Estimating \textit{CDKN2A} Carrier Probability and Personalizing Cancer Risk Assessments in Hereditary Melanoma Using MelaPRO

Wenyi Wang$^{1,2}$, Kristin B. Niendorf$^3$, Devanshi Patel$^4$, Amanda Blackford$^9$, Fabio Marroni$^{10}$, Arthur J. Sober$^5$, Giovanni Parmigiani$^{7,8}$, and Hensin Tsao$^{4,5,6}$

Abstract

Personalized cancer risk assessment remains an essential imperative in postgenomic cancer medicine. In hereditary melanoma, germline \textit{CDKN2A} mutations have been reproducibly identified in melanoma-prone kin-dreds worldwide. However, genetic risk counseling for hereditary melanoma remains clinically challenging. To address this challenge, we developed and validated MelaPRO, an algorithm that provides germline \textit{CDKN2A} mutation probabilities and melanoma risk to individuals from melanoma-prone families. MelaPRO builds on comprehensive genetic information, and uses Mendelian modeling to provide fine resolution and high accuracy. In an independent validation of 195 individuals from 167 families, MelaPRO exhibited good discrimination with a concordance index ($C$) of 0.86 [95% confidence intervals (95% CI), 0.75–0.97] and good calibration, with no significant difference between observed and predicted carriers (26; 95% CI, 20–35, as compared with 22 observed). In cross-validation, MelaPRO outperformed the existing predictive model MELPREDICT ($C$, 0.82; 95% CI, 0.61–0.93), with a difference of 0.05 (95% CI, 0.007–0.17). MelaPRO is a clinically accessible tool that can effectively provide personalized risk counseling for all members of hereditary melanoma families. Cancer Res; 70(2); 552–9. ©2010 AACR.

Introduction

In 2009, there will be an estimated 68,720 new cases of melanoma with 8,650 deaths (1). Despite decades of therapeu-tic investigation, metastatic melanoma is still considered incurable, thereby making identification of high-risk individu-als with an eye towards early detection, a cornerstone in the strategy for cure.

A fundamental goal of personalized medicine is to uncover germline variants that identify individuals at the greatest risk for disease. For melanoma, the first such mutations were found in \textit{CDKN2A} over a decade ago (2). Since then, heritable alterations in \textit{CDKN2A} (encoding two proteins; p16/Ink4a and p14/ARF) and \textit{CDK4}, the inhibitory target of p16/Ink4a, have also been found in a significant subset of melanoma-prone families (2–8). Earlier validation of a computational tool—MELPREDICT, for estimating \textit{CDKN2A} carrier probability, showed reasonable performance in ranking carriers higher than noncarriers among patients with melanoma (9). However, MELPREDICT is based on logistic regression models and therefore cannot effectively incorporate crucial biological information embedded within the pedigree structure. Moreover, it lacks the flexibility to account for variations in \textit{CDKN2A} mutation prevalence and penetrance across geographic regions (3, 4).

To this end, we developed MelaPRO, a new model to estimate the probability of carrying a mutation of \textit{CDKN2A} in melanoma families, using a general Mendelian risk prediction approach (10) that integrates Mendelian inheritance and Bayesian probability theories. This computational strategy effectively translates genetic information into a clinically useful algorithm for carrier probability estimation and has been successfully applied to develop BRCAPRO for the breast and ovarian cancer syndrome (11–14), MMRpro for the Lynch syndrome (15), and PancPRO for familial pancreatic cancer (16). In this initial validation, we show that MelaPRO exhibits strong discrimination and calibration ability, and outperforms the regression model, MELPREDICT.
Materials and Methods

Model Development

MelaPRO translates population estimates of the mutation prevalence and penetrance of CDKN2A into mutation prediction for any designated family member (the counselee), given his or her family history, and assuming autosomal dominant inheritance. The penetrance refers to the age-specific risk of developing cutaneous melanoma depending on CDKN2A carrier status and gender.

