Appendix to "Prediction of Germline Mutations and Cancer Risk in the Lynch Syndrome"

Sining Chen[1], Wenyi Wang[1], Shing Lee[2], Khedoudja Nafa[3], Johanna Lee[3], Kathy Romans[4], Patrice Watson[5], Stephen B. Gruber[6], David Euhus[7], Kenneth W. Kinzler[4], Jeremy Jass[8], Steven Gallinger[9], Noralane Lindor[10], Graham Casey[11], Nathan Ellis[12], Francis M. Giardiello[4], the Colon Cancer Family Registry, Kenneth Offit[3], Giovanni Parmigiani[1,4]

1. Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
2. Columbia University Mailman School of Public Health, New York, NY
3. Memorial Sloan-Kettering Cancer Center, New York, N.Y.
4. Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD
5. Creighton University School of Medicine, Omaha, NE
6. University of Michigan Medical School, Ann Arbor, MI
7. UT Southwestern Medical Center, Dallas, TX
8. Department of Pathology, McGill University, Montreal, Canada
9. University of Toronto, Toronto, Canada
10. Mayo Clinic College of Medicine, Rochester, Minnesota.
11. Cleveland Clinic Lerner College of Medicine, Cleveland, OH
12. Department of Medicine, University of Chicago, Chicago, IL

Corresponding author:
Sining Chen, PhD
Department of Environmental Health Sciences &
Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St.,
Room W7033A
Baltimore, MD 21205
Tel: 410-502-4266
FAX: 410-955-1811
Email: sichen@jhsph.edu

Penetrance and Prevalence Parameters

The age-specific incidence of colorectal and endometrial tumors among *MLH1* and *MSH2* carriers has been reported by a number of studies. Some studies ascertained families with multiple cancer cases, without proper correction of ascertainment bias, they are vulnerable to upward biases in the risk estimate. Thus, we estimated the cancer risk via a meta-analysis using only the population-based studies [1,2,3] and one additional analysis that adjusted for ascertainment [4]. The cancer incidence among population-based *MSH6* carriers was only recently studied[3,5]. We abstract *MSH6* penetrance from those reports. We derived mean cumulative risks by 10-year age intervals. At each interval, the mean is a weighted average of risks from applicable studies, the weights are calculated according to the widths of study-reported confidence intervals. For cancer risks among non-carriers, we use the SEER registry, which publishes authoritative and comprehensive cancer incidence data from 11 population-based registries throughout the US [6].

We summarize the results of our meta-analysis of penetrance in Figure A1. These results provide the default values used in the current version of MMRpro and all analyses in this article. These penetrance curves are also used to calculate the probability of developing CRC and EC for unaffecteds.

To test the robustness of MMRpro performance with respect to different penetrance, we re-did the validation by calculating MMRpro probabilities after increasing the penetrance by 50%, which gives a risk level comparable to that reported on high-risk families, without ascertainment correction. The validation results are nearly identical. The AUC for the higher penetrance scenario is 0.82 (0.77, 0.87) with MSI, and 0.79 (0.74, 0.85) without MSI. The O/E ratio is 1.01 (0.89, 1.13) with MSI, and 1.05

(0.92, 1.19) without. Therefore it is reasonable to infer that the performance of MMRpro is stable over a range of plausible penetrance choices.

As for prevalence, mutations on MMR genes are rare in the general population and the frequency is not known with certainty. We estimated it via the following relationship

$$\text{carrier prevalence} = \frac{\text{carrier prevalence among cases} \cdot \text{cancer incidence}}{\text{cancer incidence among carriers}}$$

