

# Genome-wide association studies and polygenic risk scores for skin cancer: clinically useful yet?

M.R. Roberts,<sup>1,2</sup> M.M. Asgari<sup>1,2</sup> and A.E. Toland <sup>3</sup>

<sup>1</sup>Department of Dermatology, Massachusetts General Hospital, Boston, MA, U.S.A.

<sup>2</sup>Department of Population Medicine, Harvard Pilgrim Healthcare Institute, Boston, MA, U.S.A.

<sup>3</sup>Department of Cancer Biology and Genetics, Comprehensive Cancer Center, Ohio State University, 998 Biomedical Research Tower, 460 W 12th Ave, Columbus, OH 43210, U.S.A.

## Summary

### Correspondence

Amanda Ewart Toland.

E-mail: [Amanda.Toland@osumc.edu](mailto:Amanda.Toland@osumc.edu)

### Accepted for publication

18 March 2019

### Funding sources

This work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (K24 AR069760 to M.M.A.) and the National Institutes of Health (R01 CA215151 and R21 CA219884 to A.E.T.). The funding agencies had no role in the development and preparation of this review.

### Conflicts of interest

None declared.

DOI 10.1111/bjd.17917

**Background** Genome-wide association studies (GWAS) have identified thousands of susceptibility variants, although most have been associated with small individual risk estimates that offer little predictive value. However, combining multiple variants into polygenic risk scores (PRS) may be more informative. Multiple studies have developed PRS composed of GWAS-identified variants for cutaneous cancers. This review highlights data from these studies.

**Objectives** To review published GWAS and PRS studies for melanoma, cutaneous squamous cell carcinoma (cSCC) and basal cell carcinoma (BCC), and discuss their potential clinical utility.

**Methods** We searched PubMed and the National Human Genome Research Institute–European Bioinformatics Institute GWAS catalogue to identify relevant studies.

**Results** Results from 21 GWAS (11 melanoma, 3 cSCC, 7 BCC) and 11 PRS studies are summarized. Six loci in pigmentation genes overlap between these three cancers (ASIP/RALY, IRF4, MC1R, OCA2, SLC45A2 and TYR). Additional loci overlap for cSCC/BCC and BCC/melanoma, but no other loci are shared between cSCC and melanoma. PRS for melanoma show roughly two-to-threefold increases in risk and modest improvements in risk prediction (2–7% increases). PRS are associated with twofold and threefold increases in risk of cSCC and BCC, respectively, with small improvements (2% increase) in predictive ability.

**Conclusions** Existing data indicate that PRS may offer small, but potentially meaningful, improvements to risk prediction. Additional research is needed to clarify the potential utility of PRS in cutaneous carcinomas. Clinical translation will require well-powered validation studies incorporating known risk factors to evaluate PRS as tools for screening.

### What's already known about this topic?

- Over 50 susceptibility loci for melanoma, basal cell carcinoma (BCC) and cutaneous squamous cell carcinoma (cSCC) have been identified in genome-wide association studies (GWAS).
- Polygenic risk scores (PRS) using variants identified from GWAS have also been developed for melanoma, BCC and cSCC, and investigated with respect to clinical risk prediction.

### What does this study add?

- This review provides an overview of GWAS findings and the potential clinical utility of PRS for melanoma, BCC and cSCC.

Genetic risk factors have been reported for melanoma, cutaneous squamous cell carcinoma (cSCC) and basal cell carcinoma (BCC).

### Genetic association studies

The role of genetic variation in the aetiology and pathogenesis of cancer and other complex chronic diseases has been investigated using population- and family-based approaches. Population-based designs investigate genotype–phenotype associations in unrelated individuals using a candidate-gene or genome-wide approach, whereas family-based designs examine related individuals using linkage analysis.<sup>1–3</sup>

Most genetic association studies to date have focused on common single nucleotide variants/polymorphisms (SNVs/SNPs), or single base-pair changes in the genome occurring with a minor allele frequency of > 1%.<sup>4</sup> Genome-wide association studies (GWAS) test variants spanning the genome for correlation with disease status. This agnostic approach is well suited to investigating common variants in polygenic diseases and permits examination of genes not previously known to be associated with a phenotype. Furthermore, GWAS have greater statistical power than single gene approaches to detect small or modest effect sizes.<sup>5</sup> As genomic technology has advanced and costs have decreased, GWAS have become a popular and efficient way to study common genetic traits.<sup>6–9</sup>

### Penetrance

Disease-associated genetic variants display variable levels of penetrance, defined as the proportion of people in a population who carry the disease-causing variant and who are also affected by the disease. With complete penetrance, all individuals harbouring the pathogenic variant develop the disease. Disorders caused by rare, highly penetrant variants typically display Mendelian patterns of inheritance, in which a single genetic locus is responsible for the disease. Most hereditary cancer disorders show incomplete penetrance and require other genetic or environmental factors for disease development.<sup>10</sup> Family-based linkage studies are well suited to studying these rare, highly penetrant alleles. In contrast, the majority of the genetic risk for complex non-Mendelian phenotypes, such as cancer, is thought to be due to lower-penetrance common variants with small effect sizes that act in concert to influence aetiology.<sup>11</sup>

### Polygenic risk scores and population-attributable risk

Genetic risk factors are potentially useful tools for preventive medicine.<sup>12</sup> However, effect sizes for any given disease-causing variants are typically small, so while any particular variant will not significantly impact risk prediction, an individual's combination of variants may be informative. Additive effects can be investigated by summing risk alleles and weighting by their effect size to create a polygenic risk score (PRS). Different approaches are used to construct PRS. One common approach incorporates variants reaching genome-wide

significance (defined as  $P < 5 \times 10^{-8}$ ). Other methods combine variants across the genome, using pruning and thresholding techniques. Scores resulting from this latter methodology are sometimes referred to as 'genetic risk scores'. Here, we use PRS, but acknowledge that the literature we review uses varying statistical methods for score construction.

Effect sizes of genetic variants can also be used to calculate the population-attributable risk (PAR), which measures the incidence reduction occurring if the causal factor(s) are eliminated in a population. Here, the PAR estimates the extent to which a disease is due to genetics. Although genetic variants cannot be eliminated, prevention strategies targeting their pathways may be possible. The PAR calculation has several known limitations, including the potential for inflated estimates, which have previously been described in detail.<sup>13</sup>

The use of PRS for risk stratification and prediction is well established.<sup>12,14–16</sup> The utility of a predictor is evaluated by estimating the area under the receiver-operating characteristic (ROC) curve (AUC), which graphically displays the true-positive rate (sensitivity) against the false-positive rate ( $1 - \text{specificity}$ ) for all possible cutpoints of a continuous predictor (Fig. 1). An AUC of 0.5 indicates that the predictor has no discriminatory ability (Fig. 1a). Curves that align more closely with the upper left corner of the ROC plot have higher AUC, maximizing the true-positive and true-negative fractions (Fig. 1b).

