CANCER GENETICS

Lynch Syndrome–Associated Variants and Cancer Rates in an Ancestrally Diverse Biobank

Rachel E. Rosenblum, MD[1]; Celina Ang, MD[2]; Sabrina A. Suckiel, MS[3,4]; Emily R. Soper, MS[3,4]; Meenakshi R. Sigireddi, MD, MPH[5]; Sinead Cullina, MS[3]; Gillian M. Belbin, PhD[3,6]; Aimee L. Lucas, MD, MS[7]; Eimear E. Kenny, PhD[3,5,6]; and Noura S. Abul-Husn, MD, PhD[3,4,5,6]

PURPOSE Limited data are available on the prevalence and clinical impact of Lynch syndrome (LS)–associated genomic variants in non-European ancestry populations. We identified and characterized individuals harboring LS-associated variants in the ancestrally diverse Bio*Me* Biobank in New York City.

PATIENTS AND METHODS Exome sequence data from 30,223 adult Bio*Me* participants were evaluated for pathogenic, likely pathogenic, and predicted loss-of-function variants in *MLH1*, *MSH2*, *MSH6*, and *PMS2*. Survey and electronic health record data from variant-positive individuals were reviewed for personal and family cancer histories.

RESULTS We identified 70 individuals (0.2%) harboring LS-associated variants in *MLH1* (n = 12; 17%), *MSH2* (n = 13; 19%), *MSH6* (n = 16; 23%), and *PMS2* (n = 29; 41%). The overall prevalence was 1 in 432, with higher prevalence among individuals of self-reported African ancestry (1 in 299) than among Hispanic/Latinx (1 in 654) or European (1 in 518) ancestries. Thirteen variant-positive individuals (19%) had a personal history, and 19 (27%) had a family history of an LS-related cancer. LS-related cancer rates were highest in individuals with *MSH6* variants (31%) and lowest in those with *PMS2* variants (7%). LS-associated variants were associated with increased risk of colorectal (odds ratio [OR], 5.0; $P = .02$) and endometrial (OR, 30.1; $P = 8.5 \times 10^{-9}$) cancers in Bio*Me*. Only 2 variant-positive individuals (3%) had a documented diagnosis of LS.

CONCLUSION We found a higher prevalence of LS-associated variants among individuals of African ancestry in New York City. Although cancer risk is significantly increased among variant-positive individuals, the majority do not harbor a clinical diagnosis of LS, suggesting underrecognition of this disease.

JCO Precis Oncol 4:1429-1444. © 2020 by American Society of Clinical Oncology

INTRODUCTION

People with genetic variants linked to hereditary cancers are at increased risk for adverse health outcomes and often escape clinical diagnosis.[1,2] Lynch syndrome (LS) is an autosomal dominant cancer syndrome caused by pathogenic variants in the mismatch repair (MMR) genes (*MLH1, MSH2, MSH6,* and *PMS2*) and deletions in the 3′ end of the *EPCAM* gene. Although mainly associated with colorectal cancer (CRC) and endometrial cancer (EC), LS also increases lifetime risk for various other cancers, including those of the ovaries, stomach, small intestine, upper urologic tract, brain, biliary system, pancreas, and cutaneous sebaceous glands.[3,4] Individuals with LS-associated variants are at increased risk of developing multiple malignancies during their lifetime and are more likely to develop cancer at an early age.

Early identification of individuals harboring LS-associated variants in MMR genes provides the opportunity toreduce cancer-related morbidity and mortality through specialized cancer screening, prophylactic surgical measures, and/or chemoprevention.[5-9] In current clinical practice, patients typically obtain genetic testing for LS when they meet clinical diagnostic criteria, including the Amsterdam II criteria or revised Bethesda guidelines.[10,11] These criteria rely on personal and family cancer history, age of cancer onset, and/or molecular tumor characteristics, with sensitivities as low as 25% and 50%, respectively.[12] As such, certain at-risk individuals may be missed. Recently, universal screening of all CRCs and ECs using microsatellite instability (MSI) and/or immunohistochemistry (IHC) for MMR protein expression has been recommended.[13-15] Although more effective than clinical diagnostic criteria, this method of testing still misses patients with LS who do not present with EC or CRC.[14]

The prevalence of LS has been estimated as 1 in 440, with the majority of cases resulting from pathogenic variants in *MLH1* and *MSH2*.[16,17] However, prevalence estimates are based on studies ascertaining predominantly patients with European ancestry (EA) who met clinical diagnostic criteria for LS or were diagnosed with CRC and/or EC.[18-21] Additionally, clinical testing

ASSOCIATED CONTENT

Appendix

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on October 16, 2020 and published at ascopubs.org/journal/po on November 23, 2020: DOI https://doi.org/10.1200/PO.20.00290

CONTEXT

Key Objective

We sought to evaluate the prevalence and clinical impact of Lynch syndrome (LS)–associated genomic variants in unselected and diverse patient populations.

Knowledge Generated

Expected pathogenic variants in mismatch repair (MMR) genes were identified in approximately 1 in 430 participants of an ancestrally diverse, electronic health record–linked biobank in New York City. The highest prevalence was seen in individuals of self-reported African ancestry (approximately 1 in 300). Variant-positive individuals had significantly increased risk of endometrial cancer and colorectal cancer but not of breast cancer. The vast majority of variant-positive individuals did not have a documented diagnosis of LS in their medical records, even when they had LS-related cancers, suggesting underrecognition of this syndrome.