Estimates of all quantities on the right hand side can be found in the literature. We consider only colorectal cancers diagnosed before age 50. A population-based estimate of the prevalence of *MLH1* and *MSH2* carriers among CRC cases younger than 50 is available[7]. After accounting for the sensitivity of the germline testing techniques used, the prevalence among younger cases is estimated to be 0.28. The overall cancer incidence by age 50 in the general population is 0.00215, from the SEER registry. The CRC risk among carriers of *MLH1* and *MSH2* (by age 50) is 0.32 according to the penetrance estimates reported earlier. The resulting carrier frequency is 0.0019 for *MLH1* and *MSH2* combined, or 0.0009 and 0.0010 for the two genes individually. Less information is available on *MSH6* mutation prevalence. In MMRpro we assume that it accounts for 15% of all HNPCC mutations, that is, 0.00036. This prevalence may be population specific [8, 9,10]

Sensitivity and specificity of MSI Testing

Results of MSI testing of tumors in any family member are incorporated into the Mendelian calculation by considering them part of an individual's phenotype information. This enters the calculations via the "probability of tumor characteristic given the individual's genotype", which in

this case is a function of the sensitivity or specificity of the tumor markers in predicting a germline mutation.

MSI-h tumor results are predictive of germline mutation in MMR genes[10]. Lynch et al.[11] suggested that about 90% of HNPCC tumors are MSI. A meta-analysis of 16 published studies[12-26] estimated the sensitivity at 81% (73% – 89%) for *MLH1* or *MSH2* [27]. The specificity is estimated at 92% (90% – 94%). The imperfect specificity occurs because sporadic colorectal cancers can display MSI from somatic hyper-methylation of the *MLH1* promoter and potentially other factors. Although these could result in differential specificity of MSI for *MLH1* and *MSH2*, we currently do not distinguish between the two. The positive predictive value of MSI in low-risk families can be low. Literature reports suggest that tumors due to *MSH6* mutations are less likely to show MSI [3].Specifically, we use a sensitivity of 73% derived from Hendriks and colleagues [28].

Abnormal IHC is reported to be highly correlated with MSI [29-31], with the added value of predicting the gene in which the mutation resides. It should be noted that the specificity of IHC and MSI may decrease with age, as somatic hyper-methylations accumulate. However, age-specific specificities are not implemented in the current version.

Sensitivity of conventional mutation detection techniques

Currently, the most widely used germline testing technique is direct sequencing. This, along with other conventional mutation detection assays, such as protein truncation[32], conformation sensitive gel electrophoresis and allele-specific oligonucleotide testing [33], direct sequencing can miss large genome rearrangements, deletions and duplications, which constitute a significant portion of MMR gene mutations. In the meta-analysis of 16 published studies, using the Hui and Walter approach[34], by grouping the tested population into a high-risk group and a low risk group, tabulating mutation

analysis results against MSI result, we estimate that these mutation analysis techniques have a sensitivity of 62% (56-68%) in detecting pathogenic mutations [35]. This estimate is consistent with evidence arising from recent technologies such as conversion, which revealed a large number of mutations not detected before[27,36,37,38.]

After specifying sensitivity and specificity (the specificity of direct sequencing is assumed to be 100%), the results of germline testing of any member within the family can be incorporated in the same manner as MSI results. In practice, an individual with a high pre-germline test probability of being a carrier, who has an inconclusive sequencing test may still harbor a mutation missed by sequencing, and, thus, often has a high post-test carrier probability.

References

[1] Dunlop MG, Farrington SM, Carothers AD, Wyllie AH, Sharp L, Burn J, Liu B, Kinzler KW, Vogelstein B. Cancer risk associated with germline DNA mismatch repair gene mutations. *Hum Mol Genet* 1997;6:105

[2] Hampel H, Stephens JA, Pukkala E, Sankila R, Aaltonen LA, Mecklin JP, de la Chapelle A. Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. Gastroenterology 129(2):415-21 2005

[3] Mark A Jenkins, Laura Baglietto, James G Dowty, Christine M Van Vliet, Letitia Smith, Leeanne J Mead, Finlay A Macrae, D. James B St John, Jeremy R Jass, Graham G Giles, John L Hopper, Melissa C Southey, Cancer risks for mismatch repair gene mutation carriers: a population-based early onset case-family study. *Clin Gastroenterol Hepatol*,4, 489-498 2006

[4] Quehenberger F, Vasen HFA, van Houwelingen HC. Risk of colorectal and endometrial cancer for carriers of mutations of the h*MLH1* and h*MSH2* gene: correction for ascertainment. *J Med Genet* 42; 491–496 2005.