The carrier probability is modeled via Bayes’ rule as follows (10): 

\[ P_{\text{genotype}|\text{history}} = P_{\text{genotype}} \times P_{\text{history}|\text{genotype}} / P_{\text{history}} \]

Here, \( P_{\text{denotes probability, genotype denotes whether the counselee carries a deleterious mutation in CDKN2A, and history denotes family history (as detailed in Table 1). The } P_{\text{genotype|history}} \text{ term is the mutation prevalence and } P_{\text{history|genotype}} \text{ is a weighted average of the probabilities of family history given each possible genotype configuration of all relatives, where the weights are the probabilities of the genotype configuration based on Mendelian transmission. This step uses the Elston-Stewart algorithm (17), as implemented in the latest version of the BayesMendel R package.}^{11}

The probability of family history given each genotype configuration can be broken down into the product of each relative’s probability of phenotype given the genotype, assuming conditional independence. Here, each probability term is calculated as either the cumulative penetrance (age specific) for affected relatives [multiple primary melanomas (MPM)] or 1 − cumulative penetrance for unaffected relatives. \( P_{\text{history}} \) is the sum of terms \( P_{\text{genotype}} \times P_{\text{history|genotype}} \) across all possible genotypes of the counselee. Risks of developing SPM and MPM for unaffected individuals are estimated by a weighted average of the carrier’s and noncarrier’s penetrance, where the weights are the carrier probabilities.

Parameter Settings

MelaPRO incorporates three distinct penetrance estimates. The GenoMEL consortium (3) collected high-risk families (more than two affected family members) and estimated separate penetrances for areas with high (HBI) and low (LBI) baseline incidences up to age 80 using logistic regression. Alternatively, the GEM Study Group (4) collected melanoma patients from the general population and estimated the penetrance in 5-year age intervals using the nonparametric kin-cohort method. We extrapolated the GenoMEL data and interpolated the GEM data using estimates from the SEER DevCan Software\(^12\) as a reference, to establish the age-specific penetrance between ages 1 and 110. We then calculated the mutation prevalence indirectly from the penetrance estimates. By Bayes’ rule, we have \( P(G) = P(G|B) \times P(B) / P(B|G) \), where \( G \) denotes being a CDKN2A carrier, and \( B \) denotes new cases per year between 2001 and 2005. We then calculated the mutation prevalence indirectly from the penetrance estimates. By Bayes’ rule, we have \( P(G) = P(G|B) \times P(B) / P(B|G) \), where \( G \) denotes being a CDKN2A carrier, and \( B \) denotes new cases per year between 2001 and 2005. From previous studies, we obtained \( P(G|B) = 0.0179 \) (4) and \( P(B) = 19.4/100,000 \)\(^13\) for the North American population. \( P(B|G) \) is a weighted average of probability distribution

Table 1. Family history as input for MelaPRO and resulting output

<table>
<thead>
<tr>
<th>Features</th>
<th>Input (for the counselee and each relative)</th>
<th>Output (for the counselee)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Parent information (IDs)</td>
<td>Probability that the counselee carries a deleterious mutation of CDKN2A</td>
</tr>
<tr>
<td>Melanoma status: 0 if unaffected, 1 if single primary melanoma, 2 if multiple primary melanomas</td>
<td>Probability that the counselee, if asymptomatic, will develop single or multiple primary melanomas in the future, in yearly intervals</td>
<td></td>
</tr>
<tr>
<td>Age at first diagnosis (in years) of melanoma if affected</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current age or age at last follow-up (in years) if unaffected</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Result of previous germline testing of CDKN2A in any family member: 0 if missing, 1 if mutation present, 2 if mutation not found</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Any input could be designated “Not Applicable” if unavailable, with the exception of the parental IDs that relate each family member to the counselee. Personalized predictions can be made for each family member by changing the designation of the counselee.

\(^{11}\) http://astor.som.jhmi.edu/BayesMendel/


function of penetrance for CDKN2A carrier, where the weights are melanoma incidences within 10-y age intervals.\textsuperscript{13}

**Accounting for MPM versus SPM**

We used $X$ to indicate the number of primary melanomas and $G$ to indicate carrier status, with $X = 1$ for SPM and $X = 2$ for MPM. The published penetrance estimates are $P_0 = \Pr(X \geq 1|G = 0)$ and $P_1 = \Pr(X \geq 1|G = 1)$. The relative risk of MPM for carriers and noncarriers among melanoma cases is $\Pr(X \geq 2|X \geq 1, G = 1) / \Pr(X \geq 2|X \geq 1, G = 0) = 1.8$ (18), and the risk ratio of having MPM versus SPM for carriers is $\Pr(X \geq 2|G = 1) / \Pr(X = 1|G = 1) = 1.14$ (by age 50; ref. 4). Based on these numbers, we estimated the MPM- and SPM-specific penetrances.