Relevance

Genomic screening for LS-associated variants in MMR genes may be an effective tool to identify individuals with elevated cancer risk.

for *MSH6* and *PMS2* was developed after testing for *MLH1* and *MSH2*, such that many earlier studies of prevalence did not account for these genes.[22] More recent epidemiologic data suggest that LS may be more common than previously appreciated, with an estimated prevalence as high as 1 in 279.[23] Improved understanding of the prevalence and phenotypic spectrum of MMR gene variants is needed, especially within non-EA populations. This study evaluated the prevalence and clinical impact of expected pathogenic MMR variants in an ancestrally diverse, unselected population biobank.

PATIENTS AND METHODS

Study Population

The Bio*Me* Biobank is an electronic health record (EHR)–linked biobank with > 55,000 participants enrolled nonselectively from the Mount Sinai Health System in New York City. The study population consisted of 30,223 consented Bio*Me* participants age 18 years or older (upon enrollment) and with exome sequence data available through a collaboration with the Regeneron Genetics Center. Details of the study population have previously been described.[1] Self-reported ancestry of participants was derived from a multiple-choice survey administered on enrollment in Bio*Me*.[1,24] This study was approved by the Icahn School of Medicine at Mount Sinai's institutional review board.

Generation of Genomic Data and MMR Variant Annotation

Sample preparation and exome sequencing were performed at the Regeneron Genetics Center.[25] Post hoc filtering of samples and quality control of sequence data were performed as described previously, resulting in a total of 30,223 samples from participants age 18 years and older.[1] Sequence data for the MMR genes (*MLH1, MSH2, MSH6,* and *PMS2*) were extracted from exome sequences. Variants were annotated with the Variant Effect Predictor and cross-referenced with the ClinVar database (accessed March 2019).[26] Exome sequencing did not capture large genomic rearrangements in the MMR genes or in the 3′ area of *EPCAM*.

Pathogenic variants in MMR genes were identified based on a pathogenic or likely pathogenic assertion in ClinVar. In addition, we considered predicted loss-of-function (pLOF) variants not classified in ClinVar, including frameshift, stop-gain, start-loss, stop-loss, or canonical splice acceptor or donor. The union of ClinVar pathogenic/likely pathogenic and pLOF variants was termed expected pathogenic (n = 44) and was used to identify variant-positive individuals in Bio*Me* for subsequent analyses. Additional quality control included manual inspection of sequence reads for variant-positive individuals (n = 23) harboring 1 of 7 multiallelic sites annotated as expected pathogenic. Of these, 12 of 15 carriers of *MSH6* c.3261dupC and 1 of 2 carriers of *MSH6* c.3261del were determined to be false positives and excluded from downstream analyses.

Evaluation of Clinical Characteristics in Variant-Positive Individuals

Individuals harboring expected pathogenic variants in *MLH1, MSH2, MSH6*, or *PMS2* in Bio*Me*, termed variant positive (N = 70), were evaluated for personal and/or family histories of LS-related and other cancers. For this evaluation, we extracted International Classification of Diseases (ICD)-9 and ICD-10 codes related to these cancers from EHRs (Appendix Table A1). Screening or diagnostic testing for CRC and EC was evaluated by extraction of Current Procedural Terminology codes for colonoscopy (45378-45398) and transvaginal or pelvic ultrasound (76830, 76831, 76856, 76857). We carried out medical record review of variant-positive individuals to further evaluate personal and family cancer histories and to evaluate for clinical diagnosis, genetic testing, or tumor screening for LS. These data were supplemented by participant survey

data on personal and family cancer histories, which were available for 18 variant-positive individuals. Personal and family cancer histories were used to determine whether variant-positive individuals fulfilled Amsterdam II criteria[10] and/or revised Bethesda guidelines[11] (Appendix Table A2). Data were summarized using medians for continuous variables and frequencies and percentages for categoric variables.

LS-Related Cancer Case-Control Studies

We evaluated CRC, EC, and breast cancer (BC) risks in LS variant-positive versus variant-negative individuals in Bio*Me*. We excluded individuals within Bio*Me* who were second-degree relatives and closer from these analyses, as previously described.[1,27] This exclusion corrects for potential bias in the statistical analysis and is consistent with best practices for genetic association analyses.[28] For each cancer type, patient cases were defined as having any ICD-9/10 code for personal history of the cancer, and controls were defined as individuals without any of these codes. For CRC, we tested for association with variant-positive (n = 64) versus variant-negative (n = 26,942) unrelated individuals in Bio*Me*, using the saddlepoint approximation to account for unbalanced case-control ratios[29] (implemented via "SPAtest" package, R [version 3.5.3] and adjusting for age, sex, and the first 5 principal components of ancestry). EC and BC were assessed in variant-positive (n = 34) versus variant-negative (n = 15,669) unrelated women, adjusting for age and the first 5 principal components.

