A comparison of models used to predict MLH1, MSH2 and MSH6 mutation carriers

C. J. Pouchet1,2*, N. Wong1,3, G. Chong4, M. J. Sheehan5, G. Schneider2, B. Rosen-Sheidley2, W. Foulkes1,6,7 & M. Tischkowitz8,9

1Department of Medical Genetics, SMBD-Jewish General Hospital, Montreal, Canada; 2Genetic Counseling Program, Brandeis University, Waltham, USA; 3Department of Human Genetics, McGill University, Montreal, Canada; 4Molecular Pathology, SMBD-Jewish General Hospital, Montreal, Canada; 5Department of Psychology, Brandeis University, Waltham, USA; 6Departments of Medicine, Human Genetics and Oncology, McGill University; 7Department of Medical Genetics, McGill University Health Centre; 8Departments of Medicine, Human Genetics and Oncology, McGill University; 9Department of Medical Genetics, Oncology and Medicine, SMBD-Jewish General Hospital, Montreal, Canada

Received 18 August 2008; revised 25 September 2008; accepted 26 September 2008

Background: MMRpro, prediction of mutations in MLH1 and MLH2 (PREMM1,2) and MMRpredict are models which were developed to predict the probability that an individual carries a Lynch syndrome-causing mutation. Each model utilizes data from personal and family histories of cancer. To date, no studies have compared these models in a cancer genetics clinic. The purpose of this study was to determine each model’s ability to predict the probability of carrying a Lynch syndrome-causing mutation in individuals with a family history of colorectal cancer and to determine their clinical applicability.

Methods: We obtained family pedigrees from 81 individuals who presented for Lynch syndrome testing due to a personal and/or family history of cancer. Data from each pedigree were entered into the models and analyzed using SPSS.

Results: We found that MMRpredict, PREMM1,2 and MMRpro showed similar performances with areas under the receiver-operating characteristic curve of 0.731, 0.765 and 0.732, respectively. MMRpro showed the least dispersion of mutation probability estimates with a P value of 0.205, compared with 0.034 for PREMM1,2 and 0.001 for MMRpredict.

Conclusion: We found all three carried out well in a cancer genetics setting, with PREMM1,2 giving slightly better estimates. There were some significant discrepancies between the models in cases where the proband had endometrial cancer.

Key words: colorectal cancer, genetic testing, Lynch syndrome, mismatch repair, prediction models

introduction

Hereditary nonpolyposis colorectal cancer, or Lynch syndrome, is the most common of the hereditary colon cancer syndromes, causing between 2% and 5% of colorectal cancer cases. Germline mutations in mismatch repair (MMR) genes, MLH1, MSH2, MSH6, PMS1 and PMS2, are most frequently implicated in the syndrome [1–3]. Individuals with Lynch syndrome have a 50%–80% lifetime risk of developing colorectal cancer, 60%–70% lifetime risk of developing endometrial cancer and <15% lifetime risk of developing other malignancies such as carcinoma of the stomach, small bowel, pancreas, ovary, renal pelvis, transitional cell carcinoma of the ureter, brain tumors and adenomas [1, 4].

In suspected cases, mutation analysis of the MMR genes is typically preceded by microsatellite instability (MSI) and/or immunochemistry (IHC) testing on the tumor block [5]. However, tumors from ~ 5% of patient with Lynch syndrome do not show MSI, and a similar number will have normal IHC staining [6–8]. Moreover, there are many unclassified variants for which the clinical significance is unknown [9]. Studies have also shown that only a fraction of individuals who should be referred for molecular evaluation are actually referred [10]. For these reasons, tools that can help predict the likelihood that an individual carries a Lynch syndrome-causing mutation would be useful in a clinical setting.

There are currently three computer-based models, which estimate the probability that an individual carries a mutation in a MMR gene. These models are MMRpro designed by Chen et al. [11], prediction of mutations in MLH1 and MLH2 (PREMM1,2) designed by Balmana et al. [12] and MMRpredict developed by Barnetson et al. [13]. The purpose of each model is the same; however, each was developed very differently.

MMRpro is based on the data obtained by meta-analyses of population-based studies. It provides an estimate that an
individual carries a germline mutation in MLH1, MSH2 or MSH6, based on the integration of estimates of mutation prevalence and penetrance of MMR genes, colorectal cancer incidence from the Surveillance, Epidemiology, and End Results registry and carrier prevalence among cases reported in the literature with data on the sensitivity and specificity of MSI and germline testing. This model calculates the likelihood that an individual carries a mutation in one of three MMR genes, MLH1, MSH2 and MSH6.