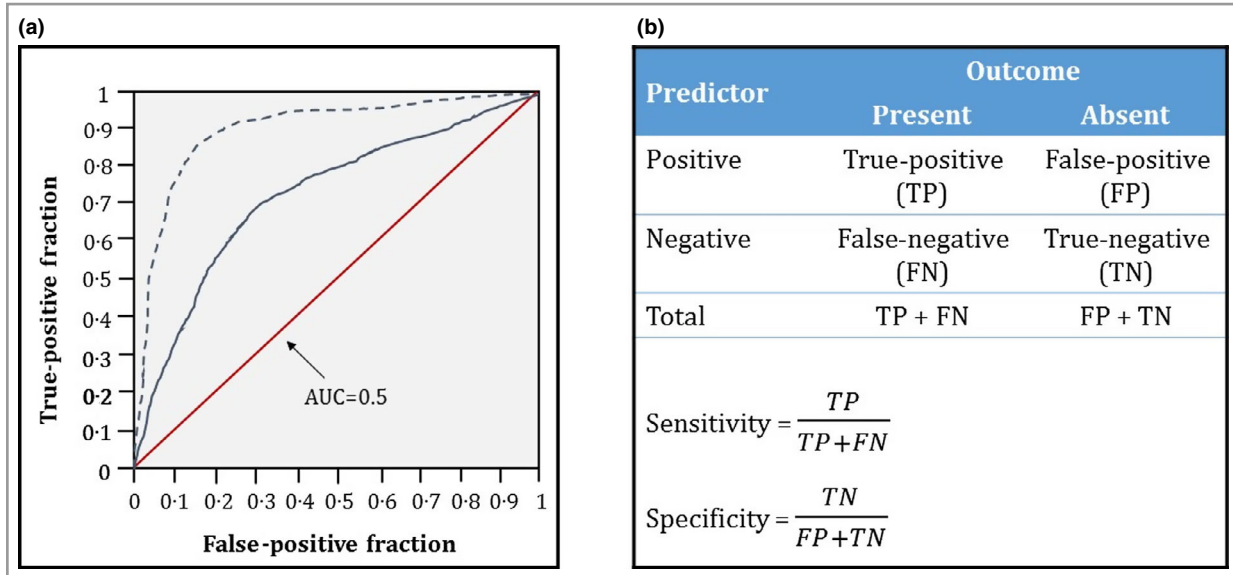
PRS have been employed across a wide variety of diseases for clinical applications including screening, diagnosis, prognosis and treatment response. Translation of PRS to the clinical setting is being facilitated by large biobanking efforts and large population-based genotyping,<sup>17–20</sup> and PRS for breast and prostate cancer have demonstrated clinical utility. While PRS are not currently clinically available for cutaneous carcinomas, research has significantly increased our understanding of genetic susceptibility to cutaneous cancers.

### Objectives

We aim to provide an overview of PRS for melanoma, cSCC and BCC. We summarize GWAS findings and studies examining PRS and PAR for cutaneous malignancies. We also highlight the potential clinical utility of adding PRS to existing prediction models for these malignancies and discuss future research directions.

### Materials and methods

GWAS and PRS/PAR studies were identified using PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>) and the following search criteria: "genome wide association study" AND ["melanoma" or "cutaneous squamous cell carcinoma" or "basal cell carcinoma"]; "polygenic risk score" AND ["melanoma" or "cutaneous squamous cell carcinoma" or "basal cell carcinoma"]; "genetic risk score" AND ["melanoma" or "cutaneous squamous cell carcinoma" or "basal cell carcinoma"]; "population attributable risk" AND ["melanoma" or "cutaneous squamous cell carcinoma" or "basal cell carcinoma"]. We



**Fig 1.** Sensitivity, specificity and receiver-operating characteristic (ROC) curves. (a) Examples of ROCs. The red line demonstrates a ROC with an area under the curve (AUC) of 0.5, indicating a predictor with no ability to discriminate cases from controls. The solid and dashed blue lines indicate increased AUC, and thus greater predictive ability. (b) For a given dichotomous outcome and predictor, the proportions of true-positive, false-positive, true-negative and false-negative cases may be defined. Sensitivity is defined as the true-positive fraction, or the proportion of cases correctly identified by the test (predictor) as having the outcome. Specificity is defined as the true-negative fraction, or the proportion of cases correctly identified by the test as not having the outcome.

also searched the National Human Genome Research Institute-European Bioinformatics Institute GWAS catalogue (<https://www.ebi.ac.uk/gwas/>) for relevant studies, using the terms “melanoma”, “cutaneous squamous cell carcinoma” and “basal cell carcinoma”. Separate searches were conducted for each phenotype. References for each identified study were searched to identify any additional citations not originally captured. We did not identify any unpublished or non-English language studies, and studies from any time period were included. A dermatology/epidemiology fellow (M.R.R.) identified relevant papers.

## Results

### Skin cancer genome-wide association studies

Numerous GWAS have revealed genetic variants associated with susceptibility for melanoma, cSCC and BCC (Fig. 2). Importantly, many of these variants do not fall within coding regions of the genes, and their function is mainly uncharacterized.