RESULTS

We evaluated variants in *MLH1, MSH2, MSH6,* and *PMS2* among 30,223 adult participants of the Bio*Me* Biobank with available exome sequencing data. Demographics of these participants have been described previously: 59% were women, the median age was 59 years, and 74% were non-EA by self-report.[1] We identified 44 expected pathogenic variants in MMR genes, including 7 *MLH1,* 5 *MSH2*, 14 *MSH6*, and 18 *PMS2* variants (Table 1). There were 70 Bio*Me* participants harboring at least one of these 44 variants. Though the majority (75%) of these variants were observed as singletons, 3 variants were observed in more than 5 individuals each, including *MSH2* c.1697del (n = 8), *MLH1* c.790+1G>A (n = 6), and *PMS2* c.137G>T (n = 6). None of these recurrent variants were observed predominantly in a single ancestry group, and none of the carriers of recurrent variants were related to one another.

Overall, 70 individuals (0.2%) in Bio*Me* harbored expected pathogenic variants in MMR genes (Table 2; Appendix Table A3). *PMS2* variants were most frequently identified (41%), followed by *MSH6* (23%), *MSH2* (19%), and *MLH1* (17%). Variant-positive individuals were 56% female, with a median age of 55 years. The overall estimated prevalence of Bio*Me* participants harboring expected pathogenic variants in MMR genes was 1 in 432. Prevalence was unchanged (1 in 428) in an unrelated subset of Bio*Me* that included only 1 individual from each first- or second-degree relationship (n = 27,816). Prevalence varied across self-reported ancestry groups and was highest in individuals of self-reported African American/African (AA) ancestry (1 in 299), with most harboring variants in *MSH6* (30%) or *PMS2* (39%). Prevalence was lower in EA (1 in 518) and Hispanic/Latinx (H/L; 1 in 654) ancestries.

We evaluated for LS-associated clinical characteristics in the 70 variant-positive individuals in Bio*Me*. The overall rate of malignancy was 34%, with the lowest rate in individuals harboring *PMS2* (21%) compared with *MSH2* (38%), *MLH1* (42%), or *MSH6* (50%) variants (Table 2). Of the 24 variant-positive individuals with cancer, 13 (54%) had a tumor type associated with LS (Appendix Table A3). LS-related cancers included cancers of the endometrium (n = 7; 18% of women), colorectum (n = 3; 4%), stomach or small bowel (n = 2; 3%), ovary or fallopian tube (n = 1; 3% of women), brain (n = 1, 1%), and cutaneous sebaceous gland (n = 1, 1%). Two individuals had 2 distinct LS-related cancers: an AA woman harboring *MSH6* c.892C>T with EC and CRC, and an H/L man harboring *MSH2* c.478C>T with small bowel and sebaceous gland cancers. Rates of LS-related cancers varied across ancestry groups, with higher rates in individuals of EA ancestry (27%) and lower rates in those of AA (13%) and H/L (13%) ancestries (Fig 1). Each of the 13 individuals with an LS-related cancer harbored a distinct MMR gene variant, with the exception of 2 individuals with *MLH1* c.790+1G>A (one with CRC and the other with gastric cancer) and 2 with *MSH2* c.1697del (one with fallopian tube cancer and the other with EC). Other cancer types observed in variant-positive individuals included BC (n = 5); skin cancers, including squamous and basal cell carcinomas and melanoma (n = 4); prostate cancer (n = 3); and lung cancer (n = 2). Two individuals had the same 2 variants, *PMS2* c.2444C>T (likely pathogenic) and c.2331dup (pLOF), for which the phase could not be determined. Neither of these individuals had EHR evidence of features consistent with constitutional MMR deficiency (which results from biallelic germline variants in MMR genes), including hematologic, brain, intestinal tract, or other cancers. Only 10 (31%) of 32 women without EC had undergone ultrasonography of the pelvis, and 15 (22%) of 67 individuals without CRC had completed a colonoscopy. Among 34 variant-positive individuals (49%) with a family history of cancer, the majority (n = 19; 56%) had first- and/or second-degree relatives diagnosed with an LS-related cancer (Appendix Table A3). Family members of individuals harboring *MLH1* variants had the highest rates of LS-related cancers (50%) or any cancer type (67%) compared with family members of individuals harboring other MMR gene variants (Table 2).

Review of medical records revealed that 2 variant-positive individuals (3%) had a diagnosis of LS. One was an H/L man harboring *MSH2* c.478C>T with duodenal adenocarcinoma

TABLE 1. Expected Pathogenic Variants in MMR Genes Identified Among 30,223 Bio*Me* Participants