**Test sensitivity and specificity.** For the genetic results, the default specificity was set at 1.0, because only known mutations were included in the analysis; putative polymorphisms (e.g., Ala148Thr) and variants of unknown significance were excluded because accurate penetrance data are not available for these alterations. As such, the model does not currently calculate the probability of detecting established polymorphisms or variants of unknown significance. Other possibilities for a false positive, such as sample confusion, can be considered negligible. Because our mutational screen does not detect deep intronic mutations and large chromosomal deletions, we set our sensitivity at 0.9 presuming that these types of deleterious changes occur in no more than 10% of the cases. It is straightforward for users to replace these estimates with different ones.

**MELPREDICT.** The MELPREDICT model (9) is a multiple logistic regression, in which the estimated carrier probability of the counselee being a mutation carrier is given by $e^{\beta X + \beta 0 + \beta 1 \times \text{(no. of counselee primaries)} + \beta 2 \times \text{(no. of additional family primaries)} + \beta 3 \times \text{ln (counselee age)}} / (1 + e^{\beta X + \beta 0 + \beta 1 \times \text{(no. of counselee primaries)} + \beta 2 \times \text{(no. of additional family primaries)} + \beta 3 \times \text{ln (counselee age)}})$. 

**Validation**

**Study population.** We used data from the Massachusetts General Hospital Melanoma and Pigmented Lesion Center. This series was not used in the development of MelaPRO and provides an independent validation. This study was performed in accordance with a protocol approved by the Massachusetts General Hospital Institutional Review Board. From April 2001 to January 2008, all patients with invasive or in situ melanoma evaluated at the Pigmented Lesion Center were screened for eligibility as follows: (a) $\geq 1$ first-degree relatives with melanoma, or (b) $\geq 2$ affected relatives with melanoma on one side of the family (first or second degree), or (c) $\geq 3$ primary cutaneous melanomas irrespective of family history. The presence and number of melanomas for counselees were confirmed via pathology reports, except for a small number of cases ($< 10\%$, data not shown). Medical record confirmation of reported family histories was pursued but limited to relatives who provided prior consent to participate in the study. We excluded two families: one because it lacked counselee information and the other because the counselee was unaffected and therefore ineligible for comparison with MELPREDICT.

**Mutation analysis.** CDKN2A exons 1\(\alpha\), 1\(\beta\), and 2 were screened for sequence variants as previously described (9).

**Data analysis.** All analyses were performed in R.\textsuperscript{14} Within each family, we assigned each CDKN2A-tested individual in turn as the counselee and calculated the probability of detecting a CDKN2A mutation using MELPREDICT (9) and all modules of MelaPRO. For the MelaPRO modules, this probability is obtained by multiplying the probability of carrying a CDKN2A mutation, provided by the model, by the sensitivity of the mutation analysis (default = 0.9). The comparison between models required additional exclusion of four cases in which tested individuals were unaffected, as MELPREDICT does not apply. We evaluated the discrimination, calibration, and accuracy performance of each model by comparing the calculated probabilities with the observed mutational status. Discrimination reflects a model’s ability to differentiate individuals with positive outcomes from those with negative outcomes. It can be visualized using receiver operating characteristic curves and summarized by the underlying area, or concordance index ($C$). Calibration is a model’s ability to make unbiased estimates of the proportion of carriers. We also used positive predictive value and negative predictive value to measure accuracy, and mean squared error for an overall comparison of performance. MELPREDICT was developed based on a subset of our validation set. Therefore, we used cross-validation (leave-one-out) to obtain evaluation statistics for MELPREDICT. In our cross-validation, we fixed the covariates selected by the original MELPREDICT model, but re-estimated the coefficients in each training set. We obtained 95% confidence intervals (95% CI) using the bootstrap (19). We also evaluated sensitivity and specificity for the descriptive classifier [summary family history criterion (FH)] defined by having at least two affected relatives. We present hypothetical but realistic family history scenarios for illustration.