[5] Buttin B, Powell M, Mutch D, et al. Penetrance and expressivity of *MSH6* germline mutations in seven kindreds not ascertained by family history. *American Journal of Human Genetics* 74(6); 1262–9 2004.

[6] National Cancer Institute: Surveillance, Epidemiology, and End Results (SEER) Program. 1997.

[7] Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med* 338(21); 1481–87 1998.

[8] Offit K. *MSH6* mutations in hereditary nonpolyposis colon cancer: another slice of the pie. *J Clin Oncol* 22(22); 4449–51 2004.

[9] Sanchez de Abajo A, de la Hoya M, van Puijenbroek M, Godino J, Díaz-Rubio E, Morreau H and Caldes T. Dual role of LOH at MMR loci in hereditary non-polyposis colorectal cancer?Oncogene (2006) 25, 2124-2130.

[10] Sanchez de Abajo A, de la Hoya M, Tosar A, Godino J, Fernandez JM, Asenjo JL, Villamil BP, Segura PP, Diaz-Rubio E, Caldes T. Low prevalence of germline h*MSH6* mutations in colorectal cancer families from Spain. World Journal of Gastroenterology 11(37);5770-5776 2005.

[11] Lynch HT, Lynch PM. Molecular screening for the Lynch syndrome–better than family history? *N Engl J Med* 352; 1920–1922 2005.

[12] Bapat B, Madlensky L, Temple L, et al. Family history characteristics, tumor microsatellite instability and germline *MSH2* and *MLH1* mutations in hereditary colorectal cancer. *Human Genetics* 104; 167–76 1999.

[13] Calistri D, Presciuttini S, Buonsanti G, et al. Microsatellite instability in colorectal cancer patients with suspected genetic predisposition. *Int J Cancer* 89; 89–91 2000.

[14] Debniak T, Kurzawski G, Gorski B, Kladny J, Domagala W, Lubinski J. Value of pedigree/clinical data, immunohistochemistry and microsatellite instability analyses in reducing the cost of determining h*MLH1* and h*MSH2* gene mutations in patients with colorectal cancer. *Eur J Cancer* 36; 49–54 2000.

[15] Dieumegard B, Grandjouan S, Sabourin JC, et al. Extensive molecular screening for hereditary non-polyposis colorectal cancer. *Br J Cancer* 82; 871–80 2000.

[16] Lamberti C, Kruse R, Ruelfs C, et al. Microsatellite instability-a useful diagnostic tool to select patients at high risk for hereditary non-polyposis colorectal cancer: a study in different groups of patients with colorectal cancer. *Gut* 44; 839–43 1999.

[17] Liu T, Wahlberg S, Burek E, Lindblom P, Rubio C, Lindblom A. Microsatellite instability as a predictor of a mutation in a DNA mismatch repair gene in familial colorectal cancer. *Genes Chromosomes Cancer* 27; 17–25 2000.

[18] Scartozzi M, Bianchi F, Rosati S, et al. Mutations of *hMLH1* and *hMSH2* in patients with suspected hereditary nonpolyposis colorectal cancer: correlation with microsatellite instability and abnormalities of mismatch repair protein expression. *J Clin Oncol* 20; 1203–8 2002.

[19] Cederquist K, Golovleva I, Emanuelsson M, Stenling R, Gronberg H. A population based cohort study of patients with multiple colon and endometrial cancer: correlation of microsatellite instability (MSI) status, age at diagnosis and cancer risk. *Int J Cancer* 91; 486–91 2001.

[20] Ponz de Leon M, Benatti P, Di Gregorio C, et al. Genetic testing among high-risk individuals in families with hereditary nonpolyposis colorectal cancer. *Br J Cancer* 90; 882–7 2004.