The PREMM1,2 model developed by Balmana et al. [12], is a model based on data supplied on the test submission forms for Myriad Genetics Inc. Information that was used to develop this model included the patient’s age, sex and specific details about personal and family cancer history. This model calculates the likelihood that an individual carries a germline mutation in two MMR genes, MLH1 and MSH2.

The MMRpredict model developed by Barnetson et al. [13] is based on the data from patients who received a colorectal cancer diagnosis under the age of 55, regardless of family history. They tested 870 patients for the presence of a germline mutation in MLH1, MSH2 or MSH6, by MSI testing followed by IHC and gene sequencing. This information was then incorporated into a two-stage model. The group determined which variables, such as age, sex, presence of family history and presence of endometrial cancer, were significant indicators of the presence of a germline mutation. The weight of each predictor was incorporated into an equation in the first stage of the model. Stage 2 of the model includes analysis of the tumor for MSI and IHC. This model calculates the likelihood that an individual carries a germline mutation in two MMR genes, MLH1 and MSH2.

Each of these three models has been validated separately, but has not been compared with each other [12–14]. Recently, the PREMM1,2 and MMRpredict models were compared with each other; however, the MMRpro model was not included in the analyses [15]. In addition, the study used the EPICOLOH cohort and had limited information on family history. Therefore, the aim of this study was to compare the performance characteristics of these models using data from a cancer genetics clinic. We wanted to examine the capacity of each model to identify those individuals at higher risk of carrying an MMR gene mutation and who should subsequently undergo genetic testing.

methods

patients

We obtained anonymized pedigrees of families that had presented clinically for Lynch syndrome testing at the Jewish General Hospital in Montreal, Quebec, Canada, and the Montreal General Hospital in Montreal, Quebec, Canada. Families included were those known to carry a germline mutation in a MMR gene, as well as those with a history of colorectal cancer and Lynch syndrome-related cancers with no known mutation. In total, 81 pedigrees were obtained, 39 in which a Lynch syndrome-causing mutation was identified and 42 in which no known mutation has been identified. When possible, cancer diagnoses in relatives were verified through medical records. We obtained ethical approval from the Institutional Review Board of McGill University, Montreal, Quebec, Canada, as well as approval from the Institutional Review Board of Brandeis University, Waltham, MA, where this study was first initiated as a Master’s thesis.

assessment of genetic risk

Data from each pedigree were entered into each of the three programs being evaluated, MMRpro, PREMM1,2 and MMRpredict. The MMRpro model is available through the CaGen computer program, the PREMM1,2 model available on the Dana-Farber Cancer Institute Web site at http://www.dfci.org/premm and the MMRpredict model available at http://www1.bgu.mrc.ac.uk/So/idata/MMRpredict.php.

When the age at cancer diagnosis was unknown, we estimated it to be 50 years. For cases in which the exact age at cancer diagnosis was unknown, we used a mid-range age, such as age 45 as an estimate. Since some models incorporated information about more relatives than other models (first- and second-degree relatives as opposed to first-degree relatives only), we therefore entered as many affected relatives into each program as possible.

statistical analyses

Receiver-operating characteristic (ROC) curves were used to determine the sensitivity and specificity of each model. We additionally conducted tests to assess the fit of the predicted carrier probabilities to the observed data. Statistical analyses were carried out with SPSS version 15.0. Following convention, all statistical tests were two sided, and statistical significance was based on a $P$ value of <0.05.

results

Of all 81 individuals included in this study, 39 (48%) were determined to have a gene mutation associated with Lynch syndrome. Twenty (51%) of the 39 individuals carried a mutation in the MLH1 gene, 14 (36%) carried a mutation in the MSH2 gene, four (10%) carried a mutation in the MSH6 gene and one (3%) carried a mutation in the PMS2 gene. Table 1 outlines the clinical characteristics of the pedigrees included in this study.

overall performance

All three models showed similar abilities to predict mutation carriers, as demonstrated by the area under the ROC curves in Figure 1 and Table 2. The three models showed similar trends in the ROC curves, suggesting similar performances. The PREMM1,2 showed a slightly larger area under the curve as compared with the other two models, with the upper and lower limits exceeding those of MMRpro and MMRpredict. At one point on the curve, the line for the MMRpredict model drops below the reference line. We are unable to explain this anomaly; however, it is likely attributable to chance and small sample size.