#### Melanoma

The earliest GWAS of melanoma risk identified susceptibility loci at 20q11.22, containing *MTH7B* and *PIGU* and 9p21, a region adjacent to *MTAP* and flanking the familial melanoma susceptibility locus *CDKN2A*.<sup>21,22</sup> Subsequently, more than 20 additional loci have been identified, including skin pigmentation genes (*MC1R*, *OCA2*, *ASIP/RALY*, *TYR*, *IRF/EXOC2* and *SLC45A2*), and loci containing genes for DNA repair (*PARP1*),

epidermal development (*CASP8*), telomere maintenance (*TERT/CLMPT1L* and *OFBC1*) and cell-cycle progression (*CCND1*, *CDK10* and *ATM*).<sup>23–31</sup>

#### Squamous cell carcinoma

Three cSCC GWAS have been published, identifying skin pigmentation genes, including 5p13 (*SLC45A2*), 6p25 (*IRF4*), 9p22 (*BNC2/CNTLN*, a putative pigmentation locus), 11q14 (*TYR*), 15q13 (*HERC2/OCA2*), 16q24 (*DEF8/MC1R*) and 20q11 (*ASIP/RALY*). Other notable loci include variants near human leucocyte antigen (HLA) class II genes (6p21, *HLA-DQA1*) and 11q23.3, containing the metastasis suppressor gene *CADM1*.<sup>32–34</sup>

#### Basal cell carcinoma

The first GWAS (2008) identified signals at 1p36 (containing *PADI4*, *PADI6*, *RCC2* and *ARHGEF10L*) and 1q42 (the nearest gene is *ras-homolog RHOU*).<sup>35</sup> Since then, other studies have identified susceptibility loci within genes involved in skin pigmentation, immune function (*HLA-DQA2*, *HLA-B*, *LPP*, *NEU1*, *ZBTB10* and *TICAM/PLIN3*) and telomere maintenance (*TERT/CLPTM1L* and *OFBC1*). Tumour progression loci include tumour suppressor gene *TP53*, *MYCN* oncogene, p13 (transcription factor *FOXP1*), 7q22.1 (*CUX1*, involved in cellular proliferation and differentiation), 7q12.3 (*TNS3*) and 6q27 (*MIR3939*). Susceptibility loci within epidermal development genes include type 3 transglutaminase (*TGM3*), which is indispensable for normal epidermal formation,<sup>36</sup> keratin 5 (*KRT5*), 10p14 (*GATA3*) and 2q33 (*CASP8/ALSC12*).<sup>37–43</sup>

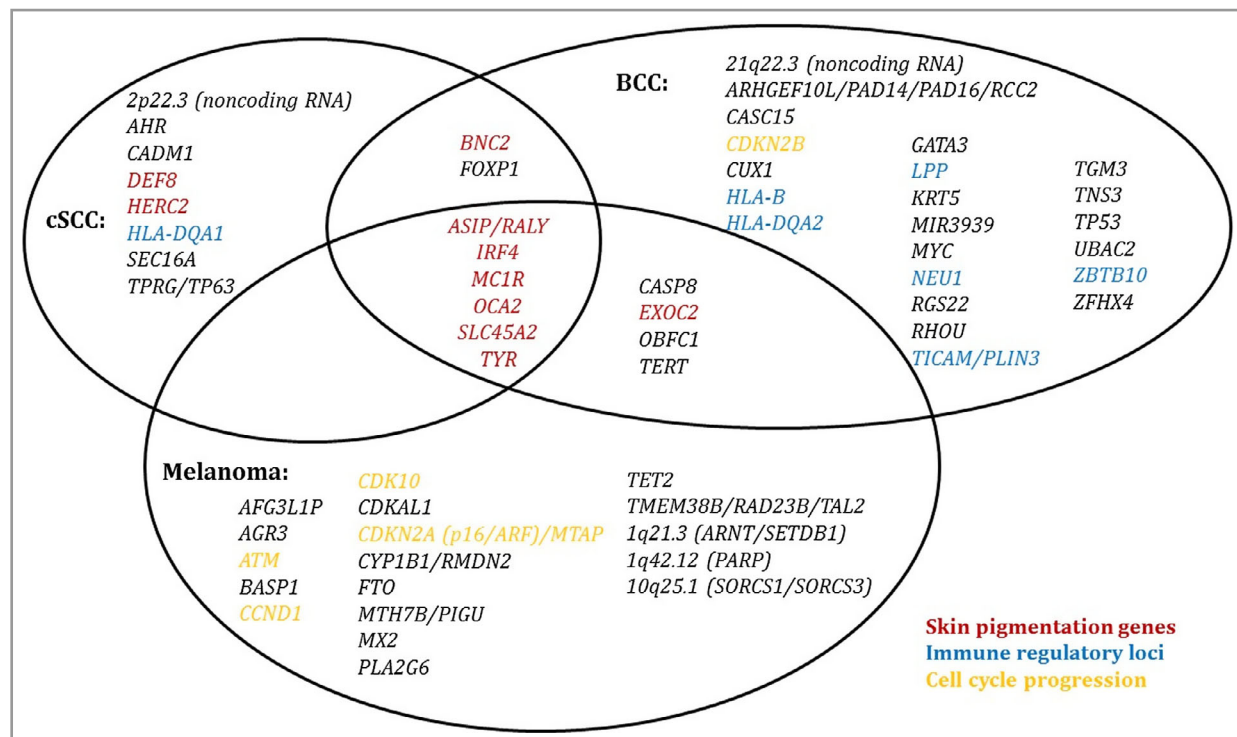


Fig 2. Susceptibility loci identified in published genome-wide association studies. The nearest genes to significant loci for melanoma, cutaneous squamous cell carcinoma (cSCC) and basal cell carcinoma (BCC) are shown. Overlapping regions indicate loci identified in more than one tumour type. Loci involved in key pathways are highlighted as follows: skin pigmentation, red; immune regulation, blue; cell-cycle progression, gold. For cSCC and BCC, loci at 2p22.3 (cSCC) and 21q22.3 (BCC) include long noncoding RNAs. For melanoma, 1q21.3 is a region spanning 10 genes, including ARNT/SETDB1 and 1q42.12 is a region including the DNA repair gene PARP. The loci at 10q25.1 lies between the SORCS1 and SORCS3 genes, which are involved in vacuolar protein production.

### Shared genetic pathways

Loci containing important pigimentary genes have been identified in all three malignancies (Fig. 2). Loci containing immune regulatory genes have also been identified for cSCC and BCC, although the specific loci do not appear to overlap between these tumour types. Similarly, loci containing cell-cycle progression genes have been identified in BCC and melanoma, but the specific loci are unique to each cancer. Other loci show some overlap between cSCC and BCC (Fig. 2), but no additional shared loci have been identified between cSCC and melanoma.

### Skin cancer polygenic risk score and population attributable risk

PRS and PAR have been used to examine the genetic aetiology of skin cancers (Tables 1 and 2).

#### Melanoma

For melanoma, PRS risk estimates range from 1.34 to 3.22. The addition of PRS to clinical risk-factor models has resulted in modest improvements to predictive value.

Using data from the Michigan Genomics Initiative, investigators conducted a phenome-wide association study, exploring

associations between PRS and multiple cancer sites. For melanoma, a 16-variant PRS was constructed, and participants in the top quartile were found to have a 2.4-fold increase in melanoma risk compared with those in the bottom quartile, after adjustment for age, sex, genotyping array and ancestry [odds ratio (OR) 2.4, 95% confidence interval (CI) 2.0–2.8].<sup>44</sup> The predictive ability of the PRS was not investigated.