[21] Terdiman JP, Gum J. R. J, Conrad PG, et al. Efficient detection of hereditary non-polyposis colorectal cancer gene carriers by screening for tumor microsatellite instability before germline genetic testing. *Gastroenterology* 120; 21–30 2001.

[22] Wang Y, Friedl W, Lamberti C, et al. HNPCC: frequent occurrence of large genomic deletions in *MSH2* and *MLH1*. *Int J Cancer* 5 2003.

[23] Percesepe A, Borghi F, Menigatti M, et al. Molecular screening for hereditary non-polyposis colorectal cancer: a prospective, population-based study. *J Clin Oncol* 19; 3944–50 2001.

[24] Salahshor S, Kressner U, Fischer H, et al. Microsatellite instability in sporadic colorectal cancer is not an independent prognostic factor. *Br J Cancer* 81; 190–3 1999.

[25] Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, Ruschoff J. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. *Cancer Res* 57; 4749–56 1997.

[26] Salovaara R, Loukola A, Kristo P, et al. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 18; 2193–200 2000.

[27] Liu B., Parsons R., Papadopoulos N., Nicolaides N.C., Lynch H.T., Watson P., Jass J.R., Dunlop M., Wyllie A., Peltomäki P., de la Chapelle A., Hamilton S.R., Vogelstein B. and Kinzler K.W. Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat Med* 1996 2 169-174

[28] Hendriks Y, Wagner A, Morreau H, et al. Cancer risk in hereditary nonpolyposis colorectal cancer due to *MSH6* mutations: impact on counseling and surveillance. *Gastroenterology* 127(1); 17–25 2004.

[29] Marcus VA, Madlensky L, Gryfe R, et al. Immunohistochemistry for *hMLH1* and *hMSH2*: a practical test for DNA mismatch repair-deficient tumors. *Am J Surg Pathol* 23; 1248–1255 1999.

[30] Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol* 20; 1043– 1048 2002.

[31] Engel C, Forberg J, Holinski-Feder E, et al. Novel strategy for optimal sequential application of clinical criteria, immunohistochemistry and microsatellite analysis in the diagnosis of hereditary nonpolyposis colorectal cancer. *Int J Cancer* 118; 115–122 2006.

[32] Powell SM, Petersen GM, Krush AJ, et al. Molecular diagnosis of familial adenomatous polyposis. *N Engl J Med* 329(27); 1982–1987 1993.

[33] Wahlberg S, Liu T, Lindblom P, Lindblom A. Various mutation screening techniques in the DNA mismatch repair genes *hMSH2* and *hMLH1*. *Genet Test* 3(3); 259–264 1999.

[34] Hui SL and Walter SD, Estimating the error rates of diagnostic tests. *Biometrics* 36, 167-171, 1980

[35] Chen S, Watson P, Parmigiani G. Accuracy of MSI testing in predicting germline mutations of *MSH2* and *MLH1*: a case study in Bayesian meta-analysis of diagnostic tests without a gold standard. *Biostatistics*, 6(3) 2005

[36] Charbonnier F, Raux G, Wang Q, Drouot N, Cordier F, Limacher JM, Saurin JC, Puisieux A, Olschwang S, Frebourg T. Detection of exon deletions and duplications of the mismatch repair genes in hereditary nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. Cancer Res. 2000 Jun 1;60(11):2760-3.

[37] Nakagawa H, Hampel H, de la Chapelle A. Identification and characterization of genomic rearrangements of *MSH2* and *MLH1* in lynch syndrome (HNPCC) by novel techniques. *Hum Mutat* 22; 258 2003.

[38] Renkonen E, Zhang Y, Lohi H, et al. Altered expression of *MLH1*, *MSH2*, and *MSH6* in predisposition to hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 21; 3629–37 2003.