We additionally wanted to compare the ability of the three models to predict the likelihood of a deleterious mutation, as shown in Figure 2 and Table 3. We accomplished this by creating five prediction categories, <10% probability of carrying a mutation, 10%–25%, 25%–50%, 50%–75% and 75%–100% and calculating the percentage of cases in that predicted risk category for which a mutation was identified.

The MMRpredict model was able to predict the probability of carrying a mutation for 74 of 81 individuals. MMRpredict could not calculate the likelihood for seven individuals. This was due to endometrial cancer in the proband without colorectal cancer ($n = 2$), a Lynch syndrome-associated cancer in the proband other than colorectal cancer or endometrial cancer ($n = 1$), polyps in the proband without colorectal cancer
and endometrial cancer and polyps in the proband with colorectal cancer \((n = 3)\). The PREMM\(_{1,2}\) model was able to predict the probability of carrying a mutation for all 81 individuals. The MMRpro model was able to predict the probability of carrying a mutation for 80 of 81 individuals and was unable to calculate the probability for one proband, who had a family history of polyps without cancer.

By creating these risk categories, we could see which model was able to better discriminate carriers from noncarriers and vice versa. One would expect a higher carrier to noncarrier ratio within the high-risk categories (50%–75% and 75%–100%) and a higher noncarrier to carrier ratio in the lower risk categories (<10% and 10%–25%). A perfect model would have only carriers in the highest risk category and no carriers in the lowest. As demonstrated below in Figure 2 and Table 3, all individuals classified within the highest risk category (75%–100%) by the PREMM\(_{1,2}\) model were carriers. Conversely, only approximately half of the individuals classified within the highest risk category by the MMRpro model were in fact carriers. Within the lowest high-risk category (<10%) for the MMRpro model, the majority of individuals who fell into this category were in fact noncarriers, whereas less than half of the individuals within the same category for the MMRpredict model were noncarriers. Given these results, PREMM\(_{1,2}\) was more accurate at detecting carriers, whereas MMRpro was more accurate at detecting noncarriers.

In addition to the overall performance of each model, we wanted to test whether any of the three models were over- or underestimating carriers. Logistic regression analyses, including results of the Hosmer and Lemeshow goodness-of-fit test, which compares observed to expected values using the chi-square distribution, indicated that dispersion was not statistically significant for MMRpro, \(\chi^2(3) = 4.58, P = 0.205\). However, dispersion was statistically significant for both PREMM\(_{1,2}\) \(\chi^2(3) = 8.68, P = 0.034\) and MMRpredict, \(\chi^2(3) = 17.40, P = 0.001\) (data not shown). As demonstrated in Table 4.

### Table 1. Clinical description of pedigrees included in study

<table>
<thead>
<tr>
<th>Clinical characteristic</th>
<th>Mutation positive</th>
<th>Mutation negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pedigrees included</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>Average number of first-degree relatives affected</td>
<td>1.86</td>
<td>1.24</td>
</tr>
<tr>
<td>Range</td>
<td>0–5</td>
<td>0–5</td>
</tr>
<tr>
<td>Average age of CRC onset in proband (years)</td>
<td>44.00</td>
<td>48.59</td>
</tr>
<tr>
<td>Range (years)</td>
<td>21–68</td>
<td>33–75</td>
</tr>
<tr>
<td>Cancers associated with MMR gene defects seen in family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(includes Lynch syndrome, Muir-Torre syndrome and Turcot syndrome)</td>
<td>Brain endometrial kidney/urinary tract ovarian pancreas sebaceous gland stomach</td>
<td>Endometrial kidney/urinary tract ovarian pancreas stomach</td>
</tr>
<tr>
<td>Number of female probands with CRC</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Number of probands with endometrial cancer</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Number of probands with CRC and endometrial cancer</td>
<td>2(^a)</td>
<td>3(^a)</td>
</tr>
<tr>
<td>Number of probands with a Lynch-associated cancer other than CRC and endometrial cancer</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Number of probands with more than one CRC diagnosis</td>
<td>6(^b)</td>
<td>0(^b)</td>
</tr>
</tbody>
</table>

\(^a\)Although the number of probands with both endometrial cancer and CRC was higher in the mutation negative category, it is important to note that the ages of onset for the two cancers were later in this category than the mutation-positive category (endometrial cancer diagnosed at 60 years and CRC diagnosed at 65 years, endometrial cancer diagnosed at 50 years and CRC diagnosed at 50 years, and endometrial cancer diagnosed at 65 years and CRC diagnosed at 67 years, for the three individuals in the mutation-negative category compared to endometrial cancer diagnosed at 47 years and CRC diagnosed at 53 years and endometrial cancer diagnosed at 55 years and CRC diagnosed at 54 years, for the two individuals in the mutation-positive category).

