Assessing the Incremental Contribution of Common Genomic Variants to Melanoma Risk Prediction in Two Population-Based Studies

Anne E. Cust¹,², Martin Drummond¹,², Peter A. Kanetsky³, Australian Melanoma Family Study Investigators⁴, Leeds Case-Control Study Investigators⁵, Alisa M. Goldstein⁴, Jennifer H. Barrett⁵, Stuart MacGregor⁶, Matthew H. Law⁶, Mark M. Iles⁵, Minh Bui⁵, John L. Hopper⁷, Myriam Brossard⁸,⁹, Florence Demenais⁸,⁹, John C. Taylor³, Clive Hoggart¹⁰, Kevin M. Brown⁴, Maria Teresa Landi⁴, Julia A. Newton-Bishop⁵, Graham J. Mann²,¹¹ and D. Timothy Bishop⁵

It is unclear to what degree genomic and traditional (phenotypic and environmental) risk factors overlap in their prediction of melanoma risk. We evaluated the incremental contribution of common genomic variants (in pigmentation, nevus, and other pathways) and their overlap with traditional risk factors, using data from two population-based case-control studies from Australia (n = 1,035) and the United Kingdom (n = 1,460) that used the same questionnaires. Polygenic risk scores were derived from 21 gene regions associated with melanoma and odds ratios from published meta-analyses. Logistic regression models were adjusted for age, sex, center, and ancestry. Adding the polygenic risk score to a model with traditional risk factors increased the area under the receiver operating characteristic curve (AUC) by 2.3% (P = 0.003) for Australia and by 2.8% (P = 0.002) for Leeds. Gene variants in the pigmentation pathway, particularly MC1R, were responsible for most of the incremental improvement. In a cross-tabulation of polygenic by traditional tertile risk scores, 59% (Australia) and 49% (Leeds) of participants were categorized in the same (concordant) tertile. Of participants with low traditional risk, 9% (Australia) and 21% (Leeds) had high polygenic risk. Testing of genomic variants can identify people who are susceptible to melanoma despite not having a traditional phenotypic risk profile.

Journal of Investigative Dermatology (2018) 138, 2617–2624; doi:10.1016/j.jid.2018.05.023

INTRODUCTION

Primary and secondary prevention strategies that reduce sun exposure and encourage increased sun protection and skin examination behaviors are important for reducing melanoma incidence and mortality, particularly for those identified as being at high risk (Aitken et al., 2010; Armstrong and Kricker, 1993; Breitbart et al., 2012; Glanz et al., 2015). To date, identifying people at high risk for melanoma has focused on factors such as family history (Olsen et al., 2010b), phenotypic characteristics (fair skin, skin that burns easily, red hair, moles and freckling (Gandini et al., 2005a), solar and artificial UV radiation exposure (Cust et al., 2011a; Gandini et al., 2005b), and previous keratinocyte skin cancers (Gandini et al., 2005b).

Common genomic variants may help identify people at high risk for melanoma, particularly those who lack these...
traditional risk factors (Berwick et al., 2014; Cust et al., 2012; Kanetsky et al., 2010). It is becoming more feasible to incorporate genomic profiles into risk prediction tools, given increased understanding of genomic risk factors and increased genomic testing in clinical practice. Development and assessment of melanoma risk prediction models that combine traditional and genomic risk factors is warranted to help improve melanoma prevention and population screening. For example, Australian general practice guidelines state that clinical assessment of melanoma risk should take into account multiple risk factors yet highlight the fact that there are no sufficiently well-validated risk models to assess the combined effects of all risk factors (Royal Australian College of General Practitioners, 2016). Accurate identification of people at high risk for melanoma will also facilitate research on targeted screening strategies for those at higher risk (US Preventive Services Task Force et al., 2016).

Few studies have examined the contribution of multiple genomic risk factors to risk prediction over and above that of traditional risk factors (Fang et al., 2013; Kypreou et al., 2016; Stefanaki et al., 2013). Limitations of previous studies have included lack of external validation, small sample size (Stefanaki et al., 2013), lack of data on nevus phenotype or included lack of external validation, small sample size (Stefanaki et al., 2013). Limitations of previous studies have included lack of external validation, small sample size (Stefanaki et al., 2013), lack of data on nevus phenotype or included lack of external validation, small sample size (Stefanaki et al., 2013), and possible confounding by ethnicity (Fang et al., 2013).

We aimed to evaluate the incremental contribution of common genomic variants to melanoma risk prediction, including the contribution of variants associated with identified biological pathways, using data from two population-based studies in geographically disparate but genetically similar populations (Australia and the UK). Both studies were developed by the GenoMEL (www.genomel.org) melanoma genetics consortium and used the same measurement protocols, facilitating external validation.

RESULTS
Participant characteristics
Demographic characteristics of participants in the Australian and Leeds, UK, case-control studies in this analysis are shown in Table 1. The studies had a similar proportion (60%) of female participants but a slightly different mix of European ethnicities; results were similar if restricted to English ethnic background. Participants were younger in the Australian study because it restricted recruitment to onset before age 40 years.

Association of polygenic risk score with melanoma risk
For each country, there was a 3-fold higher risk of melanoma for participants in the highest versus lowest tertile of polygenic risk score and a 6-fold higher risk in the highest versus lowest decile; the odds ratio (OR) per adjusted standard deviation increase in polygenic score (i.e., OPERA) was 1.75 for Australia and 1.63 for Leeds (Table 2). When evaluated by biological pathway (see Supplementary Table S1 online) the ORs for the pigmentation pathway were similar to the overall polygenic risk score, but the ORs were lower for the telomere/senescence/other pathway (about 50% increased risk for the highest vs. lowest tertile) and the nevus pathway (about 25% nonsignificantly increased risk for the highest vs. lowest tertile).

