JNCI Cancer Spectrum (2021) 5(2): pkab022

doi: 10.1093/jncics/pkab022
First published online 8 March 2021
Brief Communications

Assessment of a Polygenic Risk Score for Colorectal Cancer to Predict Risk of Lynch Syndrome Colorectal Cancer

Mark A. Jenkins, PhD,[1,2] Daniel D. Buchanan, PhD,[2,3,4] John Lai, PhD,[1,5] Enes Makalic, PhD,[1] Gillian S. Dite, PhD,[1] Aung K. Win, PhD,[12] Mark Clendenning, PhD,[2,3] Ingrid M. Winship, Fracp,[6,7] Richard B. Hayes, PhD,[8] Jeroen R. Huyghe, PhD,[9] Ulrike Peters, PhD,[9] Steven Gallinger, MD, PhD,[10] Loïc Le Marchand, MD, PhD,[11] Jane C. Figueiredo, PhD,[12] Rish K. Pai, MD, PhD,[13] Polly A. Newcomb, PhD,[14,15] James M. Church, MD, PhD,[16] Graham Casey, PhD,[17] John L. Hopper, PhD[1]

[1]Centre for Epidemiology and Biostatistics, School of Population and Global Health, The University of Melbourne, Victoria, Australia, [2]Centre for Cancer Research, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Victoria, Australia, [3]Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Victoria, Australia, [4]Genomic Medicine and Family Cancer Clinic, Royal Melbourne Hospital, Parkville, Victoria, Australia, [5]Australian Genome Research Facility, Saint Lucia, Queensland, Australia, [6]Genetic Medicine, Royal Melbourne Hospital, Parkville, Victoria, Australia, [7]Department of Medicine, The University of Melbourne, Parkville, Victoria, Australia, [8]Division of Epidemiology, New York University School of Medicine, New York, NY, USA, [9]Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, [10]Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada, [11]Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA, [12]Department of Medicine, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA, [13]Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Scottsdale, AZ, USA, [14] Department of Epidemiology, University of Washington, Seattle, WA, [15]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA, [16]Departments of Stem Cell and Regenerative Medicine and Colorectal Surgery, Sanford R Weiss MD Center for Hereditary Colorectal Neoplasia, Digestive Disease and Surgery Institute, Cleveland Clinic Lerner Research Institute, Cleveland, OH, USA and [17]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

*Correspondence to: Mark A. Jenkins, PhD, Centre for Epidemiology and Biostatistics, School of Population and Global Health, University of Melbourne, Victoria 3010, Australia (e-mail: m.jenkins@unimelb.edu.au).

Abstract

It was not known whether the polygenic risk scores (PRSs) that predict colorectal cancer could predict colorectal cancer for people with inherited pathogenic variants in DNA mismatch repair genes—people with Lynch syndrome. We tested a PRS comprising 107 established single-nucleotide polymorphisms associated with colorectal cancer in European populations for 826 European-descent carriers of pathogenic variants in DNA mismatch repair genes (293 *MLH1*, 314 *MSH2*, 126 *MSH6*, 71 *PMS2*, and 22 *EPCAM*) from the Colon Cancer Family Registry, of whom 504 had colorectal cancer. There was no evidence of an association between the PRS and colorectal cancer risk, irrespective of which DNA mismatch repair gene was mutated, or sex (all 2-sided $P > .05$). The hazard ratio per standard deviation of the PRS for colorectal cancer was 0.97 (95% confidence interval = 0.88 to 1.06; 2-sided $P = .51$). Whereas PRSs are predictive of colorectal cancer in the general population, they do not predict Lynch syndrome colorectal cancer.

Polygenic risk scores (PRSs) aggregate genetic risk variants to predict disease risk and are an emerging tool in precision medicine. For colorectal cancer, we and others have identified more than 100 single-nucleotide polymorphisms (SNPs) that, when combined as a PRS, predict colorectal cancer (1,2). A clinically important question is whether this PRS is associated with colorectal cancer risk for people who have inherited a pathogenic variant in a DNA mismatch repair (MMR) gene, that is, people with Lynch syndrome. These people have, on average, a high colorectal cancer risk, and there is evidence of unidentified genetic factors that modify their risk (3). If identified, genetic risk-modifying factors would provide an avenue for improved personalized prevention strategies for people with Lynch syndrome. A recent paper reported risks of Lynch syndrome colorectal cancer for an existing colorectal cancer PRS (4). However, the authors did not estimate these risks directly but instead simply assumed the PRS was associated with Lynch syndrome colorectal cancer, leaving the question unanswered.

