Integration of genetic and epigenetic markers for risk stratification: opportunities and challenges

Nora Pashayan[1*], Daniel Reisel[2], Martin Widschendter[2]

[1]University College London, Department of Applied Health Research, London, UK

[2]University College London, Department of Women's Cancer, London, UK

E-mail addresses:

NP: n.pashayan@ucl.ac.uk

DR: d.reisel@ucl.ac.uk

MW: m.widschwendter@ucl.ac.uk

* Correspondence to

Nora Pashayan

Department of Applied Health Research

University College London

1-19 Torrington Place, London, WC1E 7HB

Tel: +44 20 3108 3255; Fax: +44 20 7813 0280

E-mail: n.pashayan@ucl.ac.uk

Summary

Common genetic susceptibility variants could be used for risk stratification in risk-tailored cancer screening and prevention programmes. Combining genetic variants with environmental risk factors would improve risk stratification. Epigenetic changes are surrogate markers of environmental exposures during individual's lifetime. Integrating epigenetic markers, in lieu of environmental exposure data, with genetic markers would potentially improve risk stratification. Epigenetic changes are reversible and acquired gradually, providing potentials for prevention and early detection strategies. The epigenetic changes are tissue-specific and stage-of-development-specific, raising challenges in choice of sample and timing for evaluation of cancer-associated changes. The Horizon 2020 funded research programme, FORECEE, using empirical data, will investigate the value of integration of epigenomics with genomics for risk prediction and prevention of women-specific cancers.

Key words

Risk assessment, risk stratification, polygenic risk, epigenetics, cancer

To date, model-based estimates have shown the potential utility of common genetic susceptibility loci for risk-stratification in cancer prevention and screening programmes at population level [1;2]. A risk stratified screening strategy has the potential to improve the efficiency of the screening programme and reduce its adverse consequences. For example, a risk-stratified screening strategy for breast cancer with eligibility for screening based on an absolute risk that is dependent on age and polygenic risk-profile has been shown to reduce the number of women invited to screening while detecting most cancers potentially detectable by a conventional age-based screening strategy [1]. Studies in prostate cancer have shown that targeting screening to men at higher than population average risk, based on polygenic risk, could reduce the proportion of prostate cancers likely to be overdiagnosed [3;4].

A risk-stratified prevention or screening programme would involve risk assessment, then stratifying the population into several risk groups and offering the intervention differentially to each population stratum with the expectation of improving the benefit – harm balance of the intervention [5]. Interventions, like screening, would be risk-tailored, with varying start and end age, frequency, and modality of screening [6].

For risk assessment, genetic and environmental risk factors could be used. A broad definition of environment includes lifestyle, nutrition, external environmental exposures, and reproductive history [7]. To date, genome-wide association studies have identified 94 breast cancer susceptibility variants [8]. The polygenic risk profile based on these variants has area under the receiver operator characteristic curve (AUC_{-ROC}) of 0.65. AUC_{-ROC} is the probability that a test correctly identifies an individual who will develop the disease from a pair of whom one will be affected and one will remain unaffected. AUC_{-ROC} values range from 0.5 (total lack of discrimination) to 1.0 (perfect discrimination) [9]. Combining polygenic risk with environmental risk factors, like age of menarche, number of births, age of first live birth, oral contraceptive use, body mass index, alcohol, smoking, personal history of benign breast disease, and family history of breast cancer in first-degree relatives [2], would increase AUC_{-ROC} to 0.68. Even a modest increase in AUC_{-ROC} following combining genetic with environmental risk factors would substantially improve risk-stratification [2].

Epigenome as readout for environmental exposures

Age, environmental exposures (such as tobacco, alcohol, infectious agents) and endogenous stimuli (such as circulating hormones) can trigger alterations in the epigenetic pattern. These alterations affect gene expression without changing the nucleotide sequence. Epigenetic changes are generally stable and propagate over cell divisions resulting in changes in phenotype. The predominant mechanisms that may act alone or in combination to regulate gene expression over the lifetime of

an organism include DNA methylation, histone modification, microRNA (miRNA) expression and processing, and chromatin condensation [10].

