

Preliminary validation of a consumer-oriented colorectal cancer risk assessment tool compatible with the US Surgeon General's My Family Health Portrait

W. Gregory Feero, MD, PhD¹, Flavia M. Facio, MS, CGC², Emily A. Glogowski, MSc, CGC³, Heather L. Hampel, MS, CGC⁴, Jill E. Stopfer, MS, CGC⁵, Haley Eidem, BS⁶, Amy M. Pizzino, MS, CGC⁷, David K. Barton, MS⁸ and Leslie G. Biesecker, MD⁹

Purpose: This study examines the analytic validity of a software tool designed to provide individuals with risk assessments for colorectal cancer based on personal health and family history information. The software is compatible with the US Surgeon General's My Family Health Portrait (MFHP).

Methods: An algorithm for risk assessment was created using accepted colorectal risk assessment guidelines and programmed into a software tool (MFHP). Risk assessments derived from 150 pedigrees using the MFHP tool were compared with "gold standard" risk assessments developed by three expert cancer genetic counselors.

Results: Genetic counselor risk assessments showed substantial, but not perfect, agreement. MFHP risk assessments for colorectal cancer yielded a sensitivity for colorectal cancer risk of 81% (95% confidence interval: 54–96%) and specificity of 90% (95% confidence interval:

83–94%), as compared with genetic counselor pedigree review. The positive predictive value for risk for MFHP was 48% (95% confidence interval: 29–68%), whereas the negative predictive value was 98% (95% confidence interval: 93–99%). Agreement between MFHP and genetic counselor pedigree review was moderate ($\kappa = 0.54$).

Conclusion: The analytic validity of the MFHP colorectal cancer risk assessment software is similar to those of other types of screening tools used in primary care. Future investigations should explore the clinical validity and utility of the software in diverse population groups.

Genet Med advance online publication 18 December 2014

Key Words: colorectal cancer; family history; My Family Health Portrait; risk assessment tools

Patient-completed electronic family history (FH) collection tools have been proposed as a way to overcome several of the barriers to collection and use of FH information in health-care settings.^{1,2} Several patient-completed tools designed to aid in the collection and interpretation of FH have been developed and validated, often in settings such as cancer and primary-care clinics.^{3–5} A 2009 review of FH literature conducted for the US Agency for Health Care Research and Quality and the National Institutes of Health identified a paucity of validated FH tools suitable for use in primary-care or public health contexts.⁶ In 2014, the National Colorectal Cancer Roundtable found that electronic health record systems are poorly prepared to store and interpret FH relevant to colorectal cancer (CRC) risk assessment.⁷

The US Centers for Disease Control and Prevention-sponsored Family Healthware Trial investigated behavioral change related to the clinical use of a patient-completed electronic family history tool (FHT) that provided risk assessment for several conditions, including type 2 diabetes, heart disease, stroke, and colorectal, breast, and ovarian cancer.⁸ The FHT was not designed for stand-alone CRC risk assessment, but

extensive expertise was brought to bear in its development, suggesting that the FHT may be a useful point of comparison for the evaluation of other automated CRC risk assessment tools. The public version of the My Family Health Portrait (MFHP) tool was developed and launched in 2004 as part of the US Surgeon General's Family History Initiative. The tool was designed to help the public collect, organize, and share FH information with relatives and health professionals.⁹ The current version of the MFHP was designed to collect FH in a format compatible with electronic health record systems using nomenclature and data standards for storing and sharing FH; the MFHP has been studied in several settings and is freely available to the public.^{10–13}

An important limitation of the current public MFHP tool is that it does not provide users with feedback or education regarding their individual disease risk. Here, we report on the development of a CRC risk assessment module compatible with the public MFHP tool and, potentially, electronic health record systems. Moreover, we present data regarding the analytic validity of the new module and the FHT according to the analytical

¹Maine-Dartmouth Family Medicine Residency, Fairfield, Maine, USA; ²Inova Translational Medicine Institute, Falls Church, Virginia, USA; ³Clinical Genetics Service, Memorial Sloan-Kettering Cancer Center, New York, New York, USA; ⁴Division of Human Genetics, The Ohio State University, Columbus, Ohio, USA; ⁵Abramson Cancer Center, University of Pennsylvania, Philadelphia, Pennsylvania, USA; ⁶Department of Genetics, Vanderbilt University, Nashville, Tennessee, USA; ⁷White Matter Disease Program, Children's National Health System, Washington, DC, USA; ⁸Department of Biomedical Engineering, University of Michigan, Ann Arbor, Michigan, USA; ⁹Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA. Correspondence: W. Gregory Feero (wfeero@mainegeneral.org)