We conducted an analysis of 826 people with Lynch syndrome, of whom 504 had colorectal cancer, to determine

Received: 3 August 2020; **Revised:** 15 December 2020; **Accepted:** 4 March 2021

Table 1. Age at which half of the carriers developed colorectal cancer by quintile of polygenic risk score (PRS)

Participant characteristic	No.	Age (standard error), y				
		Quintile 1 (n = 165)	Quintile 2 (n = 164)	Quintile 3 (n = 166)	Quintile 4 (n = 165)	Quintile 5 (n = 166)
PRS, range		0.219-0.594	0.595-0.804	0.808-1.038	1.045-1.380	1.381-3.833
All genes and all carriers	826	49 (1.1)	49 (1.0)	49 (1.6)	48 (1.7)	48 (1.4)
Gene with pathogenic variant						
MLH1	293	46 (2.2)	45 (1.5)	54 (3.5)	41 (1.3)	46 (1.9)
MSH2	314	48 (1.4)	48 (1.2)	46 (2.4)	49 (2.3)	46 (3.0)
MSH6	126	—[a]	56 (4.3)	53 (4.6)	54 (3.3)	72 (—)
PMS2	71	55 (9.7)	57 (11)	67(15)	61 (7.5)	60 (—)
EPCAM	22	—	52 (3.8)	—	48 (—)	52 (5.1)
Sex						
Male	387	49 (1.0)	50 (1.7)	48 (1.2)	46 (1.1)	47 (1.2)
Female	439	51 (3.0)	48 (1.1)	52 (2.9)	53 (2.7)	52 (2.4)

[a]— = Insufficient data.

whether an existing colorectal cancer PRS is associated with Lynch syndrome colorectal cancer. Participants were from the Colon Cancer Family Registry (5), which recruited participants between 1998 and 2013, from the United States, Canada, Australia, and New Zealand: population-based colorectal cancer cases from state and regional population cancer registries; attendees with strong family histories of colorectal cancer at family cancer clinics; and relatives of these cases and attendees. Participants provided a blood sample, access to any colorectal tumors, ethnicity, and cancer and polyp history. For those whose colorectal cancer tumors were not accessed, attempts were made to verify colorectal cancer reports with cancer registrations, medical records, and relative reports. Participants were followed up every 5 years to update polyp and cancer history. Written informed consent was obtained from each participant, and research was approved by local institutional review boards.

Colorectal cancer case probands from the population-based families and colorectal cancer cases attending family cancer clinics were tested for germline variants in MMR genes. Relatives of identified carriers of pathogenic MMR germline variants were tested for their family-specific variant. Variants in *MLH1*, *MSH2*, *MSH6*, and *EPCAM* were identified by Sanger sequencing or denaturing high-performance liquid chromatography followed by confirmatory DNA sequencing (6). Variants in *PMS2* were identified using a modified protocol (7). Variants were classified for pathogenicity based on a 5-class system applied to variants cataloged within the InSiGHT database (8), with classes 4 and 5 considered pathogenic (9). Variants not yet classified by InSiGHT were considered pathogenic if predicted to result in a stop codon, frameshift, or large deletion, or if it removed a canonical splice site.

This analysis includes the 826 carriers identified as carrying a pathogenic variant in a DNA mismatch repair gene (293 *MLH1*, 314 *MSH2*, 126 *MSH6*, 71 *PMS2*, 22 *EPCAM*) of European descent and had undergone genome-wide SNP testing. SNP data for 462 carriers were from a previous testing, and genotyping, imputation, and quality control have been described (1). SNP data from the other 364 carriers were from a testing using the Infinium OncoArray-500K platform (Illumina, San Diego, USA) (10) and imputed using MiniMac v1.2.4 through the Michigan Imputation Server (11) using the European HRC r1.1 2016 reference. Filtering of the harmonized datasets for European participants was based on the first 2 principal components using the 1000 Genomes Project dataset as a reference (12). Genotypes for each of the 108 SNPs previously identified as being associated with colorectal cancer in European populations (1,2) were extracted from the harmonized set using PLINK v1.9 (13). One SNP was not included in the imputation reference panel (rs6928864), leaving 107 SNPs for this analysis (Supplementary Table 1, available online).