CHR:POS:REF:ALT	Gene	Function	rsID	cDNA Position[a]	Protein Position	No. of Hets
3:36996686:C:T	*MLH1*	Stop gained	rs63751428	c.184C>T	p.Gln62Ter	1
3:37012098:C:T	*MLH1*	Stop gained	rs63751615	c.676C>T	p.Arg226Ter	1
3:37014545:G:A	*MLH1*	Splice donor	rs267607789	c.790+1G>A		6
3:37017508:C:T	*MLH1*	Missense	rs63751194	c.793C>T	p.Arg265Cys	1
3:37025895:G:T	*MLH1*	Stop gained	rs63750443	c.1297G>T	p.Glu433Ter	1
3:37048610:G:A	*MLH1*	Splice donor	rs267607879	c.1989+1G>A		1
3:37048994:G:T[b]	*MLH1*	Stop gained	rs147542208	c.2080G>T	p.Glu694Ter	1
2:47410205:C:T	*MSH2*	Stop gained	rs63751426	c.478C>T	p.Gln160Ter	1
2:47463029:A:G[b]	*MSH2*	Splice acceptor	rs1573547594	c.1387-2A>G		1
2:47470995:CA:C	*MSH2*	Frameshift	rs1553367635	c.1697delA	p.Asn566fs	8
2:47475171:G:C	*MSH2*	Missense	rs63750875	c.1906G>C	p.Ala636Pro	2
2:47480869:G:T[b]	*MSH2*	Stop gained	rs749543152	c.2632G>T	p.Glu878Ter	1
2:47783495:T:A	*MSH6*	Splice donor	rs1553408469	c.260+2T>A		1
2:47798875:C:T	*MSH6*	Stop gained	rs146816935	c.892C>T	p.Arg298Ter	1
2:47799083:A:G	*MSH6*	Missense	rs1553412495	c.1100A>G	p.His367Arg	1
2:47799331:GTA:G	*MSH6*	Frameshift	rs878853702	c.1350_1351delAT	p.Phe451fs	1
2:47799684:GTT:G	*MSH6*	Frameshift	rs587783056	c.1705_1706delTT	p.Phe569fs	1
2:47799823:TC:T	*MSH6*	Frameshift	—	c.1842delC	p.Cys615fs	1
2:47800056:C:CA	*MSH6*	Frameshift	—	c.2079dupA	p.Cys694fs	1
2:47803473:C:T	*MSH6*	Missense	rs63750617	c.3226C>T	p.Arg1076Cys	2
2:47803500:A:AC	*MSH6*	Frameshift	rs267608078	c.3261dupC	p.Phe1088fs	2
2:47803500:AC:A	*MSH6*	Frameshift	rs267608078	c.3261delC	p.Phe1088fs	1
2:47803552:CTT:C	*MSH6*	Frameshift	—	c.3311_3312delTT	p.Phe1104fs	1
2:47806286:A:ATTAT[b]	*MSH6*	Frameshift	—	c.3732_3735dup	p.Ser1246fs	1
2:47806605:AAAGC:A	*MSH6*	Frameshift	rs267608120	c.3959_3962delCAAG	p.Ala1320fs	1
2:47806630:A:ATCAG	*MSH6*	Frameshift	rs267608121	c.3984_3987dup	p.Leu1330fs	1
7:5977589:G:A	*PMS2*	Missense	rs587779338	c.2444C>T	p.Ser815Leu	4
7:5977629:G:A	*PMS2*	Stop gained	rs63751466	c.2404C>T	p.Arg802Ter	3
7:5977701:A:AG[b]	*PMS2*	Frameshift	—	c.2331dupG	p.Phe778fs	2[c]
7:5986838:G:A	*PMS2*	Stop gained	rs63751422	c.1927C>T	p.Gln643Ter	1
7:5986933:A:AT	*PMS2*	Frameshift	rs63750250	c.1831dupA	p.Ile611fs	1
7:5987174:C:A	*PMS2*	Stop gained	rs878854037	c.1591G>T	p.Glu531Ter	1
7:5989821:GCTGA:G[b]	*PMS2*	Frameshift	rs757679199	c.1119_1122delTCAG	p.Gln374fs	1
7:5989956:C:A	*PMS2*	Splice acceptor	rs587780064	c.989-1G>T		1
7:5992012:G:A	*PMS2*	Stop gained	rs143277125	c.949C>T	p.Gln317Ter	1
7:5992018:G:A	*PMS2*	Stop gained	rs200640585	c.943C>T	p.Arg315Ter	2
7:5995534:C:A	*PMS2*	Missense	rs267608153	c.903G>T	p.Lys301Asn	2
7:5995612:T:C	*PMS2*	Synonymous	rs876659736	c.825A>G	p.Gln275 =	1
7:5995628:G:C	*PMS2*	Stop gained	rs786201047	c.809C>G	p.Ser270Ter	1
7:6003717:T:TC	*PMS2*	Frameshift	rs587781716	c.325dupG	p.Glu109fs	1
7:6003793:C:A	*PMS2*	Splice acceptor	rs764171734	c.251-1G>T		1
7:6005918:C:A	*PMS2*	Missense	rs121434629	c.137G>T	p.Ser46Ile	6
7:6008996:C:A	*PMS2*	Splice donor	rs587782074	c.23+1G>T		1
7:6009019:T:C	*PMS2*	Start lost	rs587779333	c.1A>G	p.Met1Val	1

Abbreviations: Hets, heterozygotes; MMR, mismatch repair; pLOF, predicted loss of function.

[a]cDNA position provided for *MLH1* ENST00000231790 (NM_000249.3), *MSH2* ENST00000233146 (NM_000251.2), *MSH6* ENST00000234420 (NM_000179.2), and *PMS2* ENST00000265849 (NM_000535.7).

[b]pLOF variants not in ClinVar.

[c]These individuals also harbor *PMS2* 7:5977589:G:A (c.2444C>T).