**Results**

**MelaPRO features.** The MelaPRO model treats melanoma family history as a diagnostic test or profile, and CDKN2A genotype as an occult condition to be diagnosed. To use MelaPRO during a typical counseling session, the counselor collects the counselee’s family history information, and enter it into MelaPRO to obtain a carrier probability and an estimate of future risk if the counselee is still free of the disease. The family history information is detailed in Table 1, and includes family members’ relationship, occurrence of cutaneous melanoma (including whether single or multiple primaries were found), age of diagnosis, or age at last contact for unaffected family members, and earlier germline testing results of any family members, if available. There is no restriction to which family member can be designated the counselee and calculated the probability of detecting a CDKN2A mutation using MELPREDICT (9) and all modules of MelaPRO. For the MelaPRO modules, this probability is obtained by multiplying the probability of carrying a CDKN2A mutation, provided by the model, by the sensitivity of the mutation analysis (default = 0.9). The comparison between models required additional exclusion of four cases in which tested individuals were unaffected, as MELPREDICT does not apply. We evaluated the discrimination, calibration, and accuracy performance of each model by comparing the calculated probabilities with the observed mutational status. Discrimination reflects a model’s ability to differentiate individuals with positive outcomes from those with negative outcomes. It can be visualized using receiver operating characteristic curves and summarized by the underlying area, or concordance index ($C$). Calibration is a model’s ability to make unbiased estimates of the proportion of carriers. We also used positive predictive value and negative predictive value to measure accuracy, and mean squared error for an overall comparison of performance. MELPREDICT was developed based on a subset of our validation set. Therefore, we used cross-validation (leave-one-out) to obtain evaluation statistics for MELPREDICT. In our cross-validation, we fixed the covariates selected by the original MELPREDICT model, but re-estimated the coefficients in each training set. We obtained 95% confidence intervals (95% CI) using the bootstrap (19). We also evaluated sensitivity and specificity for the descriptive classifier [summary family history criterion (FH)] defined by having at least two affected relatives. We present hypothetical but realistic family history scenarios for illustration.

\textsuperscript{14} R Development Core Team: A language and environment for statistical computing. R Development Core, Vienna, Austria. 2006 URL http://www.r-project.org/.
Supplementary Figure S1 shows the penetrance estimates from GenoMEL (3) and GEM (4), which applied to all melanoma diagnoses combined. We estimated the allele frequency (mutation prevalence) as 0.00015 using the HBI penetrance, 0.0003 using the LBI penetrance, and 0.0004 using the GEM penetrance. Additionally, MelaPRO incorporated an estimate that 53% of diagnoses in carriers are MPM compared with 30% for noncarriers.

MelaPRO provides three modules: MelaPRO-HBI (HBI), MelaPRO-LBI (LBI), and MelaPRO-GEM (GEM) reflecting different penetrances, now adjusted to be MPM/SPM-specific. Users choose the module that best matches the population in which the model is used to the characteristics of the original studies. To illustrate, in Fig. 1, for scenario 1 (Table 2), MelaPRO gave a probability estimate of 0.43 (HBI), 0.90 (LBI), and 0.85 (GEM). For comparison, the probabilities without the MPM/SPM adjustment are 0.27 (HBI), 0.83 (LBI), and 0.74 (GEM). Users can also specify the sensitivity (default = 0.9) and specificity (default = 1) of the germline testing method when results are available for some family members.

Software. MelaPRO is open source and is freely available as part of the BayesMendel (10) risk prediction package, along with the CancerGene (20) counseling package.

Clinical illustration. Figure 1 illustrates how MelaPRO provides high-resolution information to support clinical counseling, by presenting carrier probability estimates for several hypothetical, but realistic, scenarios. We compared our results to the descriptive classifier FH, and to MELPREDICT (9), a logistic regression model based on the number of primary melanomas in the counselee, in all other family members and the counselee’s age at diagnosis.

In the pedigree, MELPREDICT estimated a carrier probability of 0.24 as compared with MelaPRO’s estimate of 0.43 (HBI, see other modules in Table 2). MelaPRO captured the two relatives’ earlier disease onset (59 years in the general population) as an additional indication of carrier status. It also responded, with considerable increase in probabilities, to modification of the father’s disease history, whereas the total number of familial melanomas remained the same (HBI, from scenario 1 to 3: 0.43–0.77, and from scenario 1 to 4: 0.43–0.82). The additional scenarios show how carrier probabilities varied as the number of affected individuals and patients’ relationship to the counselee were changed (i.e., aunt healthy or brother affected).