\(^b\)An important distinguishing feature between the groups was that metachronous CRCs only occurred in the mutation-positive probands.

CRC, colorectal cancer.
Table 2. Area under the curve for each model

<table>
<thead>
<tr>
<th>Model</th>
<th>n</th>
<th>Area under the ROC curve (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMRpredict</td>
<td>74</td>
<td>0.731 (0.606–0.857)</td>
</tr>
<tr>
<td>PREMM1,2</td>
<td>81</td>
<td>0.765 (0.654–0.876)</td>
</tr>
<tr>
<td>MMRpro</td>
<td>80</td>
<td>0.732 (0.616–0.847)</td>
</tr>
</tbody>
</table>

This table indicates the area under the ROC curve for each model. n = total number of individuals in each category. For the MMRpredict model, the probability could not be calculated for seven individuals, and the MMRpro model could not calculate the probability for one individual (see text). ROC, receiver-operating characteristic; CI, confidence interval; PREMM1,2, prediction of mutations in MLH1 and MLH2.

Figure 2. Percentage of individuals within each categories that were found to be carriers. The figure shows the percentage of all cases within each predicted risk category, which had identified mutations. The color-coded bars show five prediction categories, and the height of each bar shows the percentage of cases in the category for which a mutation was identified.

We found that MMRpredict underestimated carriers in the <10% carrier probability category and overpredicted carriers between 10% and 75%. PREMM1,2 and MMRpro both showed some dispersion in the <10% and the 10%–25% categories, but results were not largely dispersed in the other prediction categories.

Performance within a family

The statistical tests that we carried out were able to determine the global performance characteristics of each model. However, these tests do not determine whether the three models perform similarly within the same family. When entering data into the three models, we found that with some families, the three models showed discrepancy in their predictions within the same family, and we illustrate this with two examples, one family in which a mutation was identified and one family in which no mutation was identified. We have additionally provided some hypotheses as to why these discrepancies may have occurred.

The first example is illustrated in Figure 3 which shows a mutation-positive pedigree for which the three models showed discrepancy. MMRpredict estimated the lowest probability, 1% that the proband carried an MMR gene mutation. One reason why MMRpredict may have calculated the lowest probability is that the proband does not have any affected first-degree relatives. Also, MMRpredict does not include endometrial cancer as a diagnosis in the proband and therefore could not take this into account. MMRpro estimated the highest probability, 99.9%. This may be in part due to the fact that MMRpro incorporates both endometrial and colorectal cancer in the proband. In addition, MMRpro was able to include the age of onset of all affected individuals. PREMM1,2 estimated the probability to be 60%. This estimate might have been predicted to be between MMRpredict and MMRpro since similarly to MMRpro, PREMM1,2 was able to take into account both the colorectal and endometrial cancer in the proband, as well as the colorectal in the proband’s second-degree relatives. However, similar to the MMRpredict model, PREMM1,2 was not able to include all the ages of onset for colorectal cancer for all cases, but only the youngest case.

The second example is given in Figure 4, which shows a mutation-negative pedigree for which the three models showed discrepancy. MMRpredict estimated the lowest probability, 2% that the proband carried an MMR gene mutation. Similarly to the previously discussed case, the proband had both endometrial and colorectal cancers, and MMRpredict only considered the colorectal cancer. MMRpro estimated the greatest probability, 99.4%. This could be due to the fact that as with the mutation-positive family, MMRpro was able to take into account the colorectal and endometrial cancers in the proband, as well as the colorectal cancer in four other first-degree relatives. PREMM1,2 estimated the probability that the proband carried an MMR gene mutation to be 35%. This is likely because the proband and her first-degree relatives had a later age of onset. Another factor that might account for the low risk estimate is that unlike the previous example this proband did not have any second-degree relatives who were affected.