We analyzed data as a retrospective cohort of carriers (given the pathogenic variant is present from birth) censored at age of first polypectomy, giving 37 332 years of observation. There were 141 participants with a polypectomy and no colorectal cancer, 504 participants diagnosed with colorectal cancer with no previous polypectomy, and 75 participants with colorectal cancer diagnosed after polypectomy (included in the sensitivity analysis only; see the Supplementary Methods and Supplementary Table 2, available online).

We calculated 2 PRSs: 1) weighted sum of the number of risk alleles of each participant, using the variant's per-allele odds ratio as weights (1,2), and 2) count of the total number of risk alleles. We then tested for PRS associations with colorectal cancer risk by studying time to colorectal cancer (years of age since birth) by survival analysis and Cox regression. We allowed observations to be independent across, but nonindependent within, families by using the cluster option (14) in Stata (15) to produce robust standard errors. The PRS associations were assessed as per quintile and per standard deviation. Median observation time was 44 years (interquartile range = 36-53). All P values were 2-sided, calculated by Cox regression, and significant if less than .05.

We found no difference in the age at which half of the carriers were diagnosed with colorectal cancer by quintile of PRS (see Table 1) and no association of the PRS with colorectal cancer risk (Table 2), irrespective of the MMR gene mutated, sex, or method used to calculate the PRS (all 2-sided $P \geq .05$) (see Table 2).

Lynch syndrome colorectal cancer has genotypic features involving high mutability, consistent with a different genetic etiology than non-Lynch syndrome colorectal cancer. If this genetic etiology includes polygenic factors [and there is indirect evidence that these are substantial (3)], they will not necessarily be the SNPs identified to date for colorectal cancer risk, given the vast majority of colorectal cancer is not Lynch syndrome. We did not attempt to use this study to identify SNPs associated with Lynch syndrome colorectal cancer because the number of subjects was too small for a conventional genome-wide

Table 2. Association between the polygenic risk score (PRS) and colorectal cancer risk

Participant characteristic	No. of carriers	PRS using the per-allele odds ratio		PRS using the risk allele count	
		HR per SD (95% CI)	P[a]	HR per SD (95% CI)	P[a]
All genes and all carriers	826	0.97 (0.88 to 1.06)	.51	0.99 (0.90 to 1.10)	.90
Gene with pathogenic variant					
MLH1	293	0.98 (0.86 to 1.12)	.79	0.97 (0.83 to 1.14)	.72
MSH2	314	1.02 (0.86 to 1.22)	.78	1.02 (0.88 to 1.17)	.83
MSH6	126	0.94 (0.76 to 1.16)	.55	1.02 (0.80 to 1.30)	.90
PMS2	71	0.90 (0.63 to 1.28)	.56	0.99 (0.76 to 1.31)	.97
EPCAM	22	1.40 (0.92 to 2.14)	.12	1.95 (0.94 to 4.04)	.07
Sex					
Male	387	1.01 (0.89 to 1.15)	.87	1.01 (0.90 to 1.14)	.81
Female	439	0.94 (0.83 to 1.07)	.37	0.98 (0.86 to 1.13)	.81

[a]Two-sided Cox regression, 2-sided test. CI = confidence interval; HR = hazard ratio.

associations study analysis. We are progressing a larger, collaborative Lynch syndrome–specific genome-wide association study to address this important question.

Although there is evidence that the PRS for female breast cancer might be a modifier of risk for women with pathogenic variants in *BRCA1* and *BRCA2* (16), we found no evidence that the PRS for colorectal cancer is a modifier of colorectal cancer risk because of pathogenic variants in the DNA MMR genes. Application of the PRS for colorectal cancer to Lynch syndrome is thus unwarranted.