Incremental contribution of polygenic risk score based on published risk estimates
Adding the polygenic risk score based on published ORs to a model with traditional risk factors increased the AUC by 2.3% (P = 0.003) for Australia and by 2.8% (P = 0.002) for Leeds (Table 3). The net reclassification improvement (NRI) was 0.42 (95% confidence interval [CI] = 0.30–0.54) for Australia and 0.29 (95% CI = 0.18–0.39) for Leeds; this was driven more by improvements in specificity (i.e., net movement of control individuals to a lower risk: 29% in Australia and 19% in Leeds) than in sensitivity (net movement of case individuals to a higher risk: 13% in Australia, 9% in Leeds). Single nucleotide polymorphisms (SNPs) in the pigmentation pathway, particularly MC1R, were responsible for most of the incremental improvement. Conversely, SNPs in the nevus and other pathways did not significantly improve risk prediction. Most models were well calibrated except for the addition of the nevus pathway SNPs in Australia.

Secondary analyses
Polygenic risk score and traditional risk factors had similar discrimination when compared side by side in separate models (respectively, 0.71 vs. 0.72 for Australia and 0.64 vs. 0.65 for Leeds). The incremental contribution of the polygenic risk score, including pathway-specific scores, was stronger when the models were based on risk estimates derived from the study datasets (AUC increased by 6.0% for Australia and 5.6% for Leeds), but after 10-fold

| Table 1. Characteristics of case and control individuals in the Australian Melanoma Family Study and Leeds case-control study |
|-----------------------------------------------|
| Characteristic | Australia, n (%) | Leeds, n (%) |
| Total, case and control individuals | 1,035 | 1,460 |
| Case individuals | 578 (56) | 964 (66) |
| Control individuals | 457 (44) | 496 (34) |
| Sex | | |
| Female | 623 (60) | 874 (60) |
| Male | 412 (40) | 586 (40) |
| Age at diagnosis/interview, years | | |
| 18–29 | 268 (26) | 61 (4) |
| 30–39 | 682 (66) | 198 (14) |
| 40–49 | 85 (8) | 264 (18) |
| ≥50–69 | 0 (0) | 937 (64) |
| Ethnic background | | |
| English | 649 (63) | 1,358 (93) |
| Scottish, Irish, Welsh | 52 (5) | 67 (5) |
| Other Northern European | 46 (4) | 6 (0) |
| Southern European | 11 (1) | 6 (0) |
| Eastern European | 238 (23) | 3 (0) |
| Mixed/other European | 39 (4) | 20 (1) |

1Leads case and control individuals were unsolicited for age at diagnosis. In Australia, all case individuals were younger than 40 years at diagnosis, and all population control individuals were younger than 40 years when ascertainment. Case and control individuals could be up to age 44 years at interview for this analysis.

2Self-reported.
were 0.77 (95% CI 0.74–0.75) on internal validation and 0.77 (95% CI = 0.74–0.80) on external validation.

**DISCUSSION**

Our comprehensive assessment of the contribution of common genomic variants to melanoma risk prediction showed that a polygenic risk score is strongly associated with melanoma risk and improved the classification of people at high risk of melanoma beyond that identified from traditional risk factors. The incremental improvement to risk prediction was independent of ambient sun exposure because the increases were similar for the Leeds and Australian studies.

The incremental AUCs (2.3% for Australia, 2.8% for Leeds) and NRIs (0.42 for Australia, 0.29 for Leeds), based on published risk estimates, indicate a moderate improvement to risk prediction. Similar improvements in the AUC have been shown to have independent of ambient sun exposure because the increases were similar for the Leeds and Australian studies.

The incremental AUCs (2.3% for Australia, 2.8% for Leeds) and NRIs (0.42 for Australia, 0.29 for Leeds), based on published risk estimates, indicate a moderate improvement to risk prediction. Similar improvements in the AUC have been shown to have independent of ambient sun exposure because the increases were similar for the Leeds and Australian studies.

The incremental AUCs (2.3% for Australia, 2.8% for Leeds) and NRIs (0.42 for Australia, 0.29 for Leeds), based on published risk estimates, indicate a moderate improvement to risk prediction. Similar improvements in the AUC have been shown to have independent of ambient sun exposure because the increases were similar for the Leeds and Australian studies.

The incremental AUCs (2.3% for Australia, 2.8% for Leeds) and NRIs (0.42 for Australia, 0.29 for Leeds), based on published risk estimates, indicate a moderate improvement to risk prediction. Similar improvements in the AUC have been shown to have independent of ambient sun exposure because the increases were similar for the Leeds and Australian studies.
The observed discriminative ability of the polygenic risk score for melanoma was higher than has been previously reported. A Greek study observed a 1.1% increase in the AUC when adding information from 15 SNPs to a phenotypic risk model (Kypreou et al., 2016). A US study combining three different datasets found a statistically significant 3% increase in the AUC when adding information from 15 SNPs to a phenotypic risk model (Kypreou et al., 2016). A US study combining three different datasets found a statistically significant 3% increase in the AUC when adding information from 15 SNPs to a phenotypic risk model (Kypreou et al., 2016).