Submitted 28 August 2014; accepted 5 November 2014; advance online publication 18 December 2014. doi:10.1038/gim.2014.179

validity; clinical validity; clinical utility; and ethical, legal, and social issues (ACCE) framework proposed as a method for evaluating family health history and other genomic health tools.¹⁴

MATERIALS AND METHODS

Creation of the CRC risk module

The MFHP tool output is a structured XML output file that is computable for risk assessment. An algorithm based on National Comprehensive Cancer Network and United States Preventive Services Task Force CRC risk assessment guidelines was developed to dichotomize individuals into “elevated” and “not elevated” risk categories based on data entered into the MFHP tool (**Supplementary Appendix S1** online). To avoid inducing potentially inappropriate distress in individuals without benefit of immediate access to evaluation by a health professional, individuals at “elevated” risk were not further subdivided into “high” and “moderate” risk categories by the MFHP tool. The model did not incorporate other data elements, such as proband gender, body mass index, or self-identified race, because there were no evidence-based guidelines for assigning risk for these factors in conjunction with FH data. The final algorithm was reviewed by a group of cancer risk assessment experts prior to being programmed into an open-source software tool. This input resulted in the final decision to provide a dichotomous rather than a more quantitative risk assessment and to consider all reported polyps as potential sources of risk warranting further discussion with a health-care provider. The output of the MFHP tool is risk-specific CRC information printable as letters for public users and their health-care providers.

Study population and risk algorithm validation

We used 150 patient-entered pedigrees sequentially derived from the ClinSeq cohort as the substrate for assessing the MFHP tool and the US Centers for Disease Control and Prevention FHT CRC module’s analytic performance.¹⁵ The cohort consists of self-selected volunteers with an interest in learning their personal genomic risks through genome or exome sequencing; individuals entered the study without investigator foreknowledge of cancer FHs. The cohort is more likely to be white, well-educated, and affluent than the general US population. Forty additional patient-entered pedigrees derived from the cohort served as internal controls; 20 were deliberately modified to increase CRC risk to different levels (10 “strong” risk and 10 “moderate” risk) by adding additional accepted risk factors to the pedigree data.

Three different methods were used to derive risk estimates from the test pedigrees: the MFHP tool (including the new CRC risk module, referred to as MFHP hereafter); the Centers for Disease Control and Prevention FHT (1.0 Beta Version for Research Studies); and independent review by three genetic counselors (GCs) with expertise in CRC risk assessment. The GCs were blinded to which pedigrees had been modified to alter risk, the risk assessments made by the MFHP tool, the FHT, each other’s risk assessments, and proband self-reported body-mass index and race. The GCs were instructed that they should categorize the proband’s CRC risk as “weak,” “moderate,”

or “strong” according to the criteria outlined by Yoon et al.⁸ This study was approved by the National Human Genome Research Institute Institutional Review Board.

Statistical analysis

Review of pedigrees by a panel of three expert cancer GCs was defined as the gold standard for risk of CRC in the test pedigrees.¹⁶ Risk estimates of “weak” by the FHT and GC review were considered “not elevated,” whereas “moderate” and “high” risk assignments were considered to be “elevated.” Pedigrees were scored as showing “elevated” risk for CRC only when at least two of three counselors reported risk levels as “moderate” or “strong” (true positives). The κ statistics were generated to assess GC interrater agreement for CRC assessment using the “weak,” “moderate,” and “strong” risk estimates for all 190 pedigrees. For the calculations of sensitivity, specificity, positive predictive values (PPVs), negative predictive values (NPVs), and κ statistics for the presence of CRC risk for the MFHP tool and the FHT versus the GC review, only the 150 unmodified pedigrees were used.