A second population-based study from Greece found a PRS of 26 GWAS-identified SNVs was associated with a modest increase in discriminative ability between models containing only phenotypic risk factors (age, sex, eye/hair/skin colour, phototype and tanning ability) and phenotypic factors plus PRS (phenotypic model: AUC 0.764, 95% CI 0.741–0.787; phenotypic model + PRS: AUC 0.775, 95% CI 0.752–0.797).<sup>45</sup> This study was limited by a relatively small sample size as well as lack of family history and other known melanoma risk factors.

In another study, GWAS and candidate gene results were used to create an 11-SNV PRS. The highest PRS tertile was associated with 69% increased risk of melanoma compared with individuals in the lowest tertile (OR 1.69, 95% CI 1.28–2.25), after adjustment for traditional melanoma risk factors (age, sex, eye/hair/skin colour, tanning ability and family history of melanoma).<sup>46</sup> Incorporation of the PRS into a standard risk-prediction model consisting of age, sex and pigmentation

**Table 1** Associations between polygenic risk scores and skin cancer risk, showing main results from publications examining associations between polygenic risk scores and skin cancers

Study population	Outcome	No. of SNVs in PRS	Covariates	OR (95% CI)	Reference
<b>Melanoma</b>					
Patients treated at MD Anderson Cancer Center	Melanoma risk	11	Age, sex, skin colour, eye colour, hair colour, tanning ability	2.13 (1.69–2.68); highest vs. lowest tertile	46
Nurses' Health Study	Melanoma risk	11	Age, hair colour, tanning ability, family history	1.55 (1.14–2.11); highest vs. lowest tertile	46
Health professionals follow-up study	Melanoma risk	11	Age, hair colour, eye colour, family history	1.34 (0.90–2.00); highest vs. lowest tertile	46
Case-control population from Greece	Melanoma risk	26	Age, sex, eye colour, hair colour, skin colour, phototype, tanning ability	1.88 (1.29–2.74); highest vs. middle quintile 0.73 (0.50–1.05); lowest vs. middle quintile	45
Women's Health Initiative	Lifetime melanoma risk	21	Age	1.91 (1.50–2.42); highest vs. lowest tertile	47
Women's Health Initiative	Incident melanoma risk	21	Age	1.89 (1.42–2.52); highest vs. lowest tertile <sup>a</sup>	47
Australian Melanoma Family Study	Melanoma risk	45	Age, sex, city of recruitment, European ancestry	3.22 (2.30–4.51); highest vs. lowest tertile	48
Case-control population from Leeds, U.K.	Melanoma risk	45	Age, sex, city of recruitment, European ancestry	2.84 (2.14–3.77); highest vs. lowest tertile	48
Michigan Genomics Initiative	Melanoma risk	16	Age, sex, genotyping array, European ancestry	2.4 (2.0–2.8); highest vs. lowest quartile	44
MelaNostrum Consortium	Melanoma risk	204	Age, sex, country, eye colour, hair colour, phototype, number of naevi	1.23 (1.13–1.35), per PRS quintile	49
<b>cSCC</b>					
Michigan Genomics Initiative	cSCC risk	5	Age, sex, genotyping array, European ancestry	2.0 (1.6–2.5); highest vs. lowest quartile	44
Organ transplant recipients	cSCC risk	N/A <sup>b</sup>	Age at time of transplant, era of transplant, recruitment site, European ancestry	1.97 (1.32–2.93); per 1 SD increase in the normalized PRS	51
Organ transplant recipients	Time to cSCC diagnosis	N/A <sup>b</sup>	Age at time of transplant, era of transplant, recruitment site, European ancestry	1.48 (1.25–1.74); per 2 SD increase in the normalized PRS	51
<b>BCC</b>					
Michigan Genomics Initiative	BCC risk	19	Age, sex, genotyping array, European ancestry	2.7 (2.2–3.2); highest vs. lowest quartile	44
Organ transplant recipients	BCC risk	N/A <sup>b</sup>	Age at time of transplant, era of transplant, recruitment site, European ancestry	3.03 (1.78–5.16); per 1 SD increase in the normalized PRS	51
Organ transplant recipients	Time to BCC diagnosis	N/A <sup>b</sup>	Age at time of transplant, era of transplant, recruitment site, European ancestry	1.67 (1.35–2.06); per 1 SD increase in the normalized PRS	51

BCC, basal cell carcinoma; CI, confidence interval; cSCC, cutaneous squamous cell carcinoma; GWAS, genome-wide association study; OR, odds ratio; PRS, polygenic risk score; SNV, single nucleotide variant. <sup>a</sup>Hazard ratio and 95% CI. <sup>b</sup>The number of SNVs in the PRS was not specified in the publication; PRS were constructed using SNVs reaching a predefined P-value threshold from GWAS.

**Table 2** Use of polygenic risk scores (PRS) in risk-prediction models for skin cancer: results from publications examining the ability of PRS to predict skin cancer risk