| Risk Factor Model | AUC (95% CI) | Change in AUC from Base Model P-Value | Hosmer-Lemeshow P-Value | Improvement in Sensitivity NRI (95% CI) | Improvement in Specificity NRI (95% CI) | Overall Improvement in Classification NRI (95% CI) |
|-------------------|--------------|--------------------------------------|------------------------|----------------------------------------|----------------------------------------|-----------------------------------------------|
| **Australia (N = 1,035)** | | | | | | |
| Base model with traditional risk factors | 0.72 (0.69 to 0.75) | | | | | |
| + MC1R | 0.73 (0.70 to 0.76) | 0.011 | 0.05 | 0.47 | -0.02 (-0.10 to 0.06) | 0.35 (0.26 to 0.43) | 0.33 (0.21 to 0.45) |
| + Pigmentation pathway | 0.74 (0.71 to 0.77) | 0.019 | 0.008 | 0.58 | 0.11 (0.03 to 0.19) | 0.26 (0.18 to 0.35) | 0.38 (0.26 to 0.50) |
| + Nevus pathway | 0.72 (0.69 to 0.75) | 0.001 | 0.54 | <0.01 | -0.02 (-0.11 to 0.06) | 0.07 (-0.02 to 0.16) | 0.04 (-0.08 to 0.17) |
| + Telomere, senescence, and other pathway | 0.72 (0.69 to 0.75) | 0.002 | 0.36 | 0.14 | -0.06 (-0.14 to 0.02) | 0.16 (0.07 to 0.25) | 0.10 (-0.02 to 0.22) |
| + All SNPs | 0.74 (0.71 to 0.77) | 0.023 | 0.003 | 0.23 | 0.13 (0.05 to 0.21) | 0.29 (0.20 to 0.38) | 0.42 (0.30 to 0.54) |
| **Leeds (N = 1,460)** | | | | | | |
| Base model with traditional risk factors | 0.65 (0.62 to 0.68) | | | | | |
| + MC1R | 0.67 (0.64 to 0.70) | 0.014 | 0.02 | 0.34 | -0.10 (-0.16 to -0.03) | 0.27 (0.18 to 0.35) | 0.17 (0.07 to 0.28) |
| + Pigmentation pathway | 0.68 (0.66 to 0.71) | 0.031 | 0.0005 | 0.76 | 0.09 (0.02 to 0.15) | 0.22 (0.13 to 0.30) | 0.30 (0.20 to 0.41) |
| + Nevus pathway | 0.65 (0.62 to 0.68) | 0.000 | 0.98 | 0.81 | -0.03 (-0.10 to 0.03) | 0.06 (-0.03 to 0.14) | 0.02 (-0.08 to 0.13) |
| + Telomere, senescence, and other pathway | 0.66 (0.63 to 0.69) | 0.004 | 0.33 | 0.19 | -0.01 (-0.07 to 0.06) | 0.10 (0.02 to 0.19) | 0.10 (-0.01 to 0.21) |
| + All SNPs | 0.68 (0.65 to 0.71) | 0.028 | 0.002 | 0.10 | 0.09 (0.03 to 0.16) | 0.19 (0.11 to 0.28) | 0.29 (0.18 to 0.39) |

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; NRI, net reclassification improvement; SNP, single nucleotide polymorphism.

1. Chi-square P-value for the difference in the AUC when compared with the base model.
2. Based on continuous NRI. Improvement in sensitivity is calculated from reclassification of case individuals improvement in specificity is calculated from reclassification of control individuals, and overall improvement combines the improvements in sensitivity and specificity.
3. Traditional factors include hair color, skin color, eye color, freckling as an adult, skin photosensitivity, self-reported nevi, sunbed use, keratinocyte cancer personal history, first degree family history of melanoma, vacation sun exposure, and blistering sunburns as a child, as well as the demographic and study design factors of age, sex, city of recruitment, and European ancestry.
4. Added as a polygenic risk score, comprising 45 SNPs in 21 genes. The SNPs in each pathway can overlap; the pigmentation pathway includes 14 genes (31 SNPs); nevus pathway includes 7 genes (13 SNPs); and telomere, senescence, and other pathways includes 5 genes (9 SNPs).
Discrimination may improve with the discovery of new nevus-related genes (Duffy et al., 2017), but nevus phenotype measurement will likely continue to be important for risk prediction, perhaps because it reflects early-life sun exposure and genetic susceptibility (Bataille et al., 2000). SNPs from genes involved in other pathways unrelated to pigmentation and nevi, such as telomere length or senescence, were significantly associated with melanoma risk and made a modest contribution to risk prediction.

The more parsimonious risk prediction models that we developed showed good discrimination, because the AUCs remained at 0.72 or greater after internal and external validation. Not all SNPs and traditional risk factors were significantly associated with melanoma in the Australian and Leeds datasets, and the specific risk factors selected for the model differed between countries; this is probably a reflection of the studies’ sample sizes.

Although our cross-tabulation of polygenic with traditional risk score tertiles estimated that 59% of people in Australia and 49% in Leeds have a genetic risk concordant with their traditionally estimated risk, a considerable proportion (9% and 21%, respectively) had a high polygenic risk despite having a low traditional risk. These people might be the most likely to benefit from genomic profiling given that they do not have the visible risk factors identified in public health campaigns. Conversely, a similar proportion had a high traditional risk but a low polygenic risk; knowing they have a low genetic susceptibility might worsen their sun-related behaviors. Our stratified analyses suggest that the improvement in discrimination might be better for those with a low or average traditional risk score. Studies are underway to evaluate the impact on sun-related behaviors of giving personalized melanoma genomic risk information (Kanetsky and Hay, 2017; Smit et al., 2018).