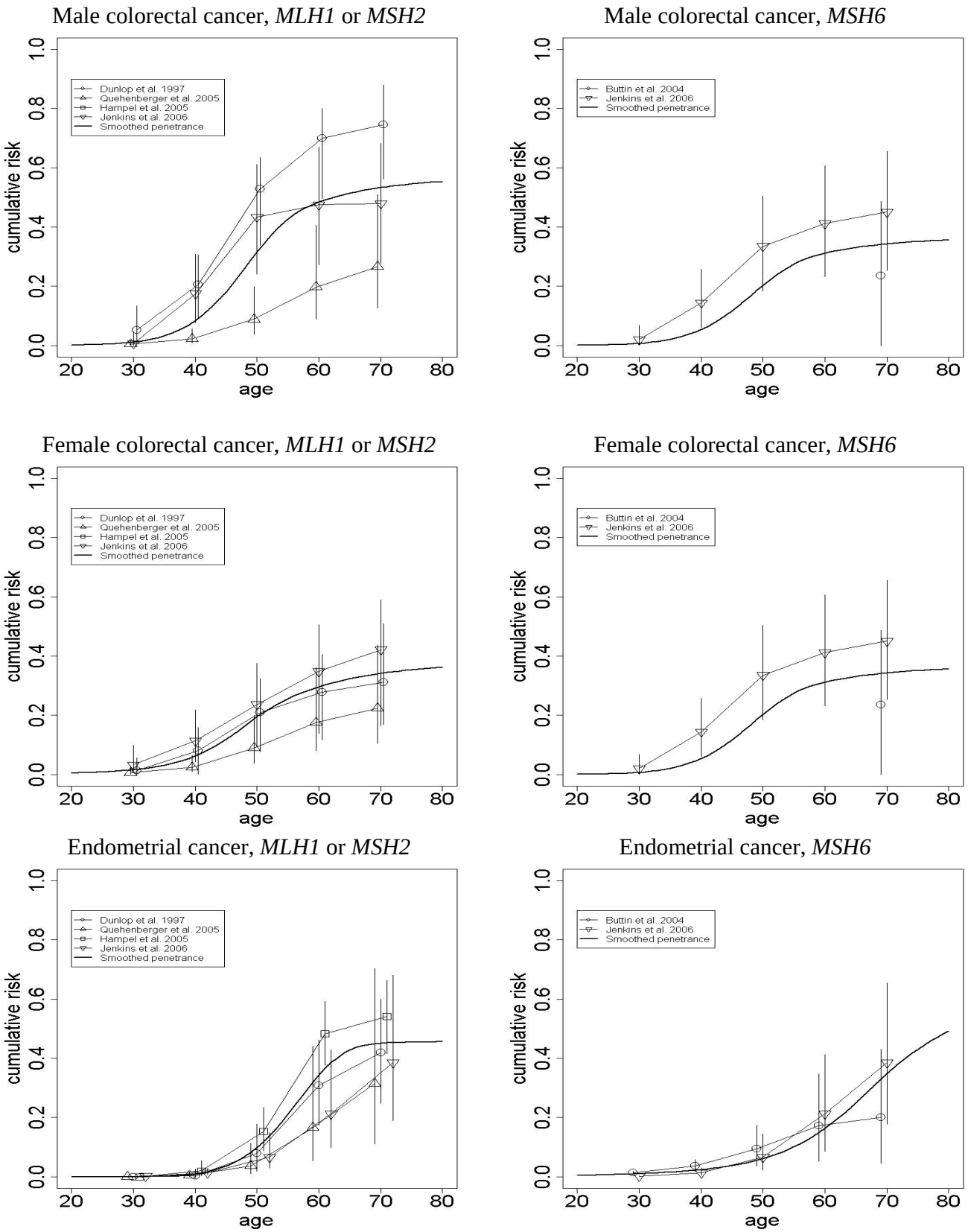

Figure A1: Default penetrance curves used in MMRpro for *MLH1*, *MSH2* and *MSH6* carriers. Estimates are derived by a meta-analyses of population-based or ascertainment-adjusted published results: the penetrance for *MLH1* and *MSH2* carriers is based on Dunlop et al.[1], Quehenberger et al.[4], Hampel et al.[2] and Jenkins et al.[3], penetrance for *MSH6* is based on Buttin et al.[17]