TABLE 2. Demographics and Clinical Characteristics of 70 Bio*Me* Participants Harboring Expected Pathogenic Variants in *MLH1, MSH2, MSH6,* or *PMS2*

Characteristic	All MMR Genes (N = 70)	*MLH1* (n = 12)	*MSH2* (n = 13)	*MSH6* (n = 16)	*PMS2* (n = 29)
Median (range) age, years	55 (24-77)	56 (40-74)	55 (26-76)	56 (30-77)	56 (24-77)
Sex					
Male	31 (44)	5 (42)	4 (31)	5 (31)	17 (59)
Female	39 (56)	7 (58)	9 (69)	11 (69)	12 (41)
Self-reported ancestry, No. (prevalence)					
African American/African (n = 6,878)	23 (1:299)	4 (1:1,720)	3 (1:2,293)	7 (1:983)	9 (1:764)
Hispanic/Latinx (n = 10,460)	16 (1:654)	4 (1:2,615)	4 (1:2,615)	4 (1:2,615)	4 (1:2,615)
European (n = 7,772)	15 (1:518)	3 (1:2,591)	2 (1:3,886)	3 (1:2,591)	7 (1:1,110)
South Asian (n = 605)	2 (1:303)	1 (1:605)	0	1 (1:605)	0
East/Southeast Asian (n = 757)	1 (1:757)	0	0	0	1 (1:757)
Native American (n = 52)	0	0	0	0	0
Multiple selected (n = 1,125)	8	0	4	1	3
Other (n = 2,343)	5	0	0	0	5
Not available (n = 231)	0	0	0	0	0
Total (N = 30,223)	70 (1:432)	12 (1:2,519)	13 (1:2,325)	16 (1:1,889)	29 (1:1,042)
LS diagnosis					
EHR diagnosis	2 (3)	1 (8)	1 (8)	0 (0)	0 (0)
Amsterdam II criteria	1 (1)	1 (8)	0 (0)	0 (0)	0 (0)
Personal cancer history					
Colorectal cancer	3 (4)	1 (8)	0 (0)	2 (13)	0 (0)
Endometrial cancer[a]	7 (18)	1 (14)	1 (11)	3 (27)	2 (17)
Any LS-related cancer[b]	13 (19)	3 (25)	3 (23)	5 (31)	2 (7)
Any cancer	24 (34)	5 (42)	5 (38)	8 (50)	6 (21)
Family cancer history					
Colorectal cancer	12 (17)	4 (33)	3 (23)	2 (13)	3 (10)
Endometrial cancer	8 (11)	3 (25)	1 (8)	1 (6)	3 (10)
Any LS-related cancer	19 (27)	6 (50)	4 (31)	2 (13)	7 (24)
Any cancer	34 (49)	8 (67)	8 (62)	5 (31)	13 (45)

NOTE. All data are No. (%) unless otherwise noted in row headings.

Abbreviations: EHR, electronic health record; LS, Lynch syndrome; MMR, mismatch repair.

[a]Only women included.

[b]LS-related cancers include tumors of the colorectum, endometrium, stomach, small bowel, ovaries, pancreas, ureter, renal pelvis, biliary tract, brain, and sebaceous gland as well as keratoacanthomas.

and sebaceous gland cancer, consistent with the Muir-Torre variant of LS. IHC for the MMR proteins of both tumors revealed a pattern of absent or reduced MSH2 and absent MSH6. This was the only individual with evidence of IHC or MSI screening in tumor tissue. The second individual with a diagnosis of LS was an EA woman harboring *MLH1* c.184C>T with EC and malignant pleural mesothelioma. This was the only individual who fulfilled Amsterdam II criteria. Of the 13 variant-positive individuals with LS-related cancers, 2 fulfilled the revised Bethesda guidelines for tumor testing of MSI: an AA woman harboring *MSH6* c.892C>T with CRC, EC, and BC; and a South Asian man harboring *MLH1* c.790+1G>A with CRC. Of the 7 ECs and 3 CRCs in our study population, tumor histology was only available for an EA woman harboring *PMS2* c.137G>T with endometrioid EC. IHC/MSI screening was not performed.

We tested for association with the most frequently observed cancers—CRC, EC, and BC—in unrelated LS variant-positive versus variant-negative individuals in Bio*Me* (Fig 2). Variant-positive individuals had significantly increased odds of CRC (odds ratio [OR], 5.0; 95% CI, 1.3 to 18.8; $P = .02$) compared with variant-negative individuals.

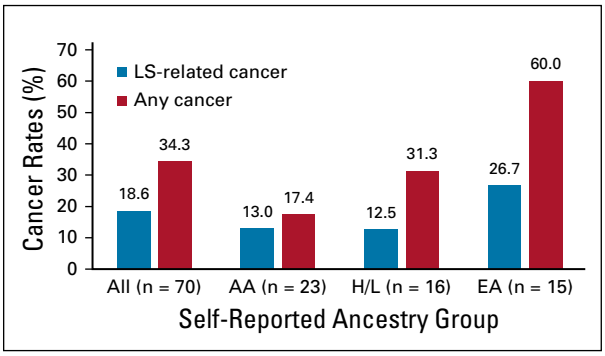

FIG 1. Rates of Lynch syndrome (LS)–related and overall cancers among self-reported ancestry groups in the Bio*Me* Biobank. Rates of LS-related cancers and any cancer type were evaluated in all 70 Bio*Me* participants harboring LS-associated variants, and across the 3 largest self-reported ancestry groups: African American/African (AA; n = 23), Hispanic/Latinx (H/L; n = 16), and European (EA; n = 15). South Asian (n = 2), East/Southeast Asian (n = 1), multiple selected (n = 8), and other (n = 5) ancestries are not shown.