Epigenetic mechanisms are implicated in the initiation and progression of cancer through tumour suppressor genes silencing and / or oncogenes activation. DNA methylation is the most studied mechanism of epigenetic gene regulation [11]. DNA methylation refers to the addition of methyl group to cytosine base that is located 5' to a guanosine base in a CpG dinucleotide. Regions of the genome rich in CpG are known as CpG islands, which are found in the promotor region of approximately half of all genes. Aberrant promotor hypermethylation associated with inappropriate gene silencing plays important role in tumour progression [12]. Whereas genome-wide decrease in methylation may lead to genomic instability and is associated with tumour progression [13].

Environmental exposures may promote tumour development by inducing both epigenetic and genetic changes (like mutations). The epigenetic make-up may play a role in the cellular response to these environmental exposures. Epigenetics may mediate or modify genetic risk [14]. As such individual's susceptibility to cancer would depend on both genetics and epigenetics make-up [15].

Opportunities and challenges

Epigenetic changes unlike the genetic ones are reversible and can be modulated for example by diet, drugs, and other environmental factors. This reversibility provides opportunity for cancer prevention strategies [10].

Environmentally triggered epigenetic changes are acquired gradually and influence tumour initiation and progression. In cervical cancer, epigenetic changes have been detected in morphologically normal cells years in advance of neoplastic transformation [16]. DNA methylation has been shown to be differentially variable between normal, intra-epithelial neoplasia, and invasive cervical cancers [17]. In esophageal cancer, DNA methylation profile has been shown to predict progression of dysplasia in Barrett's esophagus [18]. These indicate that epigenetic changes could be used for risk prediction and early diagnosis.

Epigenetic alterations can act as surrogate markers of environmental exposures during individual's lifetime. Information collected on environmental exposures via questionnaire or direct measurement is susceptible to recall bias and to inadequate capturing of exposures with short half-lives and of low biological dose. Epigenetic markers when used in lieu could overcome some of these limitations [14]. As combining polygenic risk with environmental risk factors improves risk stratification, then integration of epigenetic markers and polygenic risk into a risk assessment tool would potentially improve risk stratification in cancer screening and prevention programmes.

Unlike the genome, which is the same for all types of cells, the epigenome varies across cell types, in the same cell at different developmental stage, under the influence of different environmental exposures, and over time. For example, buccal cells exhibit significantly more smoking associated DNA methylation changes than blood cells. These changes correlate with DNA methylation changes in epithelial cancers and particularly with smoke related epithelial cancers, notably lung cancer [19]. The tissue-specificity of the epigenetic changes raises an issue with accessing tissue samples from organs like the prostate, ovary, breast, which are inaccessible by non-invasive means. Alternatively, tissue surrogates could be used. For example, epithelial cells from the uterine cervix are hormone sensitive and are likely to capture an 'epigenetic record' of breast cancer risk factors, and hence could be used as surrogate for breast cancer cells [20].

This plasticity of the epigenome poses several challenges in epigenetic testing. While polygenic risk could be determined at any point in time, a risk score that includes epigenetic markers needs to be evaluated several times to follow cancer-associated alterations. To identify a risk marker, it is important to differentiate between association and causality of epigenetic mark with disease [10], that is whether a change is adaptive response to environmental exposures or adverse effect with phenotypic consequences [15]. International efforts, such as the International Human Epigenome Consortium, and the reference epigenome, will enable comparison and identification of adverse effects [15].

Integration of different types of genomic, epigenomic, and epidemiological data would be challenging. To overcome this, there is need for engaging geneticists, bioinformaticians, statisticians, and clinicians and for developing improved strategies for handling large databases, data analysis and interpretation.

Future perspectives

The European Commission Framework Programmes, Horizon 2020, recently funded the multi-disciplinary research programme, FORECEE (Female cancer prediction using cervical omics to individualise screening and prevention) (http://cordis.europa.eu/search/result_en?q=forecee) to assess the validity, public health utility, cost-effectiveness, acceptability, and ethical, legal, social and regulatory impacts of using genomics, epigenomics, and metagenomics data in personalized screening and prevention programmes for breast, ovarian, endometrial and cervical cancers. To discover new DNA methylation signatures, an epigenome wide association study (EWAS) will be conducted using prospectively collected cervical cells from liquid based cytology cervical smears. Women's cancer risk identification test that combines DNA methylation signatures with SNPs and the microbiome will be developed and validated to predict the absolute risk of developing breast,

ovarian, endometrial or cervical cancer. In the coming four years, the multidisciplinary team with expertise in oncology, genetics, genetics statistics, omics technologies, bioinformatics, decision analysis, epidemiology, public health, health economics, ethics, risk communication, will be addressing the challenges with developing and implementing omics based cancer prevention and screening programme.