RESULTS

GC risk estimates for CRC in the ClinSeq pedigrees

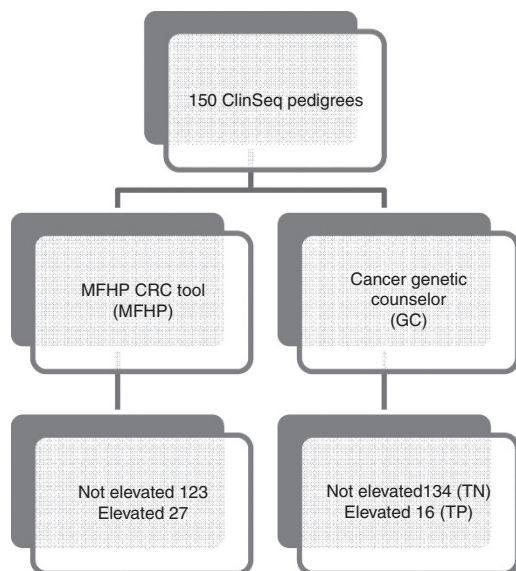
Sixteen of the 150 (11%) unmodified pedigrees were identified as having “elevated” risk by at least two of three GCs rating risk as “moderate” or “strong”; 6 of 150 (4%) were identified as having “strong” risk by at least two of three counselors. For the modified internal control pedigrees, the three counselors agreed for 10 of 10 (100%) and 6 of 10 (60%) pedigrees intended to be at “strong” or “moderate” risk, respectively. Pair-wise κ statistics for the three counselors calculated for all 190 pedigrees were 0.760, 0.776, and 0.851, indicating substantial agreement between the GCs according to the Landis–Koch guidelines.¹⁷

Comparison of the MFHP tool versus expert GC risk estimates

Twenty-seven of the 150 (18%) unmodified pedigrees were assigned “elevated” risk status by the MFHP tool. The MFHP tool had sensitivity and specificity of 81% (95% confidence interval (CI): 54–96%) and 90% (95% CI: 83–94%) for pedigrees judged to be “elevated” by GC review. The PPV and NPV were 48% (95% CI: 29–68%) and 98% (95% CI: 93–99%), respectively (**Figure 1**). All 20 of the modified pedigrees intended to be at “strong” or “moderate” risk were identified as being at “elevated” risk by the MFHP tool. The κ statistic comparing the MFHP tool and GC review suggested moderate agreement (0.54).

Comparison of the FHT versus GC risk estimates

Fourteen of 150 (9%) of the unmodified pedigrees were considered to be at “elevated” risk by the FHT; of these 14 pedigrees, 3 of 150 (6%) were classified as being at “strong” risk and 11 of 150 (7%) were rated to be at “moderate” risk. The FHT had sensitivity and specificity of 75% (95% CI: 48–93%) and 99% (95% CI: 95–100%), respectively, as compared with pedigrees judged to be “elevated” by GC review. The PPV and NPV for FHT were 86% (95% CI: 57–98%) and 97% (95% CI: 93–99%),



MFHP versus GC	GC +	GC -	Total
MFHP +	13	14	27
MFHP -	3	120	123
Total	16 (TP)	134 (TN)	150
	<u>Value (95% CI)</u>		
Sensitivity	81% (54 – 96%)		
Specificity	90% (83 – 94%)		
PPV	48% (29 – 68%)		
NPV	98% (93 – 99%)		
Kappa	0.54, moderate agreement		

Figure 1 Diagram showing the process used for developing interpretations of 150 unmodified pedigrees derived from the ClinSeq population (see text) and resulting calculated values for sensitivity, specificity, PPV, NPV, and calculated κ for the My Family Health Portrait colorectal risk algorithm for detection of elevated colorectal cancer risk using expert genetic counselor pedigree interpretation as the “gold standard” for risk interpretation. CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

respectively. The κ statistic comparing the FHT with GC review indicated substantial agreement (0.78). For the modified control pedigrees, there was agreement for 10 of 10 (100%) between the FHT and pedigrees intended to be at “strong” risk. Eight of 10 (80%) of the modified pedigrees intended to be at “moderate” risk were identified as such by the FHT; for 2 of 10 (20%) the FHT rated the risk as “strong” instead of “moderate.”

Assessment of disagreement between GC risk estimates and the MFHP

There were three unmodified pedigrees that were identified as being at “elevated” risk by GC review that were rated “not elevated” risk by the MFHP tool (“false negatives”) and 14 pedigrees for which the MFHP tool rated risk as “elevated” and GC review rated risk as “not elevated” (“false positives”). The sources of the discrepancies were diverse (Table 1). The “false-negative” pedigrees

Table 1 Potential sources of heritable risk for CRC present in pedigrees yielding “false-negative” and “false-positive” results with the MFHP in comparison with a gold standard of pedigree review by three expert GCs