External validation. We assembled a validation set containing 167 families with an average of 29 members. There were, in total, 26 carriers, 22 of which were affected with melanoma, and 603 primary melanomas. The mean number of primary melanomas in families of carriers and noncarriers was 7.9 (95% CI, 5.5–10.3) and 3.2 (95% CI, 2.9–3.5), respectively. There were 207 genotyped individuals within the 167 families. Among these, 195 were cases, with 85 males and 110 females. The mean age at diagnosis was 46.4 years (95% CI, 43.5–49.3) for males and 41.3 years (95% CI, 38.6–43.9) for females, and it was 36.6 years (95% CI, 30.7–42.6) for affected carriers and 44.4 years (95% CI, 42.3–46.4) for affected noncarriers. The proportion of mutation carriers increased with the number of primary melanomas in the counselee, the number of affected relatives, and the number of primary melanomas in relatives (Supplementary Table S1). There were a total of eight relatives from seven families affected with pancreatic cancer.

The Boston validation cohort was derived from a relatively high-incidence region (3) and is familial in ascertainment.
We deployed MelaPRO-HBI and predicted the presence of approximately 26 mutations (95% CI, 20–35); the MelaPRO-GEM module predicted 41 mutations (95% CI, 31–58), and MELPREDICT predicted 20 mutations (95% CI, 19–24; see observed/expected ratios in Table 3). Both MelaPRO-HBI and MELPREDICT showed a close correspondence with the 22 mutations observed. MelaPRO-GEM and MelaPRO-LBI predicted a substantially higher number of mutations than was observed, likely because their parameter estimates do not fit our cohort profile.

MelaPRO shows good discriminatory ability with all three modules. Figure 2 shows the receiver operating characteristic curves for MelaPRO and MELPREDICT, as well as the sensitivity and specificity based on the summary family history criterion FH. The corresponding area under the curves are presented in Table 3. The difference between the area under the curves for MelaPRO-HBI and that for MELPREDICT is 0.05 (95% CI, 0.007–0.17). Part of this difference is attributable to the gap visible at the top right portion of Fig. 2: MelaPRO achieved an estimated sensitivity of 90% at the cost of ∼70% false positives, whereas MELPREDICT provided limited discrimination at this level of sensitivity. The point corresponding to the sensitivity and specificity based on FH lay below the receiver operating characteristic curves, with an 81% sensitivity at the cost of a >40% false positive rate, whereas model-based prediction achieved higher sensitivity with 10% fewer false positives.

We also investigated the accuracy of MelaPRO and MELPREDICT predictions associated with a carrier probability cutoff of 50%. The positive predictive values were 0.70, 0.57, and 0.44 for MelaPRO-HBI, MelaPRO-GEM, and MELPREDICT, respectively. The negative predictive values were 0.97, 0.97, and 0.90 for the same three models. The mean squared error of prediction, which evaluates the overall performance of the algorithm, was significantly better in the MelaPRO-HBI (0.06; 95% CI, 0.03–0.08) and MelaPRO-GEM (0.08; 95% CI, 0.06–0.11) modules than the MelaPRO-LBI (0.19; 95% CI, 0.15–0.22) module, with the former two slightly better than

### Table 2. Comparison of MelaPRO carrier probability estimates to those provided by MELPREDICT and FH

<table>
<thead>
<tr>
<th>Scenario</th>
<th>MelaPRO GenoMEL-HBI</th>
<th>MelaPRO GenoMEL-LBI</th>
<th>MELPREDICT GEM</th>
<th>FH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) As shown</td>
<td>0.43</td>
<td>0.90</td>
<td>0.85</td>
<td>0.24</td>
</tr>
<tr>
<td>(2) As shown but the paternal aunt is healthy at 40</td>
<td>0.02</td>
<td>0.24</td>
<td>0.18</td>
<td>0.07</td>
</tr>
<tr>
<td>(3) As shown but the father’s status is unknown</td>
<td>0.77</td>
<td>0.93</td>
<td>0.87</td>
<td>0.24</td>
</tr>
<tr>
<td>(4) As shown but the paternal aunt is healthy at 76, the father is MPM at 40</td>
<td>0.82</td>
<td>0.97</td>
<td>0.94</td>
<td>0.24</td>
</tr>
<tr>
<td>(5) As shown but the brother is also SPM at 50</td>
<td>0.91</td>
<td>0.98</td>
<td>0.96</td>
<td>0.40</td>
</tr>
</tbody>
</table>