Variant-positive women had significantly increased odds of EC (OR, 30.8; 95% CI, 9.6 to 99.0; $P = 8.5 \times 10^{-9}$) but not of BC (OR, 2.5; 95% CI, 0.9 to 6.9; $P = .08$) compared with variant-negative women.

DISCUSSION

We identified and characterized individuals harboring LS-associated variants in MMR genes within a large, ancestrally diverse biobank in New York City. The overall prevalence of expected pathogenic MMR variants was ~1 in 430, similar to previous LS prevalence estimates[16,23]; however, prevalence varied across ancestry groups. Consistent with our previous work identifying genomic risk for *BRCA1/2*-related cancers,[1] we found that the vast majority of variant-positive individuals did not have an awareness of their increased cancer risk.

Expected pathogenic variants in *PMS2* and *MSH6* were most common, with frequencies of ~1 in 1,000 and 1 in 1,900. Although up to 90% of LS has been attributed to germline mutations in *MLH1* or *MSH2*, more recent studies have shown *PMS2* and *MSH6* variants to be more frequent.[16,23,30] Prior studies ascertaining patients with CRC and/or meeting clinical diagnostic criteria for LS may have underestimated the prevalence of *MSH6* and *PMS2* variants, which are thought to be less penetrant than *MLH1* and *MSH2* variants.[22,31] In this study, individuals harboring an expected pathogenic variant in *PMS2* had the lowest rates of cancer compared with other MMR genes, suggesting lower penetrance. However, individuals harboring expected pathogenic variants in *MSH6* had the highest rates of cancer among variant-positive individuals.

We estimated the prevalence of LS-associated variants in diverse populations, including self-reported AA and H/L ancestry groups, for which estimates did not previously exist. The prevalence in individuals of H/L ancestry was ~1 in 650, lower than that seen in AA and EA ancestries. The highest prevalence (~1 in 300) was in individuals of AA ancestry. Few studies have described LS in AA populations to date, despite AA patients having the highest CRC incidence and mortality of all patient populations.[32] Guindalini et al[33] previously evaluated 51 AA families with pathogenic variants or variants of uncertain significance in an MMR gene and showed that 61% of families had a deleterious variant in *MLH1*. This is distinct from our findings, in which the majority of AA individuals harbored expected pathogenic variants in *MSH6* (30%) or *PMS2* (39%).

Expected pathogenic variants in MMR genes were associated with significantly increased risk of EC and CRC in Bio*Me*. EC and CRC were the most frequently observed cancers, affecting 18% of variant-positive women and 4% of variant-positive individuals, respectively. Notably, 4 of 7 women with EC harbored *MSH6* variants. Several previous studies have suggested a higher risk for EC in women with pathogenic variants in *MSH6* compared with other MMR genes, with an estimated 26% cumulative risk of EC by age 70 years.[34-36] Our data did not support an overall association between MMR pathogenic variants and BC risk, which is the subject of ongoing debate.[37,38] Some studies have demonstrated increased BC rates in women with MMR pathogenic variants, most recently in *MSH6* and *PMS2*,[22,39] while other studies have not.[40-42] The highest rate of BC in our study was in women with *MSH6* variants (3 of 11). None of the 12 women with a *PMS2* variant had evidence of BC.

Formal diagnosis of LS was only documented in EHRs for 2 variant-positive individuals. Lack of awareness of increased cancer risk could be due in part to the diversity of this cohort, as racial and ethnic minorities access genetic counseling and testing considerably less than non-Hispanic whites.[43,44] Previous studies have shown that the Amsterdam II criteria identify approximately 13%-67% of families with LS, and the revised Bethesda guidelines detect approximately 78%-91% of patients with LS, both of which are markedly higher than 1% and 3%, respectively, seen in this study.[45] Overall, we found that 96% of variant-positive individuals did not meet any current clinical diagnostic criteria for LS testing. This is consistent with the growing concerns that medical history–based screening strategies fail to identify the majority of individuals with high genomic risk for LS and other hereditary cancers.[46] Universal tumor screening of all newly diagnosed CRCs and ECs is now recommended to increase the identification of LS.[13-15] However, this approach did not appear to enable LS diagnosis in this study: no variant-positive individuals with CRC or EC had EHR documentation of IHC or MSI screening, likely because their cancers were diagnosed before the widespread adoption of tumor screening and/or outside of the health system.