Study population	Outcome	No. of SNVs in PRS	Model	AUC (95% CI)	Reference
Patients treated at MD Anderson Cancer Center	Melanoma risk	11	PRS alone	0.62 (0.60–0.65)	46
			Risk factors (sex, age, pigmentation)	0.64 (0.61–0.66)	
			Risk factors + PRS	0.69 (0.64–0.69)	
Case-control population from Greece	Melanoma risk	26	Risk factors (sex, age, hair/skin/eye colour, phototype, tanning ability)	0.76 (0.74–0.79)	45
			Risk factors + PRS	0.78 (0.75–0.80)	
Women's Health Initiative <sup>a</sup>	Melanoma risk	21	Age	0.53 (0.50–0.57)	47
			Increase due to addition of PRS	0.07 (0.03–0.10)	
Australian Melanoma Family Study	Melanoma risk	45	Risk factors (sex, age, city of recruitment, European ancestry, hair/skin/eye colour, adult freckling, skin photosensitivity, self-reported naevi, tanning bed use, history of keratinocyte cancer, family history of melanoma, vacation sun exposure, number of blistering sunburns)	0.72 (0.69–0.75)	48
			Risk factors + PRS	0.74 (0.71–0.77)	
			Risk factors (sex, age, city of recruitment, European ancestry, hair/skin/eye colour, adult freckling, skin photosensitivity, self-reported naevi, tanning bed use, history of keratinocyte cancer, family history of melanoma, vacation sun exposure, number of blistering sunburns)	0.65 (0.62–0.68)	
Case-control population from Leeds, U.K.	Melanoma risk	45	Risk factors + PRS	0.68 (0.65–0.71)	49
			Risk factors (age, sex, country, eye colour, hair colour, phototype, number of naevi)	0.801 (0.789–0.813)	
			Risk factors + PRS	0.810 (0.798–0.822)	
MelaNostrum Consortium	Melanoma risk	204	PRS, European ancestry	0.66	51
			Age, era of transplant, site	0.79	
			Age, era of transplant, site, ancestry, PRS	0.81	
Organ transplant recipients <sup>b</sup>	NMSC risk	N/A	Age, history of actinic keratosis, history of in situ or invasive cSCC, tendency to sunburn (based on total count of five sunburn sensitivity loci and a two-locus sunburn sensitivity haplotype) and the total count of risk alleles from 16 loci	Men: 0.856 (0.847–0.864) Women: 0.840 (0.830–0.850)	52
Kaiser Permanente Northern California <sup>c</sup>	3-year risk of cSCC	16	Age, history of actinic keratosis, history of in situ or invasive cSCC, tendency to sunburn (based on total count of five sunburn sensitivity loci and a two-locus sunburn sensitivity haplotype) and the total count of risk alleles from 16 loci	Men: 0.856 (0.847–0.864) Women: 0.840 (0.830–0.850)	52

AUC, area under the curve; CI, confidence interval; cSCC, cutaneous squamous cell carcinoma; NMSC, nonmelanoma skin cancer; PRS, polygenic risk score; SNV, single nucleotide variant. <sup>a</sup>Results for incident melanoma were provided in the publication. Only the incremental increase in AUC from the addition of the PRS was available, not the AUC for the model containing age + PRS. <sup>b</sup>The number of SNVs in the PRS was not specified in the publication; PRS were constructed using SNVs reaching a pre-defined P-value threshold from GWAS. No confidence intervals were available for the AUCs. <sup>c</sup>The model incorporates an unweighted PRS.



marginally improved prediction (AUC increase of 0.03,  $P < 0.001$ ). The authors also explored reclassification of melanoma risk based on known risk factors and the PRS. For participants with predicted risk of 20–50% based on phenotypic risk factors, 17% of cases and 22.5% of controls were reclassified as having predicted risk  $> 50\%$  when the PRS was added.<sup>46</sup>

In the Women's Health Initiative, PRS were calculated based on 21 GWAS SNVs. Postmenopausal women with the highest PRS tertile had increased melanoma prevalence (OR 1.91, 95% CI 1.50–2.42) and incidence (hazard ratio 1.89, 95% CI 1.42–2.52), compared with the lowest tertile. Small improvements in risk prediction were observed [AUC increase of 0.075 (prevalent), 0.068 (incident)].<sup>47</sup>

In Australian and U.K. participants, the highest tertile of a PRS derived from 45 SNVs in 21 loci was associated with increased risk of melanoma (Australia: OR 3.22, 95% CI 2.30–4.51; UK: OR 2.84, 95% CI 2.14–3.77). In both populations, addition of the PRS to a phenotypic risk factor model [age, sex, city, ancestry, eye/hair/skin colour, adult freckling, photosensitivity, self-reported naevi, sunbed use, history of keratinocyte carcinomas (KC), family history of melanoma, sun exposure and blistering sunburns] resulted in an approximately 2–3% AUC increase (Australia: AUC change of 0.023; UK: AUC change of 0.028).<sup>48</sup>

Finally, different methods for creating PRS have been investigated using data from the Melanoma Meta-Analysis Consortium, the largest meta-analysis of melanoma GWAS currently available.<sup>49</sup> Four PRS were constructed, ranging from 18 to 204 SNVs (AUC 0.628–0.644). The best-performing PRS, containing 204 variants, was validated in the MelaNostrum Consortium, a southern European population. When adjusted for age, sex, country, eye/hair colour, phototype and number of naevi, the 204-SNV PRS was associated with a 23% increase in melanoma risk, per PRS quintile (OR 1.23, 95% CI 1.13–1.35). The addition of the PRS to a risk-prediction model containing those covariates resulted in a 0.8% change in the AUC (covariates + PRS: AUC 0.810, 95% CI 0.798–0.822). Based on the variants in the PRS, the authors also calculated a PAR of 26%.

## Squamous cell carcinoma

A recent study, using summary-level GWAS data, estimated a PAR of 62% and showed that the relative risk for cSCC increases with higher percentiles of a 21-SNV polygenic score, with men having higher risk than women at all percentiles.<sup>50</sup> Because this study did not use individual-level data, phenotypic risk factors were not incorporated.

Two studies have investigated PRS using individual-level data, finding approximately twofold increases in cSCC risk.<sup>44,51</sup> One, using data from the Michigan Genomics Initiative biorepository, found that a 5-variant PRS was associated with a twofold increased risk of cSCC, after adjustment for age, sex, genotyping array and ancestry (OR 2.0, 95% CI 1.6–2.5, for the highest vs. lowest quartile of PRS).<sup>44</sup> The second, examining a population of renal transplant recipients,

observed an association of similar magnitude after adjustment for age at transplant, era of transplant, recruitment site and ancestry (OR 1.97, 95% CI 1.32–2.93, per 1 SD increase in the normalized PRS).<sup>51</sup> In this population, the predictive ability of the PRS was also explored, although this analysis grouped all KCs together. A small increase in the AUC was observed when the PRS was added to a model containing age, era of transplant and recruitment site as predictors (model with PRS and ancestry: AUC 0.66; model with age, era of transplant and recruitment site: AUC 0.79; model with age, era of transplant, recruitment site, ancestry and PRS: AUC 0.81).<sup>51</sup> Neither study included known risk factors for cSCC, which may influence predictive power.

A recently developed risk-prediction tool, 'cSCCscore', assigns the probability of developing one or more cSCCs within 3 years, and includes age, tendency to sunburn, history of actinic keratosis and cSCC, and inherited predisposition, based on the total count of risk alleles at 16 GWAS-identified loci.<sup>52</sup> This model resulted in an AUC of 0.84 (95% CI 0.83–0.85) for women and 0.86 (95% CI 0.85–0.86) for men, indicating strong discriminatory ability between those who do and do not develop a new cSCC within 3 years.