False negative	False positive
GC risk higher than MFHP	MFHP risk higher than GC
2° GasCA; 2° OvCA	2° CRC; 1° PanCA
1° BrCa, age: 40–49 years; 2° PrCA	2° CRC, age: 40–49 years
Patient data entry error	2° with both CRC and BrCA
1° EndoCA, age: 30–39 years	2° CRC, age: 50–59 years
	2° CRC; 2° CRC, both age unknown
	2° CRC; 2° CRC, both age unknown
	1° ColP, age: 40–49 years
	1° GasCA
	2° With both CRC and BrCA
	2° CRC, age: 40–49 years
	1° AdenoCA; 2° PanCA; 2° PanCA; 2° EndoCA
	1° ColP; 2° ColP; 2° CRC
	Proband ColP; 1° ColP
	1° PrCA, 2° GasCA, 2° EndoCA

Each cell represents a single pedigree; individuals are separated by a semicolon. Unspecified age ranges for cancer are either unknown or were 60 years of age or older. 1°, First-degree relative; 2°, second-degree relative; AdenoCA, adenocarcinoma, site unspecified; BrCA, breast cancer; ColP, colon polyp; CRC, colorectal cancer; EndoCA, endometrial/uterine cancer; GasCA, gastric cancer; GC, cancer genetic counselor; MFHP, My Family Health Portrait CRC risk assessment tool; PanCA, pancreatic cancer; PrCA, prostate cancer.

shared the characteristic of a lack of CRC cases or colorectal polyps in the proband or relatives. For one of the pedigrees, the error source was a failure to enter endometrial cancer as a structured term due to misspelling. The “false-positive” pedigrees often appeared to arise from differing assignment of risk arising from personal or family histories of colon polyps or a history of CRC or CRC-associated cancers in second-degree relatives.

DISCUSSION

This study provides preliminary evidence supporting the analytic validity of an automated, consumer-oriented screening algorithm for heritable CRC risk assessment that is compatible with the US Surgeon General’s MFHP. To our knowledge, it is the first study to formally evaluate the predictive value of an FH risk algorithm designed specifically to work with the MFHP tool. It is important to note that the “gold standard” in this study is not development of disease but rather the risk prediction arrived at by a group of three expert cancer GCs. The PPV values for the MFHP and the FHT exceed the 10% threshold that is considered an acceptable value for PPV for a genomic screening test.¹⁸ Because only a fraction of individuals with elevated CRC will actually develop disease, the PPV regarding development of clinical disease is likely to be lower than the values observed in this study, and the NPV is likely to be higher, for both the MFHP and FHT tools.

The MFHP algorithm was not designed to be diagnostic of CRC syndromes; rather, it was designed to be a potential component of an educational tool that helps individuals and their

health-care providers stratify risk for heritable CRC in order to have an informed conversation regarding further evaluation. As a screening tool intended for use by the general public, it was designed to have a higher sensitivity at the cost of specificity, based on the rationale that missing elevated risk of colon cancer incurs a greater penalty than encouraging additional patients to discuss their risk of CRC with a health-care provider. This choice could be debated because available evidence suggests FH is not a strong predictor of absolute risk for CRC in the population.¹⁹

Evaluation of the pedigrees for which the risk ratings were discordant between GC review and the MFHP tool suggests that the MFHP tool assigns “elevated” risk to pedigrees with second-degree relatives having CRC or multiple Lynch syndrome–related cancers, or a personal history or FH of colon polyps. The GCs might not attribute much risk to histories of colon polyps lacking details of pathology. The FHT also does not fully incorporate personal and FH of colon polyps when assigning risk. Assuming such polyps are benign may underestimate risk, and at least one counselor desired additional information.²⁰ Two of three instances for which the MFHP tool did not assign elevated risk represent somewhat gray areas of risk assessment for CRC. The lack of clarity in guidelines for such gray areas is reflected in the inconsistency in GC risk assignment; for two of three of the “false-negative” pedigrees, at least one expert GC rated the risk as “weak” and another rated the risk as “strong.”

This study has a number of important limitations. Measured internal agreement between the counselors suggests that expert GCs do not always concur when independently assigning CRC risk to pedigrees. The discrepancies between risk estimates provided by GC review and the MFHP algorithm highlight that, currently, there is no universally accepted “gold standard” for the detection of CRC risk in FHs in routine clinical care. Similar to the 2005 study by Qureshi et al.¹⁶, this study uses review by expert GCs to arrive at a proxy gold standard for CRC risk. Additional limitations of this study include the small sample size, the nature of the ClinSeq population (which may not represent the full diversity of the US population), and that the study was only able to examine the analytic validity (and not the clinical utility) of the MFHP CRC algorithm. Future studies should seek to replicate these findings and define the clinical validity and clinical utility of the MFHP tool in additional patient populations.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