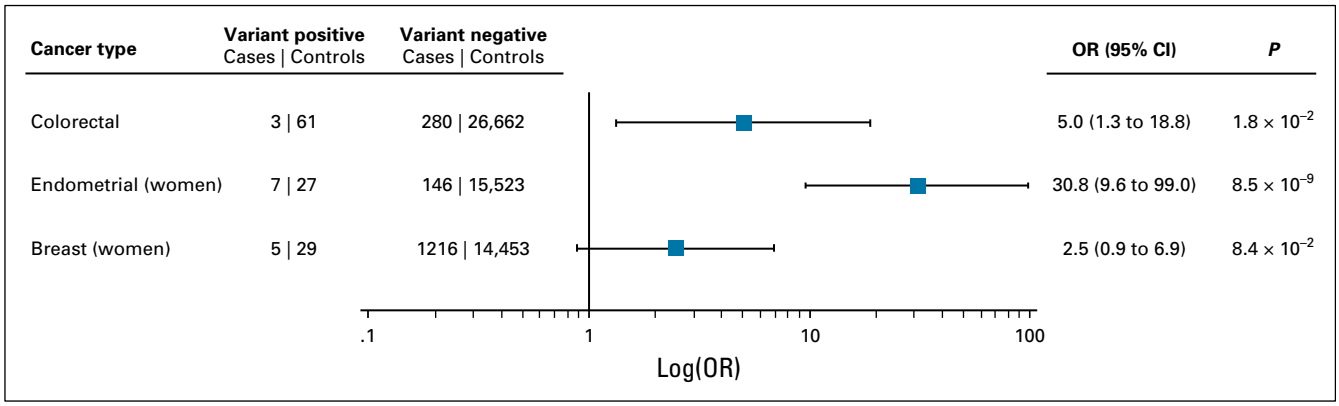

FIG 2. Lynch syndrome (LS)–associated variants are associated with increased risk of colorectal cancer (CRC) and endometrial cancer. CRC, endometrial cancer, and breast cancer cases and controls were defined using International Classification of Diseases (ICD)-9 and ICD-10 codes extracted from electronic health records. We excluded second-degree relatives and closer from these analyses. The odds ratio (OR) for CRC was calculated by saddlepoint approximation, adjusting for age, sex, and the first 5 principal components of ancestry. ORs for endometrial and breast cancers were similarly calculated in women only, adjusting for age and the first 5 principal components.

This study had some limitations. First, personal and family cancer histories were obtained retrospectively from EHR data, which may be incomplete and/or inaccurate. This might downwardly bias the estimation of penetrance of LS-associated variants. Additionally, we used exome sequencing to identify expected pathogenic variants in MMR genes; however, this does not capture large genomic rearrangements, such as multiexon deletions or duplications in MMR genes or in *EPCAM*, which can account for a large portion of LS.[30,47] Detection of pathogenic *PMS2* variants may have been confounded by numerous homologous pseudogenes, particularly *PMS2CL*, potentially resulting in a portion of *PMS2* variant-positive individuals having false positive results.[48] Of note, there were 13 individuals harboring expected pathogenic variants at multiallelic sites who were excluded from this study after manual review of the variants. However, we noted that 1 of these individuals had a personal history of LS-related cancer, and 2 others had first-degree relatives with LS-related cancers. Therefore, the quality control stringency applied may have excluded true variant-positive individuals from the study. Validation by Sanger sequencing or other methods would be needed to confirm the presence of lower-quality variants identified by research exome sequencing. Finally, this study did not include variants of uncertain significance, which can be observed more frequently in non-EA populations[1]; some of these may eventually be reclassified as pathogenic as knowledge about LS-associated variants continues to grow.

We found that the vast majority of variant-positive individuals did not have EHR evidence of a LS diagnosis, even when they had LS-related cancers, suggesting underrecognition of this syndrome. Genomic screening for LS-associated genomic variants in unselected and diverse populations can be used to identify individuals at elevated risk of CRC, EC, and other cancers. Prospective studies using this genomics-first approach are needed to better understand the penetrance of MMR variants and to inform targeted clinical strategies for the prevention and earlier diagnosis of LS-related cancers.

AFFILIATIONS

[1]Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY
[2]Division of Hematology and Medical Oncology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY
[3]The Institute for Genomic Health, Icahn School of Medicine at Mount Sinai, New York, NY
[4]Division of Genomic Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY
[5]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY
[6]Division of General Internal Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY
[7]Henry D. Janowitz Division of Gastroenterology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY

CORRESPONDING AUTHOR

Noura S. Abul-Husn, MD, PhD, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Pl, Box 1003, New York, NY 10029; Twitter: @NouraAbulHusn; e-mail: noura.abul-husn@mssm.edu.

PRIOR PRESENTATION

Presented in part at the Annual Meeting of the American Society of Clinical Oncology, Chicago, IL, May 31-June 4, 2019.

SUPPORT

Supported by dedicated funding to the Institute for Genomic Health by the Icahn School of Medicine at Mount Sinai. S.A.S., G.M.B., E.E.K., and N.S.A-H. are supported by the National Institutes of Health, National Human Genome Research Institute (NHGRI) and by National Institute on Minority Health and Health Disparities Grant No. U01 HG009610.

E.E.K. and N.S.A.-H. are supported by NHGRI Grant No. U01 HG011176. E.E.K. is supported by NHGRI Grants No. R01 HG010297, U01 HG009080, UM1 HG0089001, and U01 HG007417; National Heart, Lung, Blood Institute Grants No. R01 HL104608, X01 HL1345; and National Institute of Diabetes and Kidney and Digestive Disease Grant No. R01 DK110113. A.L.L. is supported by American Cancer Society Grant No. 129387-MRSG-16-015-01-CPHPS. Exome sequencing of Bio*Me* participants was performed in collaboration with the Regeneron Genetics Center.