There are several strengths of our analysis, including the population-based design, external validation in independent datasets using the same questionnaires, relatively large

### Table 4. Cross-tabulation of polygenic risk score versus traditional risk score categorized in tertiles

| Traditional Risk Score | Polygenic Risk Score, n (%) | Australia | Leeds |
|------------------------|-----------------------------|----------|-------|
| Tertile 1 (Lower Risk) | Tertile 2 (Average Risk) | Tertile 3 (Higher Risk) | Total |
| Tertile 1 (lower risk) | 223 (65) | 91 (26) | 30 (9) | 344 (100) |
| Tertile 2 (average risk) | 94 (27) | 160 (46) | 91 (26) | 345 (100) |
| Tertile 3 (higher risk) | 27 (8) | 94 (27) | 225 (65) | 346 (100) |
| Total | 344 | 345 | 346 | 1,035 |

### Table 5. Development of a risk prediction model for each dataset using model selection

| Variable Selected | Australian Model, Odds Ratio (95% CI) | Leeds Model, Odds Ratio (95% CI) |
|-------------------|---------------------------------------|---------------------------------|
| Traditional risk factors | | |
| Family history of melanoma | | |
| None | 1.00 | 1.00 |
| 1 or more relatives | 1.61 (1.05–2.48) | 3.38 (1.33–8.59) |
| Hair color | | |
| Dark brown/black | 1.00 | 1.00 |
| Light brown | 1.01 (0.71–1.44) | 1.15 (0.79–1.68) |
| Fair or blonde | 1.82 (1.16–2.86) | 2.13 (1.32–3.42) |
| Red | 2.76 (1.36–5.60) | 1.86 (0.96–3.58) |
| Nevus density | | |
| None | 1.00 | 1.00 |
| Few | 1.19 (0.57–2.48) | 1.84 (1.09–3.11) |
| Some | 3.13 (1.50–6.52) | 3.93 (2.27–6.79) |
| Many | 5.36 (2.43–11.83) | 4.64 (2.36–9.15) |
| Nonmelanoma skin cancer | | |
| No | 1.00 | 1.00 |
| Yes | 2.28 (1.19–4.37) | 3.86 (0.77–19.40) |
| Blistering sunburn as a child | | |
| None | 1.00 | — |
| 1 or more episodes | 0.80 (0.58–1.11) | — |
| Sunbed use | | |
| None | 1.00 | — |
| 1–10 sessions | 0.96 (0.61–1.51) | — |
| >10 sessions | 1.79 (1.01–3.20) | — |
| Freckling as an adult | | |
| None/very few | — | 1.00 |
| Few/soon/many | 0.73 (0.52–1.02) | — |
| Eye color | | |
| Brown or black | — | 1.00 |
| Green or hazel | — | 1.05 (0.68–1.63) |
| Blue or grey | — | 1.39 (0.89–2.16) |
| Sun exposure hours on weekends and vacations | | |
| Quartile 1 (lower exposure) | 1.00 | — |
| Quartile 2 | 0.52 (0.34–0.81) | — |
| Quartile 3 | 0.61 (0.39–0.96) | — |
| Quartile 4 (higher exposure) | 0.44 (0.27–0.72) | — |
| Genomic variants | | |
| rs7412746 (ARNT) | 0.85 (0.67–1.06) | 0.81 (0.65–1.02) |
| rs62211989 (ASIP) | 1.90 (1.33–2.71) | 1.91 (1.28–2.84) |
| R151C (MC1R) | 2.59 (1.77–3.79) | 2.75 (1.83–4.13) |
| R160W (MC1R) | 1.47 (1.00–2.16) | 1.70 (1.15–2.51) |
| rs2487999 (OBFC1) | 1.40 (0.95–2.06) | 1.37 (0.93–2.01) |
| rs132985 (PLA2G6) | 1.19 (0.95–1.48) | 0.85 (0.68–1.06) |
| rs1393350 (TYR) | 1.32 (1.03–1.70) | 1.20 (0.95–1.52) |
| rs6949072 (ARGR3) | 1.28 (0.94–1.74) | — |
| rs7274597 (ASIP) | 0.50 (0.31–0.81) | — |
| rs76699054 (CCND1) | 1.40 (0.93–2.12) | — |
| rs12527588 (CDKAL1) | 1.56 (0.95–2.55) | — |
| rs3731217 (CDKN2A) | 0.79 (0.57–1.09) | — |
| D84E (MC1R) | 2.18 (1.02–4.67) | — |
| H155T (MC1R) | 2.60 (1.09–6.18) | — |
| V66L (MC1R) | 1.74 (1.21–2.50) | — |
| V92M (MC1R) | 1.70 (1.17–2.49) | — |
| rs45430 (MX2) | 0.72 (0.57–0.90) | — |
| rs3219090 (PARP1) | 0.73 (0.58–0.93) | — |

(continued)
The Australian Melanoma Family Study was a multicenter, population-based, case-control family study of invasive cutaneous melanoma diagnosed between ages 18 and 39 years. Recruitment of case (n = 629) and control (n = 535) participants was locally coordinated in Sydney, Melbourne, and Brisbane, Australia. The study design, recruitment, data collection, and participant characteristics have been described (Cust et al., 2009), and more details for both studies are provided in the Supplementary Materials online.