**NOTE:** The resulting probabilities (for a counselee indicated by the arrow in Fig. 1) correspond to five different variations of the pedigree in Fig. 1, with each variation corresponding to a row in the table. The FH column is yes when the counselee has at least two affected relatives.

### Table 3. Summary of validation results

<table>
<thead>
<tr>
<th></th>
<th>Concordance index C*</th>
<th>Observed/expected ratio†</th>
<th>Mean squared error</th>
</tr>
</thead>
<tbody>
<tr>
<td>MelaPRO-HBI</td>
<td>0.86 (0.75–0.97)</td>
<td>0.85 (0.62–1.08)</td>
<td>0.06 (0.03–0.08)</td>
</tr>
<tr>
<td>MelaPRO-LBI</td>
<td>0.86 (0.74–0.97)</td>
<td>0.31 (0.20–0.42)</td>
<td>0.19 (0.15–0.22)</td>
</tr>
<tr>
<td>MelaPRO-GEM</td>
<td>0.83 (0.67–0.98)</td>
<td>0.54 (0.38–0.72)</td>
<td>0.08 (0.06–0.11)</td>
</tr>
<tr>
<td>MELPREDICT‡</td>
<td>0.82 (0.61–0.93)</td>
<td>1.10 (0.92–1.17)</td>
<td>0.09 (0.04–0.12)</td>
</tr>
<tr>
<td>Difference between MelaPRO-HBI and MelaPRO-LBI</td>
<td>0.007 (–0.01 to 0.02)</td>
<td>Not applicable§</td>
<td>–0.13 (–0.16 to –0.10)</td>
</tr>
<tr>
<td>Difference between MelaPRO-HBI and MelaPRO-GEM</td>
<td>0.04 (–0.008 to 0.09)</td>
<td>Not applicable</td>
<td>–0.03 (–0.04 to –0.02)</td>
</tr>
<tr>
<td>Difference between MelaPRO-HBI and MELPREDICT</td>
<td>0.05 (0.007 to 0.17)</td>
<td>Not applicable</td>
<td>–0.03 (–0.06 to 0.002)</td>
</tr>
</tbody>
</table>

**FH:** ≥2 affected relatives

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.81</td>
<td>0.58</td>
</tr>
</tbody>
</table>

*The concordance index C is equal to the area under the ROC curve.
†The ratio between the observed number of carriers and the total number of predicted carriers.
‡Leave-one-out cross-validation on Boston data using fixed covariates for MELPREDICT.
§Computing the difference in this case is not appropriate, as each ratio should be compared with the reference value of one.
MELPREDICT (0.09; 95% CI, 0.04–0.12; see Table 3). We then considered how often MelaPRO led to a reclassification compared with MELPREDICT and FH. As shown in Table 4, the reclassification fraction ranged from 4% to 34%. MelaPRO-HBI correctly reclassified 5, 10, and 65 more individuals than MelaPRO-GEM, MELPREDICT, and FH, respectively. The 50% threshold was chosen for illustrative purposes only and is not based on any clinical recommendations.

Discussion

One’s ability to create a personalized risk portfolio for patients with hereditary melanoma remains a formidable challenge. To this end, we have developed and successfully validated MelaPRO for individualized CDKN2A carrier estimation. This open source tool delivers a useful and easily deployable instrument for cancer risk counselors who wish to frame a more informative discussion for individuals pursuing CDKN2A genetic testing. Our results indicate that MelaPRO provides high resolution and accurate risk assessment, discriminating between individuals with or without germline mutations in CDKN2A.