Prediction of a specific gene

Of the three models, MMRpro is unique in its ability to provide an estimate for the specific MMR gene that is most likely to be mutated. MMRpro provides separate estimates for the likelihood of a mutation in MLH1, MSH2 and MSH6, but does not include a risk of carrying a PMS2 mutation, which is also known, though rarely, to be associated with Lynch syndrome. MMRpro provides the estimate by first determining the probability that an individual carries any Lynch syndrome-associated mutation and then breaking down the probability into three separate probabilities, one for each MMR gene.

In this analysis, we included 37 of 39 individuals known to be carriers in our study sample. MMRpro could not evaluate a proband with a family history of polyps without cancer, and the second proband was excluded due to the presence of a PMS2 mutation. We compared the predictions generated by
MMRpro to the actual genetic test results for those individuals who had an identifiable mutation (see Table 5) and found that MMRpro was capable of correctly predicting the specific gene mutations for 13 of 37 individuals (35% accuracy).

**discussion**

We have examined three models, which are currently used to determine the probability that an individual carries a germline MMR gene mutation, based on their personal and family history of cancer. All three models evaluated gave comparable areas under the ROC curve, as demonstrated in Figure 1, suggesting that overall they are very similar in the way they rank individuals according to their carrier probability. The PREMM1,2 model did have the greatest area under the curve (0.765) as compared with MMRpredict (0.731) and MMRpro (0.732), suggesting that the overall performance of PREMM1,2 is slightly better than that of the other two models. We also found that MMRpredict showed significant dispersion in under- and/or overestimating carriers at a \( P \) value of 0.01, while the PREMM1,2 model showed marginally significant dispersion with a \( P \) value of 0.034. These data would further suggest that the PREMM1,2 model may have a better overall performance than the other two models.

**Table 3.** Number of individuals who were in each probability category

<table>
<thead>
<tr>
<th>Model</th>
<th>Predicted probability (%)</th>
<th>Number of individuals</th>
<th>Number of carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(total)</td>
<td>(carriers)</td>
<td>Observed (O)</td>
</tr>
<tr>
<td>MMRpredict</td>
<td>&lt;10%</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>10–25%</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>25–50%</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>50–75%</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>75–100%</td>
<td>26</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>74</td>
<td>36</td>
</tr>
<tr>
<td>PREMM1,2</td>
<td>&lt;10%</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>10–25%</td>
<td>26</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>25–50%</td>
<td>23</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>50–75%</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>75–100%</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>81</td>
<td>39</td>
</tr>
<tr>
<td>MMRpro</td>
<td>&lt;10%</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>10–25%</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>25–50%</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>50–75%</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>75–100%</td>
<td>50</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>80</td>
<td>38</td>
</tr>
</tbody>
</table>

This table shows the actual number of individuals who were predicted by the various models to be within each category. ‘Total’ designates the number of all individuals included in the study that fell into each category, while ‘carriers’ designate only the MMR gene mutation carriers in that category. ‘Not possible’ indicates that the particular model could not calculate the probability that the individual carried a mutation (see text). PREMM1,2, prediction of mutations in MLH1 and MLH2.

**Table 4.** Observed and expected numbers of mutation carriers as predicted by the different models, for categories based on each model’s predicted range of probability of being a carrier

<table>
<thead>
<tr>
<th>Model</th>
<th>Predicted probability (%)</th>
<th>Number of individuals</th>
<th>Number of carriers</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMRpredict</td>
<td>&lt;10%</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>10%–25%</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>25%–50%</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>50%–75%</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>75%–100%</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>81</td>
<td>39</td>
</tr>
</tbody>
</table>

This table shows the observed and expected values in each probability category for each of the three models. The expected − observed/observed values suggest whether a model is over- or underestimating carriers. A highly positive value would suggest that a model is overestimating carriers, whereas a highly negative value would suggest that a model is underestimating carriers. N/A indicates that the result was an undefined value, which could not be quantified. PREMM1,2, prediction of mutations in MLH1 and MLH2.
In general, the models had much better sensitivity than specificity. This may in part be due to the fact that those people who present clinically have a strong family history of cancer, which would suggest a cancer syndrome, making it more difficult to differentiate between those who do and do not carry a mutation. The strengths and weaknesses of each model are summarized in Table 6. Although this study was conducted with a relatively small sample size ($n=81$), the diverse population included in it is representative of the patient population presenting at many urban health-care centers.