Conclusion

Cancer is a genetic and epigenetic disease. Combining genetic and epigenetic markers provides huge potentials for risk stratification in cancer control programmes. Multi-disciplinary efforts are needed to overcome the challenges.

Financial disclosure / Acknowledgement

NP is Cancer Research UK Clinician Scientist Fellow.

References

(1) Pashayan N, Duffy S W, Chowdhury S, Dent T, Burton H, Neal D E et al. Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. *Br J Cancer* 2011; 104(10):1656-1663.

(2) Garcia-Closas M, Gunsoy N B, Chatterjee N. Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer. *J Natl Cancer Inst* 2014; 106(11).

(3) Pashayan N, Duffy S W, Neal D E, Hamdy F C, Donovan J L, Martin R M et al. Implications of polygenic risk-stratified screening for prostate cancer on overdiagnosis. *Genet Med* 2015.

(4) Pashayan N, Pharoah P D, Schleutker J, Talala K, Tammela T L, Maattanen L et al. Reducing overdiagnosis by polygenic risk-stratified screening: findings from the Finnish section of the ERSPC. *Br J Cancer* 2015.

(5) Burton H S G P P a Z R. Time to revisit Geoffrey Rose: strategies for prevention in the genomic era? *Ital J Public Health* 2012; 9(4):e8665-1-e8665-9.

(6) 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.

(7) Widschwendter M, Jones A, Teschendorff A E. Epigenetics makes its mark on women-specific cancers--an opportunity to redefine oncological approaches? *Gynecol Oncol* 2013; 128(1):134-143.

(8) Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush M J et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* 2015; 47(4):373-380.

(9) van Z M, van der Net J B, Kundu S, Freedman A N, van Duijn C M, Janssens A C. Genome-based prediction of breast cancer risk in the general population: a modeling study based on meta-analyses of genetic associations. *Cancer Epidemiol Biomarkers Prev* 2011; 20(1):9-22.

(10) Verma M, Khoury M J, Ioannidis J P. Opportunities and challenges for selected emerging technologies in cancer epidemiology: mitochondrial, epigenomic, metabolomic, and telomerase profiling. *Cancer Epidemiol Biomarkers Prev* 2013; 22(2):189-200.

(11) Widschwendter M, Jones P A. DNA methylation and breast carcinogenesis. *Oncogene* 2002; 21(35):5462-5482.

(12) Jones P A, Baylin S B. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002; 3(6):415-428.

(13) Hamilton J P. Epigenetics: principles and practice. *Dig Dis* 2011; 29(2):130-135.

(14) Ladd-Acosta C, Fallin M D. The role of epigenetics in genetic and environmental epidemiology. *Epigenomics* 2015.

(15) Herceg Z, Lambert M P, van V K, Demetriou C, Vineis P, Smith M T et al. Towards incorporating epigenetic mechanisms into carcinogen identification and evaluation. *Carcinogenesis* 2013; 34(9):1955-1967.

(16) Teschendorff A E, Jones A, Fiegl H, Sargent A, Zhuang J J, Kitchener H C et al. Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation. *Genome Med* 2012; 4(3):24.

(17) Teschendorff A E, Liu X, Caren H, Pollard S M, Beck S, Widschwendter M et al. The dynamics of DNA methylation covariation patterns in carcinogenesis. *PLoS Comput Biol* 2014; 10(7):e1003709.

(18) Jin Z, Cheng Y, Gu W, Zheng Y, Sato F, Mori Y et al. A multicenter, double-blinded validation study of methylation biomarkers for progression prediction in Barrett's esophagus. *Cancer Res* 2009; 69(10):4112-4115.

(19) Teschendorff A E, Yang Z, Wong A, Pipinikas C P, Jiao Y, Jones A et al. Correlation of Smoking-Associated DNA Methylation Changes in Buccal Cells With DNA Methylation Changes in Epithelial Cancer. *JAMA Oncol* 2015; 1(4):476-485.

(20) Anjum S, Fourkala E O, Zikan M, Wong A, Gentry-Maharaj A, Jones A et al. A BRCA1-mutation associated DNA methylation signature in blood cells predicts sporadic breast cancer incidence and survival. *Genome Med* 2014; 6(6):47.