This project was supported by the Genetic Disease Research Branch and the Genomic Healthcare Branch, National Human Genome Research Institute, National Institutes of Health. The authors thank the following individuals for contributions made to developing the risk assessment algorithm and for their thoughtful comments on the manuscript: Andrew Freedman, Mitchell Gail, Donald Hadley, David Lanier, Colleen McBride, Gurvanet

Randhawa, Maren Scheuner, Rodolfo Valdez, and Daniel Watterdorf. Some of the data in this manuscript were presented in abstract form at the 2011 meeting of the American Society of Human Genetics in Montreal, Canada.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Yoon PW, Scheuner MT, Khoury MJ. Research priorities for evaluating family history in the prevention of common chronic diseases. *Am J Prev Med* 2003;24:128–135.
2. Rich EC, Burke W, Heaton CJ, et al. Reconsidering the family history in primary care. *J Gen Intern Med* 2004;19:273–280.
3. Qureshi N, Carroll JC, Wilson B, et al. The current state of cancer family history collection tools in primary care: a systematic review. *Genet Med* 2009;11:495–506.
4. Reid GT, Walter FM, Brisbane JM, Emery JD. Family history questionnaires designed for clinical use: a systematic review. *Public Health Genomics* 2009;12:73–83.
5. Orlando LA, Buchanan AH, Hahn SE, et al. Development and validation of a primary care-based family health history and decision support program (MeTree). *NC Med J* 2013;74:287–296.
6. Qureshi N, Wilson B, Santaguida P, et al. *Family History and Improving Health. Evidence Report/Technology Assessment*. Agency for Healthcare Research and Quality. 2009:1–135. <http://www.ahrq.gov/research/findings/evidence-based-reports/famhimp-evidence-report.pdf>.
7. Special report: exome sequencing for clinical diagnosis of patients with suspected genetic disorders. http://www.bcbs.com/blueresources/tec/vols/28/28_03.pdf. Accessed 19 December 2013.
8. Yoon PW, Scheuner MT, Jorgensen C, Khoury MJ. Developing Family Healthware, a family history screening tool to prevent common chronic diseases. *Prev Chronic Dis* 2009;6:A33.
9. Facio FM, Feero WG, Linn A, Oden N, Manickam K, Biesecker LG. Validation of My Family Health Portrait for six common heritable conditions. *Genet Med* 2010;12:370–375.
10. Feero WG, Bigley MB, Brinner KM; Family Health History Multi-Stakeholder Workgroup of the American Health Information Community. New standards and enhanced utility for family health history information in the electronic health record: an update from the American Health Information Community's Family Health History Multi-Stakeholder Workgroup. *J Am Med Inform Assoc* 2008;15:723–728.
11. Owens KM, Marvin ML, Gelehrter TD, Ruffin MT 4th, Uhlmann WR. Clinical use of the Surgeon General's “My Family Health Portrait” (MFHP) tool: opinions of future health care providers. *J Genet Couns* 2011;20:510–525.
12. Kanetzke EE, Lynch J, Prows CA, Siegel RM, Myers MF. Perceived utility of parent-generated family health history as a health promotion tool in pediatric practice. *Clin Pediatr (Phila)* 2011;50:720–728.
13. The US Surgeon General's My Family Health Portrait. <https://familyhistory.hhs.gov/FHH/html/index.html>. Accessed 2 July 2012.
14. Valdez R, Yoon PW, Qureshi N, Green RF, Khoury MJ. Family history in public health practice: a genomic tool for disease prevention and health promotion. *Annu Rev Publ Health* 2010;31:69–87 1 p following 87.
15. Biesecker LG, Mullikin JC, Facio FM, et al.; NISC Comparative Sequencing Program. The ClinSeq Project: piloting large-scale genome sequencing for research in genomic medicine. *Genome Res* 2009;19:1665–1674.
16. Qureshi N, Bethea J, Modell B, et al. Collecting genetic information in primary care: evaluating a new family history tool. *Fam Pract* 2005;22:663–669.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
18. Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. The predictive capacity of personal genome sequencing. *Sci Transl Med* 2012;4:133ra158.
19. Taylor DP, Stoddard GJ, Burt RW, et al. How well does family history predict who will get colorectal cancer? Implications for cancer screening and counseling. *Genet Med* 2011;13:385–391.
20. Madlensky L, Daftary D, Burnett T, et al. Accuracy of colorectal polyp self-reports: findings from the colon cancer family registry. *Cancer Epidemiol Biomarkers Prev* 2007;16:1898–1901.