The Leeds case-control study recruited population-based incident histopathologically confirmed invasive melanoma cases (n = 2,184) in patients aged between 18 and 82 years and population-ascertained control individuals (n = 513) (Newton-Bishop et al., 2011; Randerson-Moor et al., 2009). This analysis focuses on those case individuals whose measurement protocols exactly matched the Australian study (n = 964).

Approval for the study was obtained from the ethics committees of the coordinating centers’ institutions in Australia and Leeds and the cancer registries. All participants provided written informed consent.

Self-reported personal sun exposure

Comprehensive data on lifetime sun exposure was collected by telephone interview. Questions referred to the frequency of sunburn and time spent outdoors between 9 a.m. and 5 p.m. separately for weekdays, weekends, and vacations in warmer months and in cooler months (Cust et al., 2011b; Newton-Bishop et al., 2011). Demographic information, ancestry, diagnoses of keratinocyte and other cancers, and family history information were also collected.

Pigmentary and nevus phenotype

Participants reported the skin color of their inside upper arm, eye color, natural hair color at age 18 years, freckling using Gallagher’s freckle chart (Lee et al., 2005), ability to tan, propensity to sunburn, usual tanning and sunburn response to prolonged or repeated exposure of skin to sunlight in summer, and nevus density (described pictorially as none, few, some, many) (Cust et al., 2009).

Selection of gene variants

We selected 21 genes/loci (45 SNPs) that had a confirmed association with melanoma risk in genome-wide association studies (Law et al., 2015) or for one gene using whole-genome sequencing approaches (MTF rs149617956 variant) (Yokoyama et al., 2011) (see Supplementary Table S3 online). Polygenic risk scores summarized the combined effects of the SNPs using a published method (Mavaddat et al., 2015) (for more technical details, see Supplementary Materials).

Statistical analysis

Analytic dataset. We excluded participants who did not give a blood sample or who failed genotyping, were missing data on traditional risk factors for melanoma, had germline CDKN2A pathogenic mutations or non-European ethnicity (on self-report or principal components analysis), and Australian controls 45 years or older at interview (because all Australian cases were diagnosed when patients were younger than 40 years). The analysis dataset thus

### Table 5. Continued

| Variable Selected | Australian Model, Odds Ratio (95% CI) | Leeds Model, Odds Ratio (95% CI) |
|-------------------|--------------------------------------|---------------------------------|
| rs2736100 (TERT)  | 0.74 (0.59–0.94)                      | —                               |
| rs34585474 (AGTR3) | 1.27 (0.89–1.83)                      | —                               |
| rs7781130 (AGTR3) | 1.59 (0.85–2.96)                      | —                               |
| rs1801516 (ATM)  | 0.77 (0.55–1.07)                      | —                               |
| rs7006353 (CASPB8) | 1.27 (0.99–1.63)                      | —                               |
| rs7776158 (CDKAL1) | 1.36 (1.06–1.75)                      | —                               |
| rs16953002 (FTO) | 1.27 (0.73–2.16)                      | —                               |
| R163Q (MC1R)     | 0.62 (0.36–1.09)                      | —                               |
| D294H (MC1R)     | 1.11 (1.62–10.46)                     | —                               |
| rs6517661 (AX2)  | 0.75 (0.53–1.02)                      | —                               |
| rs113908778 (RAD23B) | 0.55 (0.24–1.24)                | —                               |
| rs4436178 (RAD23B) | 1.87 (0.82–4.26)                      | —                               |

Abbreviation: CI, confidence interval.

1 A risk prediction model was developed separately for each dataset using a backward selection process in which traditional and genomic risk factors with P < 0.20 were retained in the multivariable model in addition to forced variables age, sex, city of recruitment, and European ancestry. The same genetic variants and traditional risk factors were assessed for inclusion in both models.

2 Odds ratios derived from the respective dataset, adjusted for all other variables in the model. For genomic variants, the per-allele odds ratio is presented. Values left blank indicate that the factor was not included in the final model for that dataset (Australia/Leeds). The areas under the curve for the Australian model were 0.80 (95% CI = 0.77–0.83) from the development model, 0.77 (95% CI = 0.74–0.80) from internal validation (10-fold cross-validation), and 0.72 (95% CI = 0.69–0.75) from external validation using the Leeds dataset. The areas under the curve for the Leeds model were 0.77 (95% CI = 0.73–0.80) from the development model, 0.72 (95% CI = 0.69–0.75) from internal validation, and 0.77 (95% CI = 0.74–0.80) from external validation using the Australian dataset. Both models were well calibrated in the external datasets (Hosmer-Lemeshow P = 0.57 for both).

discrimination improvement, but SNPs in other genes further improved risk prediction. Prediction models based on both genomic and traditional risk factors could increase the yet unproven capacity of models based only on traditional risk factors to motivate melanoma risk reduction behaviors.

### MATERIAL AND METHODS

**Study samples**

The Australian Melanoma Family Study was a multicenter, population-based, case-control family study of invasive cutaneous melanoma diagnosed between ages 18 and 39 years. Recruitment of case (n = 629) and control (n = 535) participants was locally coordinated in Sydney, Melbourne, and Brisbane, Australia. The study design, recruitment, data collection, and participant characteristics have been described (Cust et al., 2009), and more details for both studies are provided in the Supplementary Materials online.