An ideal personalized risk model would rely on a menu of modules that best fit the clinical profile. Because geographic location and other unknown genetic factors which may cosegregate with melanoma families seem to influence both the penetrance and prevalence of CDKN2A mutations (4), we constructed three distinct MelaPRO modules based on separate penetrance estimates: GenoMEL-HBI, GenoMEL-LBI, and GEM. This is a first step towards accounting for both genetic and environmental factors. We also derived the corresponding mutation prevalence of CDKN2A based on penetrance. As more data emerges, these estimates can be easily updated, so that the model will continue to operate using the best available information. For example, using allele frequency that is specific to Europe might further improve the performance of GenoMEL-LBI in this population. The prevalence estimate (0.00015) for MelaPRO-HBI matches that from Bishop and colleagues (3). Sensitivity analyses with variation in prevalence (between 0.00015 and 0.0004) showed similar discrimination performance in all modules, and higher observed/expected ratios with lower prevalence values for GEM and LBI. In the Boston validation set, the GenoMEL-HBI presented significantly better performance than others (Table 3), suggesting the utility of geography and ascertainment-specific modules.

MelaPRO treats individuals with SPM and MPM differently, providing higher resolution and more accurate CDKN2A

Figure 2. Receiver operating characteristic curves for MelaPRO-HBI, MelaPRO-LBI, MelaPRO-GEM, and MELPREDICT (using leave-one-out cross-validation) on the Boston validation set. Also shown are the true positive and false positive fractions associated with the summary family history criterion (FH).
Table 4. Comparison of MelaPRO probabilities with MELPREDICT probabilities and FH by reclassification rate

<table>
<thead>
<tr>
<th>Carriers</th>
<th>MelaPRO-HBI</th>
<th>MelaPRO-HBI</th>
<th>MelaPRO-HBI</th>
<th>MelaPRO-HBI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pr ≥ 0.5</td>
<td>Pr &lt; 0.5</td>
<td>Pr ≥ 0.5</td>
<td>Pr &lt; 0.5</td>
</tr>
<tr>
<td>MelaPRO-GEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥0.5</td>
<td>16</td>
<td>1‡</td>
<td>7</td>
<td>6‡</td>
</tr>
<tr>
<td>&lt;0.5</td>
<td>0†</td>
<td>5</td>
<td>0†</td>
<td>160</td>
</tr>
<tr>
<td>MELPREDICT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥0.5</td>
<td>4</td>
<td>0†</td>
<td>3</td>
<td>2†</td>
</tr>
<tr>
<td>&lt;0.5</td>
<td>12‡</td>
<td>6</td>
<td>4†</td>
<td>164</td>
</tr>
<tr>
<td>FH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>16</td>
<td>1‡</td>
<td>7</td>
<td>66‡</td>
</tr>
<tr>
<td>−</td>
<td>0†</td>
<td>5</td>
<td>0†</td>
<td>100</td>
</tr>
</tbody>
</table>

NOTE: MelaPRO-LBI is not included, as the study population does not apply.

*Reclassification rate is defined as the percentage of individuals that were classified differently by MelaPRO-HBI (categories † and ‡) as compared with the other models.

†Values in italics denote a category in which MelaPRO-HBI classifies differently as compared with other methods and carrier status.

‡Denotes a category in which MelaPRO-HBI correctly classifies as compared with other methods.

mutation risk. With the Boston validation data, the MelaPRO-HBI model without MPM adjustment gave higher probabilities to SPM families, where the counselees are often noncarriers. Overall, it gave a lower observed/expected ratio of 0.73, and a slightly lower C index of 0.84. Our assumption of constant MPM/SPM risk ratios across ages can be modified as more data becomes available. Similarly, the current MelaPRO provides the basis for more refined models incorporating polygenic effects, risk modifiers, and biomarkers as their role becomes clarified. Future iterations of MelaPRO will also incorporate two known risk factors: MC1R status and history of pancreatic cancer in the family.

MelaPRO captures the full pedigree data, including information on affected and unaffected family members, and is therefore able to further discriminate between individuals at higher and lower risk with the same number of affected family members. Part of the Boston validation set was used to develop MELPREDICT, specifically for choosing the covariates in the final model. Although the cross-validation should correct for part of the optimism that is associated with internal validations, it does not account for variability across studies. Therefore, the gap between MelaPRO’s and MELPREDICT’s performances would likely be wider in a new independent set of families. The improvement of 0.05 in the concordance index C corresponds to real advances clinically at a personal level, as evidenced by the positive predictive value/negative predictive value and reclassification results. Lastly, MELPREDICT is not applicable to unaffected individuals in melanoma-prone families. MelaPRO is more powerful as a clinical instrument because it is applicable to the entire family and calculates pre-disease estimates of carrier probability and melanoma risk.