Funding

The Colon Cancer Family Registry (CCFR, www.coloncfr.org) is supported in part by funding from the National Cancer Institute at the National Institutes of Health (U01 CA167551). Support for case ascertainment was provided in part from the Surveillance, Epidemiology, and End Results (SEER) Program and the following US state cancer registries: AZ, CO, MN, NC, NH; and by the Victoria Cancer Registry (Australia) and Ontario Cancer Registry (Canada). SNP genotyping was supported by funding from the Canadian Cancer Society, the Ontario Ministry of Research and Innovation, and funding from the National Cancer Institute at the National Institutes of Health (U01 CA122839 and R01 CA143247), and by the National Health and Medical Research Council, Australia.

Notes

Role of the funder: The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures: The authors declare no conflicts of interest.

Author contributions: All authors of this research article have directly participated in the planning, execution, and/or analysis of this study. Mark Jenkins: project administration, recruitment, funding acquisition, formal analysis, design, writing—original draft, writing—review and editing. Daniel Buchanan: genetic testing, curation of molecular data, writing—review and editing. John Lai: imputation and quality control of genomic data, writing—reviewing and editing. Enes Makalic: analytical methods, writing—review and editing. Gillian Dite: analytical methods, writing—review and editing. Aung Win: analytical methods, writing—review and editing. Mark Clendenning: genetic testing, curation of molecular data, writing—review and editing. Ingrid Winship: clinical interpretation, recruitment, writing—review and editing. Richard Hayes: analytical methods, writing—review and editing. Jeroen Huyghe: analytical methods, writing—review and editing. Ulrike Peters: analytical methods, writing—review and editing. Steven Gallinger: project administration, recruitment, funding acquisition, writing—review and editing. Loïc Le Marchand: project administration, recruitment, funding acquisition, writing—review and editing. Jane Figueiredo: project administration, recruitment, funding acquisition, writing—review and editing. Rish Pai: project administration, writing—original draft, writing—review and editing. Polly Newcomb: project administration, recruitment, funding acquisition, writing—review and editing. James Church: project administration, recruitment, funding acquisition, writing—review and editing. Graham Casey: project administration, recruitment, funding acquisition, writing—review and editing. John Hopper: project administration, recruitment, funding acquisition, writing—review and editing.

Acknowledgements: We thank the men and women who participated in this research.

Data Availability

The data from this study cannot be shared publicly due to ethical and privacy reasons. The data are available on reasonable request to the Colon Cancer Family Registry www.coloncfr.org.

References

1. Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet*. 2019;51(1):76–87.
2. Law PJ, Timofeeva M, Fernandez-Rozadilla C, et al. PRACTICAL consortium. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat Commun*. 2019;10(1):2154.
3. Dowty JG, Win AK, Buchanan DD, et al. Cancer risks for MLH1 and MSH2 mutation carriers. *Hum Mutat*. 2013;34(3):490–497.
4. Fahed AC, Wang M, Homburger JR, et al. Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun*. 2020;11(1):3635.
5. Jenkins MA, Win AK, Templeton AS, et al. Colon Cancer Family Registry Cohort Investigators. Cohort profile: the Colon Cancer Family Registry Cohort (CCFRC). *Int J Epidemiol*. 2018;47(2):387–388.
6. Win AK, Jenkins MA, Dowty JG, et al. Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer Epidemiol Biomarkers Prev*. 2017;26(3):404–412.

7. Senter L, Clendenning M, Sotamaa K, et al. The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations. *Gastroenterology*. 2008;135(2):419–428.
8. InSIGHt. InSiGHT variant databases. http://insight-group.org/variants/databases/. Accessed January 15, 2020.
9. Thompson BA, Spurdle AB, Plazzer JP, et al.; on behalf of InSiGHT. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet*. 2014;46(2):107–115.
10. Amos CI, Dennis J, Wang Z, et al. The OncoArray Consortium: a network for understanding the genetic architecture of common cancers. *Cancer Epidemiol Biomarkers Prev*. 2017;26(1):126–135.
11. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284–1287.
12. , Auton A, Brooks LD, Durbin RM, et al. Genomes Project C. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74.
13. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–575.
14. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1. Berkeley, CA: University of California Press; 1967:221–233.
15. StataCorp. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC; 2019.
16. Kuchenbaecker KB, McGuffog L, Barrowdale D, et al. Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in BRCA1 and BRCA2 mutation carriers. *J Natl Cancer Inst*. 2017;109(7):djw302. doi: 10.1093/jnci/djw302 [published Online First: 2017/04/05]