The Leeds case-control study recruited population-based incident histopathologically confirmed invasive melanoma cases (n = 2,184) in patients aged between 18 and 82 years and population-ascertained control individuals (n = 513) (Newton-Bishop et al., 2011; Randerson-Moor et al., 2009). This analysis focuses on those case individuals whose measurement protocols exactly matched the Australian study (n = 964).

Approval for the study was obtained from the ethics committees of the coordinating centers’ institutions in Australia and Leeds and the cancer registries. All participants provided written informed consent.

### Self-reported personal sun exposure

Comprehensive data on lifetime sun exposure was collected by telephone interview. Questions referred to the frequency of sunburn and time spent outdoors between 9 a.m. and 5 p.m. separately for weekdays, weekends, and vacations in warmer months and in cooler months (Cust et al., 2011b; Newton-Bishop et al., 2011). Demographic information, ancestry, diagnoses of keratinocyte and other cancers, and family history information were also collected.

### Pigmentary and nevus phenotype

Participants reported the skin color of their inside upper arm, eye color, natural hair color at age 18 years, freckling using Gallagher’s freckle chart (Lee et al., 2005), ability to tan, propensity to sunburn, usual tanning and sunburn response to prolonged or repeated exposure of skin to sunlight in summer, and nevus density (described pictorially as none, few, some, many) (Cust et al., 2009).

### Selection of gene variants

We selected 21 genes/loci (45 SNPs) that had a confirmed association with melanoma risk in genome-wide association studies (Law et al., 2015) or for one gene using whole-genome sequencing approaches (MTF rs149617956 variant) (Yokoyama et al., 2011) (see Supplementary Table S3 online). Polygenic risk scores summarized the combined effects of the SNPs using a published method (Mavaddat et al., 2015) (for more technical details, see Supplementary Materials).

### Statistical analysis

**Analytic dataset.** We excluded participants who did not give a blood sample or who failed genotyping, were missing data on traditional risk factors for melanoma, had germline CDKN2A pathogenic mutations or non-European ethnicity (on self-report or principal components analysis), and Australian controls 45 years or older at interview (because all Australian cases were diagnosed when patients were younger than 40 years). The analysis dataset thus...
included 1,035 participants (578 case individuals, 457 control individuals) from Australia and 1,460 participants (964 case individuals, 496 control individuals) from Leeds.

**Description of models.** The primary analysis used ORs derived from published meta-analyses to prevent overfitting, that is, over-estimation of the prediction accuracy that can occur when estimating the ORs from the same dataset that the predictions are made from (Steyerberg, 2009; Wray et al., 2013). The Australian and Leeds samples contributed data to the meta-analyses, but their data represented less than 10% of the total sample. In secondary analyses presented in Supplementary Table S2, we show the results based on ORs derived from the study datasets. The published ORs and ORs derived from the datasets are shown in Supplementary Table S3 for genomic variants and Supplementary Table S4 online for traditional risk factors. The published ORs for the traditional risk factors were from fully adjusted models in meta-analyses, and we categorized our variables the same as the published data. The published ORs for genomic variants were obtained from a meta-analysis of genome-wide association studies (Law et al., 2015), using pooled ORs from a fixed effects model or random effects where there was evidence of heterogeneity (I² > 31%). Because the association of MC1R with melanoma risk is modified by phenotype (Pasquali et al., 2015), we incorporated phenotype-stratified ORs for each of the MC1R variants for models that included traditional risk factors; for this stratification, participants were classified as having a sun-sensitive phenotype if they had one or more of freckles (few, some, many), red hair, or skin that usually or always burns.

Supplementary Table S3 shows the biological pathways through which each gene is thought to influence melanoma risk: pigmentation (14 genes); nevus (7 genes); and telomere, senescence, and other pathways (5 genes). Classification of pathways was based on associations of SNPs with each of these traits (Choi et al., 2017; Codd et al., 2013; Duffy et al., 2017; Law et al., 2015), and genes could be allocated to multiple pathways.

Logistic regression models were used to assess associations between melanoma and traditional and genomic risk factors, adjusted for demographic and study design factors: age, sex, city of recruitment (Australia only), and self-reported European ancestry (British, other Northern European, Southern/Eastern European, mixed/other European); ancestry was included as a covariate to minimize confounding by ethnicity. The base risk prediction model included well-established independent traditional risk factors and UV exposure variables (hair color, skin color, eye color, freckling as an adult, nevus density, reported skin photoprotection, personal history of keratinocyte cancer, first-degree family history of melanoma, blistering sunburns as a child, vacation sun exposure, and sunbed use). The incremental contribution of melanoma risk SNPs was assessed overall, by biological pathway, and for MC1R alone.

**Model performance.** The ability of the model to discriminate between case and control individuals was evaluated by calculating the AUC, NRI, and OR per standard deviation (adjusted for age and sex using the OPERA method [Hopper, 2015]). The AUC is the probability that the predicted risk is higher for a case individual than for a control individual and ranges from 0.5 (equivalent to a coin toss) to 1.0 (perfect discrimination). The NRI quantifies overall improvement in model sensitivity and specificity; it quantifies movement of case and control individuals to higher or lower predicted risk probabilities when a new risk model is applied (Leeming et al., 2014). Because there are no established risk thresholds for melanoma, we used category-free continuous NRI (Pencina et al., 2011). We used the Hosmer-Lemeshow goodness-of-fit test to assess calibration, that is, the agreement between observed and predicted probabilities of melanoma (Steyerberg, 2009). For the secondary analyses that used ORs derived from the study datasets, we performed 10-fold cross-validation as a measure of internal validation to help correct for overfitting (Steyerberg, 2009).