## Basal cell carcinoma

A 29-SNV PRS for BCC, created using summary statistics from published GWAS, revealed a higher relative risk for BCC with increasing percentiles of the PRS (unpublished data). A two-fold increased risk of BCC was observed at the 86th percentile of the PRS for men and the 95th percentile for women, indicating that 14% of men and 5% of women can be identified at clinically increased risk of BCC using the PRS. Using five loci, the PAR for BCC was estimated at 45%,<sup>35</sup> and when variants from four additional loci were included (KRT5, 9p21, 7q32, TERT-CLPTMIL), this estimate increased to 74%,<sup>43</sup> suggesting that these additional loci account for a substantial proportion of the PAR.

PRS developed using individual-level data have also been associated with an approximate threefold increase in risk of BCC.<sup>44,51</sup> A 19-variant PRS, developed using data from the Michigan Genomics Initiative, was significantly associated with BCC risk (OR 2.7, 95% CI 2.2–3.2; top vs. bottom quartile, adjusted for age, sex, ancestry and genotyping array).<sup>44</sup> In a study of renal transplant recipients,<sup>51</sup> a threefold increase in risk of BCC was observed after adjustment for age at transplant, era of transplant, recruitment site and ancestry (OR 3.03, 95% CI 1.78–5.16, per 1 SD increase in the normalized PRS). The predictive power of a PRS for BCC risk has not been explored in the general population.

## Discussion

### Nongenetic skin cancer risk-prediction models

Several groups have developed risk-prediction models for melanoma, cSCC and BCC using demographic, lifestyle and

clinical risk factors. For melanoma, models frequently contain age, sex, hair colour, sunburn history, number of naevi, skin type (including skin colour and tanning ability), freckle density, first-degree family history of melanoma, history of KCs and measures of sun/ultraviolet radiation (UVR) exposure as predictors, with varying discriminatory ability (AUC 0.62–0.86).<sup>53–56</sup> Prediction models for KCs, particularly among organ transplant recipients (OTRs), who are at substantially greater risk of developing post-transplant cutaneous carcinomas, have also been evaluated.<sup>57</sup> Predictors include age at transplant, sex, eye colour, skin type, daily UVR exposure, climate, childhood sunburn, and pretransplant KC, Bowen disease or actinic keratosis.<sup>58–60</sup> For KC risk prediction in the general population, a model containing age, sex, ethnicity, skin colour, tanning ability, freckling tendency, number of childhood sunburns (age < 10 years), previously excised or destroyed skin lesions, and smoking status achieved an AUC of 0.80 (95% CI 0.79–0.81).<sup>61</sup>

### Potential clinical use of polygenic risk scores in cutaneous malignancies and future directions

Risk-prediction models based on demographic, lifestyle and clinical variables have resulted in varying degrees of discriminatory power and adding PRS to models containing these features modestly improves predictive value. In the general population, PRS may be used to identify individuals at both ends of the risk spectrum: those who are at higher risk for skin cancer and would benefit from routine screening, and those who are at very low risk, for whom reduced frequency of skin screening may lead to fewer unnecessary procedures and alleviate patient anxiety. Despite the potential for PRS to improve skin cancer risk stratification beyond that currently afforded by nongenetic risk factors, the available data do not yet support their use in routine clinical practice. Given that relatively few studies have investigated PRS, particularly for KCs, more data are needed before conclusions as to the efficacy of PRS can be drawn. If additional studies, ideally incorporating both genetic and nongenetic risk factors, demonstrate substantial increases in the AUC, a stronger argument for the use of PRS in clinical practice may be made. This is supported by the demonstrated clinical utility of PRS for breast and prostate cancers, for which commercial panels are now available. Clinical trials to evaluate the ability of PRS to inform breast cancer screening recommendations are currently underway.<sup>62,63</sup> We propose that similar utility could be gained by adding PRS to risk-prediction methods for cutaneous malignancies, which may then be used to inform screening recommendations for high- and low-risk individuals. Realizing this potential will require large cohorts with detailed genetic and nongenetic risk factor data for model development and validation.

One area where PRS may show particular value is in improving risk stratification for OTRs. Fitzpatrick skin phototype and polymorphisms in the *OCA2/HERC2* and *IRF4* pigmentation genes have been associated with development of post-transplant cSCC.<sup>64–66</sup> Additional research is needed to determine whether PRS can improve risk prediction above

established risk factors. PRS may also explain some of the heritability in families containing multiple skin cancer diagnoses and aid in pathological diagnoses. Evidence suggests that histological factors may affect tumour aggressiveness and patient prognosis in melanoma and KCs.<sup>67,68</sup> One study, for example, found that two SNVs in *IRF4* and *PLA2G6* were associated with *BRAF/NRAS* melanoma subtypes.<sup>69</sup> However, PRS studies to date have combined histological subtypes together, which could lead to confounding. Additional studies are needed to clarify the relationships between PRS and histological subtypes in cutaneous malignancies.

Although current PRS studies have identified only modest gains in predictive ability, even small improvements may be informative, and incorporation of additional variants and exposure variables may further refine risk models. Model optimization will require additional validation studies in large cohorts containing nongenetic risk factors. As GWAS for cutaneous carcinomas have been performed exclusively on European populations, where incidence is highest, PRS using these data may not be generalizable to other populations. Investigations including other races/ethnicities are needed to maximize clinical benefit.

Finally, while available research indicates that knowledge of one's genomic susceptibility does not influence preventative behaviour,<sup>70</sup> skin cancer prevention has not specifically been examined. Additional studies are needed to evaluate how knowledge of genetic risk may affect skin cancer prevention, as well as how to effectively communicate this information to patients.<sup>71,72</sup>

### Conclusion

Our understanding of the genetic architecture of cutaneous malignancies has increased substantially, although much work remains before clinical implementation of PRS becomes viable for the field of dermatology. Large, well-powered studies comprehensively evaluating PRS in conjunction with known phenotypic and clinical risk factors will be needed to determine clinical value.