In the current study, we built and evaluated MelaPRO on known deleterious mutations whereas excluding known polymorphisms (e.g., Ala148Thr). However, MelaPRO quantifies the degree of genetic segregation in melanoma families and may give high probabilities to carriers of variants of unknown significance that have similar effects to the known variants in CDKN2A. Going forward, MelaPRO can further accommodate errors in classification of variants as deleterious or polymorphic by changing the sensitivity and specificity accordingly. In broader terms, what is critically needed is a robust biochemical or genetic assay for p16/Ink4a and p14/ARF functionality, which will fundamentally improve the accuracy of risk predictions.

From the clinical perspective, MelaPRO can be easily incorporated into any genetic counseling session. Most melanoma clinicians appreciate the importance of family history in a qualitative but not necessarily quantitative sense. However, in clinics, individuals are frequently referred for genetic counseling without regard for pedigree structure or counselee affection status—all features that can be captured for better estimation through MelaPRO. One current thought is that melanoma patients who have either an affected first-degree relative or more than one affected relative on one side of the family, or unaffected individuals with two or more cases of melanoma in close relatives may benefit from a genetic risk assessment. In some situations, MPM patients without family history may also consider counseling. In contrast, unaffected relatives from single-case kindreds who present to their physicians for routine mole checks comprise the largest group of at-risk individuals with a “family history” of melanoma; however, these individuals in general do not need genetic consultation because the likelihood of harboring a germline...
CDKN2A mutation is likely to be close to 1% (18). Likewise, a melanoma patient with a single, distant family history of melanoma, especially if not substantiated by a medical record, would not routinely need genetic risk counseling unless special circumstances exist. Beyond the valuable exercise of counseling, the decision to undergo CDKN2A germline testing should be made in conjunction with a trained professional who can integrate the genetic, psychological, and social implications of genetic testing.

Our ability to assess calibration and discrimination is limited by the 22 CDKN2A mutation carriers found in the validation set. We excluded four carriers, as they were ineligible for MELPREDICT analysis. We also evaluated the performance of MelaPRO alone, using all carriers and obtained similar results. Although suboptimal for analysis, the 11% mutation rate among all individuals with a family history of melanoma is probably appropriate for general clinical use (3, 4, 21). In addition, because most counselees in the validation set were melanoma patients, we could not properly evaluate the model on unaffected individuals. Finally, MelaPRO does not explicitly account for germline CDK4 variants. However, because CDK4 mutations are thought to be exclusive to CDKN2A mutations and because extant data suggest that the CDK4 mutation phenotype is identical to the CDKN2A mutation phenotype (22), MelaPRO assumed CDKN2A and CDK4 as a single genetic unit without resorting to a two-locus model. There were no CDK4 kindreds among our families.

In summary, we have developed and validated a risk prediction model, MelaPRO, whose central goal is to enhance melanoma risk counseling by providing accurate pre-test assessment of CDKN2A carrier probability. MelaPRO’s architecture is of such flexibility that when data become available through ongoing and proposed gene/gene and gene/environment studies, such biological information can be readily assimilated into the model, achieving higher resolution and higher accuracy in risk assessment.

Disclosure of Potential Conflicts of Interest

H. Tsao: consultant/advisory board, SciBASE. The other authors disclosed no potential conflicts of interest.

Grant Support

American Cancer Society grant RSG MGO-112970 (H. Tsao and G. Parmigiani); National Cancer Institute grant nos. P50 CA-93683 (H. Tsao) and R01CA105090-01A1 (G. Parmigiani); and the generous philanthropic donors to the Massachusetts General Hospital.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 7/15/09; revised 10/31/09; accepted 11/16/09; published OnlineFirst 1/12/10.

References

Estimating CDKN2A Carrier Probability and Personalizing Cancer Risk Assessments in Hereditary Melanoma Using MelaPRO


Cancer Res  Published OnlineFirst January 12, 2010.

Updated version Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-09-2653

Supplementary Material Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2010/01/12/0008-5472.CAN-09-2653.DC1

E-mail alerts Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.