 126 890	Age 35–79 years; <i>n</i> = 30 936
<b>Breast cancer</b>		
Age-based screening (47–73 years)	57	72
Personalized screening (age 35–79 years and 10-year absolute risk $\geq 2.5\%$ )		
Currently known variants (variance = 0.28)	43	71
Variants explaining 50% of familial risk (variance = 0.72)	35	72
Variants explaining 100% of familial risk (variance = 1.44)	28	76
<b>Prostate cancer</b>		
Age-based screening (55–79 years)	63	96
Personalized screening (age 45–79 years and 10-year absolute risk $\geq 2.0\%$ )		
Currently known variants (variance = 0.44)	51	92
Variants explaining 50% of familial risk (variance = 0.79)	45	91
Variants explaining 100% of familial risk (variance = 1.58)	34	89

<sup>a</sup>Estimates are based on the population and cancer registrations in 2002–2006 in England.



**Fig. 1** Reclassification of women into different risk groups. *a*, Eligibility for screening: a population of 100 women, 35–79 years of age, by age group and risk threshold (i), eligible for screening based on age alone (ii) or on age and polygenic risk (iii). *b*, Potentially screen-detectable cases: a population of 100 women, 35–79 years of age, with breast cancer by age group and risk threshold (i), detectable by screening based on age alone (ii) or on age and polygenic risk (iii).

However, to estimate the true benefits of personalized screening, we need to understand whether and how tumour subtypes, screening test sensitivity, the natural history of cancer and the probability of overdiagnosis vary by polygenic risk profile. Further evidence from empirical data is needed.

#### Challenges of implementing personalized screening

In addition to considering the utility of polygenic risk stratification for personalized screening, there are many other issues to consider before a risk-tailored programme could be put into practice. These are interdependent and include cost-effectiveness, public and professional acceptability, and

organizational and a wide range of ethical, legal and social issues.

#### Evaluation of cost-effectiveness

A screening programme becomes viable if it does more good than harm at a reasonable cost [16]. A personalized screening strategy involves the additional cost of polygenic risk profiling set against the potential savings from repeat screening and diagnostic work-up of false-positive results. Our preliminary cost-utility analysis comparing age-based screening to a personalized screening programme for prostate cancer indicates that personalized screening would cost less and would increase the number of quality-adjusted life years gained.

*Exploring organizational and ethical, legal and social issues*

We convened international multidisciplinary stakeholder workshops, followed by ongoing detailed research and policy analysis of the key issues identified [17–19]. The workshop participants included clinicians, geneticists, public health specialists, epidemiologists, ethicists, lawyers, social scientists, screening programme managers, a journalist and public representatives. In addition to panel discussions and small working groups, we used hexagon modelling, a systematic technique developed by IDON [20], to brainstorm, capture and determine the priority of the issues raised. The key issues are summarized below.

An overriding concern was whether eligibility for screening based on a risk score that includes genetic profiling alongside environmental and lifestyle risk factors would be acceptable to the public, health professionals and policy makers. Particularly where there is already an established screening programme, such as that for breast cancer in the UK, not or less frequently offering screening to lower-risk groups may be perceived as service rationing and unfair discrimination based on a genetic profile that is unalterable and therefore not under control of the individual. The DNA testing involved in risk profiling may itself cause concerns amongst the public about who might access these data and whether they could be used for purposes other than for health care or public health, such as for medical research or forensic investigation, or to inform decisions about insurance or employment.

The greatly increased complexity of a polygenic risk-based screening programme could result in a further set of issues. These are evident in both the design and delivery of the programme. A major organizational challenge will be to incorporate the advances of the rapidly evolving field of genomics and the changes in individuals' environmental and lifestyle risk factors, including their family history of disease, into a dynamic risk estimation tool [21]. In terms of the environmental and lifestyle risk factors, it is critical to determine how, when, with what frequency and by whom this information will be obtained and updated. For the genetic data, it will be necessary to decide the age of sampling and testing, details of the variants to be included, whether rare highly penetrant mutations such as *BRCA1/2* will be included and whether to use a bespoke set of variants related to the particular condition or to use whole-genome sequencing for a

panel of common chronic diseases. In the latter case, the relevant information regarding prevention programmes could be extracted as and when required [22].

The programme design also needs to set out the different patterns of screening for which people in differing risk strata will be eligible. Depending on further research evidence, lower-risk groups may be offered screening less frequently and higher-risk groups may be offered screening with more sensitive (and possibly more expensive) tests, such as with magnetic resonance imaging instead of mammography for breast cancer screening [23]. To adapt services already in place or to develop new ones, substantial organizational effort will be required supported by a range of evaluative research to ensure that the level of extra value merits the additional complexity and cost.

Finally, the delivery of the programme and, in particular, the interaction of those at the frontline of provision with participants must be considered [24]. Individuals need to be able to make an informed decision about whether to accept an offer to participate. To support this and to protect the public, it is important to establish policies, for example, regarding: (i) the amount and type of information to be provided to the public, (ii) how the different levels of risk will be communicated, (iii) the potential benefits and possible harms of screening interventions for different risk strata and (iv) regulations on data safeguarding. Health professionals involved in service delivery will need training to better understand how the risk scores are determined and the evidence on which different management options are based, to communicate these effectively and to be able to provide support to patients at various stages of the programme and with different outcomes.