**Cross-tabulation of polygenic risk score versus traditional risk score.** Traditional and polygenic risk scores (each based on published risk estimates) were categorized into tertiles and cross-tabulated to compare the concordance of genomic with traditional risk. The traditional risk score variable was derived from the predicted probability of melanoma for each individual, based on beta-values from a logistic regression model with melanoma (case-control) status as the outcome and traditional risk factors as predictors.

Data were analyzed using SAS, version 9.4 (SAS Institute, Cary NC), and statistical significance was inferred at two-sided P less than 0.05. We reported the study according to published guidelines (Collins et al., 2015; Janssens et al., 2011).

**ORCIDs**

Anne E Cust: http://orcid.org/0000-0002-5331-6370
D Timothy Bishop: http://orcid.org/0000-0002-8752-8785

**CONFLICT OF INTEREST**

SM received grant funding from the Australian Research Council, Australian National Health and Medical Research Council. The other authors state no conflicts of interest.

**ACKNOWLEDGMENTS**

This work was supported by the National Health and Medical Research Council of Australia (NHMRC) (566946, 107359, 211172, 402761), the Cancer Council New South Wales (77/00, 06/10), Cancer Council Victoria, Cancer Council Queensland (371), and the US National Institutes of Health (NIH grant ROI CA83115 to the international Melanoma Genetics Consortium [GenoMEC]). The Leeds group was funded by program grant from Cancer Research UK C588/A19167 and EU under Framework 6 LSHM-CT-2005-018702. AEC received Career Development Fellowships from the NHMRC (1147843) and Cancer Institute NSW (15/CDF/1-14). This research was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Division of Cancer Epidemiology and Genetics.

**SUPPLEMENTARY MATERIAL**

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2018.05.023.

**REFERENCES**

Aitken JF, Elwood M, Baade PD, Youl P, English D. Clinical whole-body skin examination reduces the incidence of thick melanomas. Int J Cancer 2010;126:450–8.

Armstrong BK, Kricker A. How much melanoma is caused by sun exposure? Melanoma Res 1993;3:395–401.

Bataille V, Snieder H, MacGregor AJ, Sasieni P, Spector TD. Genetics of risk factors. Pigment Cell Melanoma Res 2014;27:488–9.

Berwick M, Macarthur J, Orlow I, Kanetsky P, Begg CB, Luo L, et al. MITF E318K’s effect on melanoma risk independent of, but modified by, other risk factors. Pigment Cell Melanoma Res 2014;27:488–9.

Breitbart BW, Waldmann A, Nolte S, Capellaro M, Greinert R, Volkmann B, et al. Systematic skin cancer screening in northern Germany. J Am Acad Dermatol 2012;66:201–11.

Cho HG, Ransohoff KJ, Yang L, Hedlin H, Assimes T, Han J, 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.

Choi J, Xu M, Makowski MM, Zhang T, Law MH, Kovacs MA, et al. A common intrinsic variant of PARP1 confers melanoma risk and mediates melanocyte growth via regulation of MITF. Nat Genet 2017;49:1326–35.

D Timothy Bishop: http://orcid.org/0000-0002-8752-8785

AE Cust et al.

Common Genomic Variants and Melanoma Risk Prediction
Common Genomic Variants and Melanoma Risk Prediction

Codd V, Nelson CP, Albrecht E, Mangino M, Deelen J, Buxton JL, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet 2013;45:422–7.

Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;350:g7594.

Cust AE, Armstrong BK, Goumas C, Jenkins MA, Schmid H, Hopper JL, et al. Sunbed use during adolescence and early adulthood is associated with increased risk of early-onset melanoma. Int J Cancer 2011a;128:2425–35.

Cust AE, Goumas C, Holland EA, Agha-Hamilton C, Atkin JF, Armstrong BK, et al. MC1R genotypes and risk of melanoma before age 40 years: a population-based case-control-family study. Int J Cancer 2012;131:E269–81.

Cust AE, Goumas C, Vuong K, Davies JR, Barrett JH, Holland EA, et al. MC1R genotype as a predictor of early-onset melanoma, compared with self-reported and physician-measured traditional risk factors: an Australian case-control-family study. BMC Cancer 2013;13:406.

Cust AE, Jenkins MA, Goumas C, Armstrong BK, Schmid H, Atkin JF, et al. Early-life sun exposure and risk of melanoma before age 40 years. Cancer Causes Control 2011b;22:885–97.

Cust AE, Schmid H, Maskilli JA, Jetam J, Ferguson M, Holland EA, et al. Population-based, case-control-family design to investigate genetic and environmental influences on melanoma risk: Australian Melanoma Family Study. Am J Epidemiol 2009;170:1541–54.

Duffy D, Liu G, Li X, Sanna M, Iles M, Jacobs LC, et al. Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. bioRxiv 2017. https://doi.org/10.1101/173112 (accessed 13 February 2018).

Fang S, Han J, Zhang M, Wang LE, Wei Q, Amos CI, et al. Joint effect of multiple common SNPs predicts melanoma susceptibility. PLoS One 2013;8(12):e85642.

Gandini S, Sera F, Catrarozza MS, Pasquini P, Abeni D, Boyle P, et al. Meta-analysis of risk factors for cutaneous melanoma: I. common and atypical naevi. Eur J Cancer 2005a;41:28–44.

Gandini S, Sera F, Catrarozza MS, Pasquini P, Piccioni O, Boyle P, et al. Meta-analysis of risk factors for cutaneous melanoma: II. sun exposure. Eur J Cancer 2005b;41:45–60.