Genetic risk evaluation of colon cancer families can be complicated by a number of confounding factors, including small family size, incomplete penetrance, variable expressivity, familial factors including environment and genes, as well as a general population risk for sporadic cancers. The models compared in this study are useful tools for health professionals involved in risk evaluation to navigate these complexities. Although these models are quite helpful for health professionals in identifying families that are at high risk of carrying a Lynch syndrome-causing mutation, those using these models should be cautioned that they are not perfect. It is important to appreciate that certain models will work better under specific circumstances. For example, MMRpro should not be used in cases where there is a strong family history of Lynch syndrome-related cancers other than colorectal and endometrial cancers since this model is unable to integrate the presence of these cancers in its risk estimates. However, this model might be used when there are many first- and second-degree relatives affected with colorectal and/or endometrial cancers. Table 6 provides a summary of the strengths, weaknesses and utilities of each model that we identified in our patient group.
The identification of individuals at high risk of carrying an MMR gene mutation is important clinically as studies have shown that screening and prophylactic surgery in individuals who carry an MMR gene mutation, or who are at increased risk without an identified mutation, greatly decreases the incidence of cancer and consequently mortality in those individuals [16, 17]. We have shown in this study that while the overall performance of the models is broadly similar, important discrepancies can arise with particular clinical scenarios and we recommend that users are fully familiar with strengths and limitations of the models to ensure that individuals from colon cancer families receive the most appropriate care and management.

**Table 6.** Summary of MMRpredict, PREMM1,2 and MMRpro

<table>
<thead>
<tr>
<th></th>
<th>MMRpro</th>
<th>PREMM1,2</th>
<th>MMRpredict</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takes into account colorectal cancer</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Takes into account endometrial cancer</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Takes into account other HNPCC-related cancers</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Takes into account previous MSI test results</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Takes into account previous IHC test results</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Takes into account previous genetic testing</td>
<td>Yes</td>
<td>No</td>
<td>Yes, but does not change probabilities</td>
</tr>
<tr>
<td>Takes into account age of proband and other affected individuals</td>
<td>Takes into account ages of affected individuals</td>
<td>Only takes into account youngest age of onset</td>
<td>Evaluates based on less than or greater than 50 years old</td>
</tr>
<tr>
<td>Takes into account degree of relationship to proband</td>
<td>Yes</td>
<td>Yes; first- and second-degree relatives assessed separately</td>
<td>Yes; but only takes into account first-degree relatives of proband</td>
</tr>
<tr>
<td>Provides estimates for mutations in</td>
<td>MLH1, MSH2 and MSH6; relies on estimations of mutation prevalence and penetrance by age and sex</td>
<td>MLH1 and MSH2 only</td>
<td>MLH1 and MSH2 only</td>
</tr>
<tr>
<td>Strengths</td>
<td>Calculates risk of three HNPCC-causing genes; provides separate risks for all three genes; can recalculate probability of finding a mutation if MSI testing is negative; can calculate residual probability of finding a mutation when germline testing has found none</td>
<td>Takes into account individuals with HNPCC-related cancers; distinguishes between first- and second-degree relatives; user-friendly and easily accessible online; takes into account more than one colorectal cancer in the proband or relative; takes into account polyps in proband</td>
<td>Takes into account whether proband is male or female; takes into account tumor location; takes into account if proband has been diagnosed with more than one CRC</td>
</tr>
<tr>
<td>Weaknesses</td>
<td>Does not take into account other HNPCC-related tumors other than colorectal and endometrial; does not take into account more than one colorectal cancer or polyps in the proband or family members; need a first-degree relative affected in order to calculate risk to proband; when developing model, only high-risk families were looked at to derive data</td>
<td>Does not take into account previous test results (genetic or MSI); provides only a combined estimate of a mutation in either MLH1 or MSH2; data collected through myriad genetics forms and relies on accuracy of information provided; does not take into account if proband is male or female, but takes into account whether or not proband has an endometrial cancer diagnosis</td>
<td>Tends to overestimate and underestimate probabilities by calculating risk at 100% or 0%; model has not been validated in patients &gt;55 years old; model does not take into account ages of family members</td>
</tr>
<tr>
<td>Utility</td>
<td>Good for calculating risk of a mutation when both MSI testing and germline testing are negative and strong suspicion of HNPCC remains</td>
<td>Useful in families that have HNPCC-related cancers in addition to colorectal endometrial cancer</td>
<td>Useful for individuals with early-onset CRC without any HNPCC-associated cancers</td>
</tr>
</tbody>
</table>

PREMM1,2, prediction of mutations in MLH1 and MLH2; CRC, colorectal cancer; HNPCC, hereditary nonpolyposis colorectal cancer.

**References**