**Conclusion**

Polygenic risk stratification would potentially improve the effectiveness and cost-effectiveness of screening programmes. However, in terms of set-up and delivery, a risk-tailored screening programme is much more complex than a programme with eligibility based on age alone. Any decision to initiate a personalized screening programme should first address the organizational and ethical, legal and social issues and commit to public engagement and education and to work with the health professionals delivering the programme.

From a wider perspective, the COGS project illustrates the enterprise of public health genomics. We have shown how the practice of public health genomics builds on a detailed understanding of the underlying science and technology and incorporates elements of population sciences such as epidemiology, health economics and consideration of ethical, legal and social issues as well as policy development and change management.

We have illustrated the complexity of knowledge that must be integrated and the detailed planning, consultation and development work that would be required to underpin such a new programme. As the definition of public health genomics suggests, the 'responsible translation of genomic technology for the improvement of population health', in its widest sense, will demand no less.

#### Conflict of interest statement

The authors have no conflict of interests to declare.

#### Acknowledgements

This work is part of the Collaborative Oncological Gene-environment Study ([www.cogseu.org](http://www.cogseu.org)), funded by the Seventh Framework Programme of the European Commission under grant agreement 223175 (HEALTH-F2-2009-223175). NP is supported by a Cancer Research UK Clinician Scientist Fellowship.

#### References

- Collaborative Oncological Gene-environment Study. <http://www.cogseu.org/> (last accessed on 06-05-2013).
- Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med* 2008; **358**: 2796–803.
- Zimmern RL. Genomics and individuals in public health practice: are we luddites or can we meet the challenge? *J Public Health (Oxf)* 2011; **33**: 477–82.
- Burke W, Khoury MJ, Stewart A, Zimmern RL. The path from genome-based research to population health: development of an international public health genomics network. *Genet Med* 2006; **8**: 451–8.
- Bellagio Group. Report of an expert workshop held at the Rockefeller Foundation Study. 2005. <http://www.phgfoundation.org/search/?q=bellagio+> (last accessed on 04-10-2012).
- Pashayan N, Duffy SW, Chowdhury S, Dent T, Burton H, Neal DE *et al.* Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. *Br J Cancer* 2011; **104**: 1656–63.
- Janssens AC, van Duijn CM. Genome-based prediction of common diseases: advances and prospects. *Hum Mol Genet* 2008; **17**: R166–73.
- Hingorani AD, Shah T, Kumari M, Sofat R, Smeeth L. Translating genomics into improved healthcare. *BMJ* 2010; **341**: c5945.
- Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* 2002; **31**: 33–6.
- Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst* 2008; **100**: 1037–41.
- Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013; **45**: 353–61.
- Eeles RA, Olama AA, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M *et al.* Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat Genet* 2013; **45**: 385–91.
- Risch N. The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. *Cancer Epidemiol Biomarkers Prev* 2001; **10**: 733–41.
- Fredholm H, Eaker S, Frisell J, Holmberg L, Fredriksson I, Lindman H. Breast cancer in young women: poor survival despite intensive treatment. *PLoS ONE* 2009; **4**: e7695.
- Lin DW, Porter M, Montgomery B. Treatment and survival outcomes in young men diagnosed with prostate cancer: a Population-based Cohort Study. *Cancer* 2009; **115**: 2863–71.
- Gray JA, Patnick J, Blanks RG. Maximising benefit and minimising harm of screening. *BMJ* 2008; **336**: 480–3.
- Chowdhury S, Dent T, Pashayan N, Hall A, Lyratzopoulos G, Hallowell N *et al.* Incorporating genomics into breast and prostate cancer screening: assessing the implications. *Genet Med* 2013; **15**: 423–432.
- Hall AE, Chowdhury S, Pashayan N, Hallowell N, Pharoah P, Burton H. What ethical and legal principles should guide the genotyping of children as part of a personalised screening programme for common cancer?. *J Med Ethics* 2013; doi:10.1136/medethics-2012-101079.
- Dent T, Jbilou J, Rafi I, Segnan N, Tornberg S, Chowdhury S *et al.* Stratified cancer screening: the practicalities of implementation. *Public Health Genomics* 2013; **15**: 423–432.
- Tayko PRM, Saponkij U. The power of hexagon as an organization development process. *AU-GSB e-J* 2011; **4**: 8–17.
- Howell A, Astley S, Warwick J, Stavrinou P, Sahin S, Ingham S *et al.* Prevention of breast cancer in the context of a national breast screening programme. *J Intern Med* 2012; **271**: 321–30.
- Biesecker LG. Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: lessons from the ClinSeq project. *Genet Med* 2012; **14**: 393–8.
- Morrow M, Waters J, Morris E. MRI for breast cancer screening, diagnosis, and treatment. *Lancet* 2011; **378**: 1804–11.
- Evans DG, Warwick J, Astley SM, Stavrinou P, Sahin S, Ingham S *et al.* Assessing individual breast cancer risk within the U.K. National Health Service Breast Screening Program: a new paradigm for cancer prevention. *Cancer Prev Res (Phila)* 2012; **5**: 943–51.

Correspondence: Nora Pashayan, Department of Applied Health Research, University College London, 1–19 Torrington Place, London, WC1E 6BT, UK.  
(fax: +44-20-7813-0280; e-mail: [n.pashayan@ucl.ac.uk](mailto:n.pashayan@ucl.ac.uk)).