Garcia-Closas M, Gunson NB, Chatterjee N. Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer. J Natl Cancer Inst 2014;106(11):dju305.

Glanz K, Volpicelli K, Jepson C, Ming ME, Schuchter LM, Armstrong K. Effects of tailored risk communications for skin cancer risk prediction and detection: the PennSCAPE Randomized Trial. Cancer Epidemiol Biomarkers Prev 2015;24:415–21.

Hopper JL. Odds per adjusted standard deviation: comparing strengths of association for risk factors measured on different scales and across diseases and populations. Am J Epidemiol 2015;182:863–7.

Howell A, Anderson AS, Clarke RB, Duffy SW, Evans DG, Garcia-Closas M, et al. Risk determination and prevention of breast cancer. Breast Cancer Res 2014;16:446.

Janssens AC, Ioannidis JP, van Duijn CM, Little J, Khoury MJ. Strengthening the reporting of genetic risk prediction studies: the GRIPS statement. Eur J Clin Investig 2011;41:1004–9.

Kanetsky PA, Hay JL. Marshalling the translational potential of MC1R for precision risk assessment of melanoma. Cancer Prev Res 2017;11(3). canprevres.0255.2017.

Kanetsky PA, Panossian S, Elder DE, Guerry D, Ming ME, Schuchter L, et al. Does MC1R genotype convey information about melanoma risk beyond risk phenotypes? Cancer 2010;116:2416–28.

Kypreou KP, Stefanaki I, Antonopoulou K, Karagianni F, Nitrigos G, Zaras A, 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.

Law MH, Bishop DT, Lee JE, Brossard M, Martin NG, Moses EK, et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. Nat Genet 2015;47:987–95.

Lee TK, Rivers JK, Gallagher RP. Site-specific protective effect of broadband sunscreen on nevus development among white schoolchildren in a randomized trial. J Am Acad Dermatol 2005;52:786–92.

Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician’s guide. Ann Intern Med 2014;160:122–31.

Lu Y, Ek WE, Whitehead D, Vaughan TL, Spurhse AB, Easton DF, et al. Most common ‘sporadic’ cancers have a significant germline genetic component. Hum Mol Genet 2014;23:6112–8.

Mavaddat N, Pharaoh PD, Michailidou K, Tyrer J, Brook MN, Boilla MK, et al. Prediction of breast cancer risk based on profiling with common genetic variants. J Natl Cancer Inst 2015;107(3):dxj036.

Newton-Bishop JA, Chang YM, Elliott F, Chan M, Leake S, Karpavicius B, et al. Relationship between sun exposure and melanoma risk for tumours in different body sites in a large case-control study in a temperate climate. Eur J Cancer 2011;47:732–41.

Olsen CM, Carroll HJ, Whitehead DC. Estimating the attributable fraction for cancer: a meta-analysis of nevi and melanoma. Cancer Prev Res 2010a;3:233–45.

Olsen CM, Carroll HJ, Whitehead DC. Familial melanoma: a meta-analysis and estimates of attributable fraction. Cancer Epidemiol Biomarkers Prev 2010b;19:65–73.

Pasquali E, Garcia-Barton JC, Fargnoli MC, Gandini S, Maisonneuve P, Bagnardi V, et al. MC1R variants increased the risk of sporadic cutaneous melanoma in darker-pigmented Caucasians: a pooled-analysis from the M-SKIP project. Int J Cancer 2015;136:618–31.

Pencina MJ, D’Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med 2011;30:11–21.

Penn LA, Qin M, Zhang E, Ng E, Shao Y, Berwick M, et al. Development of a melanoma risk prediction model incorporating MC1R genotype and indoor tanning exposure: impact of mole phenotype on model performance. PLoS One 2014;9(7):e101507.

Randerson-Moor JA, Taylor JC, Elliott F, Chang YM, Beswick S, Kukalich K, et al. Vitamin D receptor gene polymorphisms, serum 25-hydroxyvitamin D levels, and melanoma: UK case-control comparisons and a meta-analysis of published VDR data. Eur J Cancer 2009;45:3271–81.

Royal Australian College of General Practitioners. Guidelines for preventive activities in general practice, 9th ed. East Melbourne, Australia: Royal Australian College of General Practitioners; 2016.

Smit AK, Newson AJ, Morton RL, Kimlin M, Keogh L, Law MH, et al. The melanoma genomics managing your risk study: a protocol for a randomized controlled trial evaluating the impact of personal genomic risk information on skin cancer prevention behaviors. Contemp Clin Trials 2018;70:106–16.

Stefanaki I, Panagiotou OA, Kodela E, Gogas H, Kypreou KP, Chatzinassiou F, et al. Replication and predictive value of SNPs associated with melanoma and pigmentation traits in a southern European case-control study. PLoS One 2013(8):2:035712.

Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.

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

US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, Curry S, Davidson KW, Ebell M, et al. Screening for skin cancer: US Preventive Services Task Force recommendation statement. JAMA 2016;314:429–35.

Usher-Smith JA, Walter FM, Emery JD, Win AK, Griffin SJ. Risk prediction models for colorectal cancer: a systematic review. Cancer Prev Res 2016;9:13–26.

Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. Nat Rev Genet 2013;14:507–15.

Yokoyama S, Woods SL, Boyle GM, Aoude LG, MacGregor S, Zismann V, et al. A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. Nature 2011;480(7375):99–103.

This work is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/