### References

- Lewis CM. Genetic association studies: design, analysis and interpretation. *Brief Bioinform* 2002; **3**:146–53.
- Dueker ND, Pericak-Vance MA. Analysis of genetic linkage data for Mendelian traits. *Curr Protoc Hum Genet* 2014; **83**:1.4.1–31.
- Bailey-Wilson JE, Wilson AF. Linkage analysis in the next-generation sequencing era. *Hum Hered* 2011; **72**:228–36.
- Brookes AJ. The essence of SNPs. *Gene* 1999; **234**:177–86.
- Witte JS. Genome-wide association studies and beyond. *Annu Rev Public Health* 2010; **31**:9–20.
- Zondervan KT, Cardon LR. Designing candidate gene and genome-wide case-control association studies. *Nat Protoc* 2007; **2**:2492–501.
- Anderson CA, Pettersson FH, Clarke GM *et al.* Data quality control in genetic case-control association studies. *Nat Protoc* 2010; **5**:1564–73.
- De R, Bush WS, Moore JH. Bioinformatics challenges in genome-wide association studies (GWAS). *Methods Mol Biol* 2014; **1168**:63–81.



- 9 McCarthy MI, Abecasis GR, Cardon LR *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**:356–69.
- 10 Pomerantz MM, Freedman ML. The genetics of cancer risk. *Cancer J* 2011; **17**:416–22.
- 11 Bush WS, Moore JH. Chapter 11: Genome-wide association studies. *PLoS Comput Biol* 2012; **8**:e1002822.
- 12 Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* 2018; **19**:581–90.
- 13 Witte JS, Visscher PM, Wray NR. The contribution of genetic variants to disease depends on the ruler. *Nat Rev Genet* 2014; **15**:765–76.
- 14 Lewis CM, Vassos E. Prospects for using risk scores in polygenic medicine. *Genome Med* 2017; **9**:96.
- 15 Maher BS. Polygenic scores in epidemiology: risk prediction, etiology, and clinical utility. *Curr Epidemiol Rep* 2015; **2**:239–44.
- 16 Kraft P, Wacholder S, Cornelis MC *et al.* Beyond odds ratios: communicating disease risk based on genetic profiles. *Nat Rev Genet* 2009; **10**:264–9.
- 17 Carey DJ, Fetterolf SN, Davis FD *et al.* The Geisinger MyCode community health initiative: an electronic health record-linked biobank for precision medicine research. *Genet Med* 2016; **18**:906–13.
- 18 Buchanan AH, Manickam K, Meyer MN *et al.* Early cancer diagnoses through BRCA1/2 screening of unselected adult biobank participants. *Genet Med* 2018; **20**:554–8.
- 19 Kincaid E. Geisinger says DNA sequencing as preventative care is ready for the clinic. *Forbes* (7 May 2018). Available at: <https://www.forbes.com/sites/elliekincaid/2018/05/07/geisinger-says-dna-sequencing-as-preventative-care-is-ready-for-the-clinic/#af0dd6b4-d636> (accessed 16 August 2018).
- 20 Geisinger News Release. DNA sequencing to become part of Geisinger's routine clinical care. (7 May 2018). Available at: <https://www.geisinger.org/about-geisinger/news-and-media/news-releases/2018/05/07/12/18/dna-sequencing-to-become-part-of-geisingers-routine-clinical-care> (accessed 16 August 2018).
- 21 Brown KM, Macgregor S, Montgomery GW *et al.* Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 2008; **40**:838–40.
- 22 Bishop DT, Demenais F, Iles MM *et al.* Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 2009; **41**:920–5.
- 23 Amos CI, Wang L-E, Lee JE *et al.* Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet* 2011; **20**:5012–23.
- 24 Macgregor S, Montgomery GW, Liu JZ *et al.* Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat Genet* 2011; **43**:1114–18.
- 25 Iles MM, Law MH, Stacey SN *et al.* A variant in FTO shows association with melanoma risk not due to BMI. *Nat Genet* 2013; **45**:428–32, 432e1.
- 26 Law MH, Bishop DT, Lee JE *et al.* Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nat Genet* 2015; **47**:987–95.
- 27 Barrett JH, Iles MM, Harland M *et al.* Genome-wide association study identifies three new melanoma susceptibility loci. *Nat Genet* 2011; **43**:1108–13.
- 28 Ransohoff KJ, Wu W, Cho HG *et al.* Two-stage genome-wide association study identifies a novel susceptibility locus associated with melanoma. *Oncotarget* 2017; **8**:17586–92.
- 29 Antonopoulou K, Stefanaki I, Lill CM *et al.* Updated field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma: the MelGene database. *J Invest Dermatol* 2015; **135**:1074–9.
- 30 Teerlink C, Farnham J, Allen-Brady K *et al.* A unique genome-wide association analysis in extended Utah high-risk pedigrees identifies a novel melanoma risk variant on chromosome arm 10q. *Hum Genet* 2012; **131**:77–85.
- 31 Song F, Amos CI, Lee JE *et al.* Identification of a melanoma susceptibility locus and somatic mutation in TET2. *Carcinogenesis* 2014; **35**:2097–101.
- 32 Asgari MM, Wang W, Ioannidis NM *et al.* Identification of susceptibility loci for cutaneous squamous cell carcinoma. *J Invest Dermatol* 2016; **136**:930–7.
- 33 Siiskonen SJ, Zhang M, Li W-Q *et al.* A genome-wide association study of cutaneous squamous cell carcinoma among European descendants. *Cancer Epidemiol Biomarkers Prev* 2016; **25**:714–20.
- 34 Chahal HS, Lin Y, Ransohoff KJ *et al.* Genome-wide association study identifies novel susceptibility loci for cutaneous squamous cell carcinoma. *Nat Commun* 2016; **7**:12048.
- 35 Stacey SN, Gudbjartsson DF, Sulem P *et al.* Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. *Nat Genet* 2008; **40**:1313–18.
- 36 Matsuki M, Yamashita F, Ishida-Yamamoto A *et al.* Defective stratum corneum and early neonatal death in mice lacking the gene for transglutaminase 1 (keratinocyte transglutaminase). *Proc Natl Acad Sci U S A* 1998; **95**:1044–9.
- 37 Chahal HS, Wu W, Ransohoff KJ *et al.* Genome-wide association study identifies 14 novel risk alleles associated with basal cell carcinoma. *Nat Commun* 2016; **7**:12510.
- 38 Stacey SN, Helgason H, Gudjonsson SA *et al.* New basal cell carcinoma susceptibility loci. *Nat Commun* 2015; **6**:6825.
- 39 Stacey SN, Sulem P, Gudbjartsson DF *et al.* Germline sequence variants in TGM3 and RGS22 confer risk of basal cell carcinoma. *Hum Mol Genet* 2014; **23**:3045–53.
- 40 Nan H, Xu M, Kraft P *et al.* Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011; **20**:3718–24.
- 41 Stacey SN, Sulem P, Jonasdottir A *et al.* A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat Genet* 2011; **43**:1098–103.
- 42 Rafnar T, Sulem P, Stacey SN *et al.* Sequence variants at the TERT-CLPTMIL locus associate with many cancer types. *Nat Genet* 2009; **41**:221–7.
- 43 Stacey SN, Sulem P, Masson G *et al.* New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**:909.
- 44 Fritsche LG, Gruber SB, Wu Z *et al.* Association of polygenic risk scores for multiple cancers in a phenome-wide study: results from the Michigan Genomics Initiative. *Am J Hum Genet* 2018; **102**:1048–61.
- 45 Kypreou KP, Stefanaki I, Antonopoulou K *et al.* Prediction of melanoma risk in a southern European population based on a weighted genetic risk score. *J Invest Dermatol* 2016; **136**:690–5.
- 46 Fang S, Han J, Zhang M *et al.* Joint effect of multiple common SNPs predicts melanoma susceptibility. *PLoS One* 2013; **8**:e85642.
- 47 Cho HG, Ransohoff KJ, Yang L *et al.* Melanoma risk prediction using a multilocus genetic risk score in the Women's Health Initiative cohort. *J Am Acad Dermatol* 2018; **79**:36–41.e10.
- 48 Cust AE, Drummond M, Kanetsky PA *et al.* Assessing the incremental contribution of common genomic variants to melanoma risk prediction in two population-based studies. *J Invest Dermatol* 2018; **138**:2617–24.
- 49 Gu F, Chen T-H, Pfeiffer RM *et al.* Combining common genetic variants and non-genetic risk factors to predict risk of cutaneous melanoma. *Hum Mol Genet* 2018; **27**:4145–56.

- 50 Sordillo JE, Kraft P, Wu AC, Asgari MM. Quantifying the polygenic contribution to cutaneous squamous cell carcinoma risk. *J Invest Dermatol* 2018; **138**:1507–10.
- 51 Stapleton CP, Birdwell KA, McKnight AJ *et al.* Polygenic risk score as a determinant of risk of non-melanoma skin cancer in a European-descent renal transplant cohort. *Am J Transplant* 2019; **19**:801–10.
- 52 Wang W, Jorgenson E, Ioannidis NM *et al.* A prediction tool to facilitate risk-stratified screening for squamous cell skin cancer. *J Invest Dermatol* 2018; **138**:2589–94.
- 53 Vuong K, Armstrong BK, Weiderpass E *et al.* Development and external validation of a melanoma risk prediction model based on self-assessed risk factors. *JAMA Dermatol* 2016; **152**:889.
- 54 Vuong K, McGeechan K, Armstrong BK, Cust AE. Risk prediction models for incident primary cutaneous melanoma. *JAMA Dermatol* 2014; **150**:434.
- 55 Olsen CM, Pandeya N, Thompson BS *et al.*; QSkin Study. Risk stratification for melanoma: models derived and validated in a purpose-designed prospective cohort. *J Natl Cancer Inst* 2018; **110**:1075–83.
- 56 Usher-Smith JA, Emery J, Kassianos AP, Walter FM. Risk prediction models for melanoma: a systematic review. *Cancer Epidemiol Biomarkers Prev* 2014; **23**:1450–63.
- 57 Lowenstein SE, Garrett G, Toland AE *et al.* Risk prediction tools for keratinocyte carcinoma after solid organ transplantation: a review of the literature. *Br J Dermatol* 2017; **177**:1202–7.
- 58 Harden PN, Fryer AA, Reece S *et al.* Annual incidence and predicted risk of nonmelanoma skin cancer in renal transplant recipients. *Transplant Proc* 2001; **33**:1302–4.
- 59 Urwin HR, Jones PW, Harden PN *et al.* Predicting risk of non-melanoma skin cancer and premalignant skin lesions in renal transplant recipients. *Transplantation* 2009; **87**:1667–71.
- 60 Carroll RP, Ramsay HM, Fryer AA *et al.* Incidence and prediction of nonmelanoma skin cancer post-renal transplantation: a prospective study in Queensland, Australia. *Am J Kidney Dis* 2003; **41**:676–83.
- 61 Whiteman DC, Thompson BS, Thrift AP *et al.* A model to predict the risk of keratinocyte carcinomas. *J Invest Dermatol* 2016; **136**:1247–54.
- 62 Shieh Y, Eklund M, Madlensky L *et al.* Breast cancer screening in the precision medicine era: risk-based screening in a population-based trial. *J Natl Cancer Inst* 2017; **109**:djw290.
- 63 Esserman LJ; WISDOM Study and Athena Investigators. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer* 2017; **3**:34.
- 64 Wei L, Allain DC, Bernhardt MN *et al.* Variants at the OCA2/HERC2 locus affect time to first cutaneous squamous cell carcinoma in solid organ transplant recipients collected using two different study designs. *Br J Dermatol* 2017; **177**:1066–73.
- 65 Asgari MM, Toland AE, Arron ST. IRF4 polymorphism is associated with cutaneous squamous cell carcinoma in organ transplant recipients: a pigment-independent phenomenon. *J Invest Dermatol* 2017; **137**:251–3.
- 66 Gogia R, Binstock M, Hirose R *et al.* Fitzpatrick skin phototype is an independent predictor of squamous cell carcinoma risk after solid organ transplantation. *J Am Acad Dermatol* 2013; **68**:585–91.
- 67 Walling HW, Fosko SW, Geraminejad PA *et al.* Aggressive basal cell carcinoma: presentation, pathogenesis, and management. *Cancer Metastasis Rev* 2004; **23**:389–402.
- 68 Lattanzi M, Lee Y, Simpson D *et al.* Primary melanoma histologic subtype: impact on survival and response to therapy. *J Natl Cancer Inst* 2019; **111**:180–8.
- 69 Thomas NE, Edmiston SN, Orlov I *et al.*; GEM Study Group. Inherited genetic variants associated with melanoma BRAF/NRAS subtypes. *J Invest Dermatol* 2018; **138**:2398–2404.
- 70 Hollands GJ, French DP, Griffin SJ *et al.* The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis. *BMJ* 2016; **352**:i1102.
- 71 Lautenbach DM, Christensen KD, Sparks JA, Green RC. Communicating genetic risk information for common disorders in the era of genomic medicine. *Annu Rev Genomics Hum Genet* 2013; **14**:491–513.
- 72 Lea DH, Kaphingst KA, Bowen D *et al.* Communicating genetic and genomic information: health literacy and numeracy considerations. *Public Health Genomics* 2011; **14**:279–89.