AUTHOR CONTRIBUTIONS

Conception and design: Celina Ang, Aimee L. Lucas, Eimear E. Kenny, Noura S. Abul-Husn
Collection and assembly of data: Rachel E. Rosenblum, Meenakshi R. Siggireddi, Sinead Cullina, Gillian M. Belbin, Eimear E. Kenny, Noura S. Abul-Husn
Data analysis and interpretation: Rachel E. Rosenblum, Celina Ang, Sabrina A. Suckiel, Emily R. Soper, Meenakshi R. Sigireddi, Gillian M. Belbin, Aimee L. Lucas, Eimear E. Kenny, Noura S. Abul-Husn
Financial support: Eimear E. Kenny, Noura S. Abul-Husn
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Aimee L. Lucas
Research Funding: Immunovia

Eimear E. Kenny
Honoraria: Illumina, Regeneron

Noura S. Abul-Husn
Honoraria: Genentech
Employment: Regeneron

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

We thank participants of the Bio*Me* Biobank for their permission to use their health and genomic information.

REFERENCES

1. Abul-Husn NS, Soper ER, Odgis JA, et al: Exome sequencing reveals a high prevalence of *BRCA1* and *BRCA2* founder variants in a diverse population-based biobank. Genome Med 12:2, 2019
2. Manickam K, Buchanan AH, Schwartz MLB, et al: Exome sequencing-based screening for *BRCA1/2* expected pathogenic variants among adult biobank participants. JAMA Netw Open 1:e182140, 2018
3. Watson P, Lynch HT: The tumor spectrum in HNPCC. Anticancer Res 14:1635-1639, 1994
4. Watson P, Vasen HFA, Mecklin JP, et al: The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. Int J Cancer 123:444-449, 2008
5. Järvinen HJ, Aarnio M, Mustonen H, et al: Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology 118:829-834, 2000
6. de Jong AE, Hendriks YM, Kleibeuker JH, et al: Decrease in mortality in Lynch syndrome families because of surveillance. Gastroenterology 130:665-671, 2006
7. Lindor NM, Petersen GM, Hadley DW, et al: Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: A systematic review. JAMA 296:1507-1517, 2006
8. Schmeler KM, Lynch HT, Chen LM, et al: Prophylactic surgery to reduce the risk of gynecologic cancers in the Lynch syndrome. N Engl J Med 354:261-269, 2006
9. Burn J, Sheth H, Elliott F, et al: Cancer prevention with aspirin in hereditary colorectal cancer (Lynch syndrome), 10-year follow-up and registry-based 20-year data in the CAPP2 study: A double-blind, randomised, placebo-controlled trial. Lancet 395:1855-1863, 2020
10. Vasen HF, Watson P, Mecklin JP, et al: New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 116:1453-1456, 1999
11. Umar A, Boland CR, Terdiman JP, et al: Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 96:261-268, 2004
12. Tranø G, Sjursen W, Wasmuth HH, et al: Performance of clinical guidelines compared with molecular tumour screening methods in identifying possible Lynch syndrome among colorectal cancer patients: A Norwegian population-based study. Br J Cancer 102:482-488, 2010
13. Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group: Recommendations from the EGAPP Working Group: Genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. Genet Med 11:35-41, 2009
14. Palomaki GE, McClain MR, Melillo S, et al: EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. Genet Med 11:42-65, 2009
15. Committee on Practice Bulletins: ACOG Practice Bulletin No. 147: Lynch syndrome. Obstet Gynecol 124:1042-1054, 2014
16. Chen S, Wang W, Lee S, et al: Prediction of germline mutations and cancer risk in the Lynch syndrome. JAMA 296:1479-1487, 2006
17. Peltomäki P, Vasen H: Mutations associated with HNPCC predisposition: Update of ICG-HNPCC/INSiGHT mutation database. Dis Markers 20:269-276, 2004
18. Yurgelun MB, Kulke MH, Fuchs CS, et al: Cancer susceptibility gene mutations in individuals with colorectal cancer. J Clin Oncol 35:1086-1095, 2017
19. Hampel H, Frankel WL, Martin E, et al: Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). N Engl J Med 352:1851-1860, 2005
20. Moreira L, Balaguer F, Lindor N, et al: Identification of Lynch syndrome among patients with colorectal cancer. JAMA 308:1555-1565, 2012
21. Hampel H, Frankel W, Panescu J, et al: Screening for Lynch syndrome (hereditary nonpolyposis colorectal cancer) among endometrial cancer patients. Cancer Res 66:7810-7817, 2006
22. Espenschied CR, LaDuca H, Li S, et al: Multigene panel testing provides a new perspective on Lynch syndrome. J Clin Oncol 35:2568-2575, 2017

23. Win AK, Jenkins MA, Dowty JG, et al: Prevalence and penetrance of major genes and polygenes for colorectal cancer. Cancer Epidemiol Biomarkers Prev 26:404-412, 2017
24. Belbin GM Wenric S, Cullina S, et al: Towards a fine-scale population health monitoring System. bioRxiv:780668, 2019
25. Dewey FE, Gusarova V, O'Dushlaine C, et al: Inactivating variants in *ANGPTL4* and risk of coronary artery disease. N Engl J Med 374:1123-1133, 2016
26. Landrum MJ, Lee JM, Riley GR, et al: ClinVar: Public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res 42(D1): D980-D985, 2014
27. Manichaikul A, Mychaleckyj JC, Rich SS, et al: Robust relationship inference in genome-wide association studies. Bioinformatics 26:2867-2873, 2010
28. Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661-678, 2007
29. Dey R, Schmidt EM, Abecasis GR, et al: A Fast and accurate algorithm to test for binary phenotypes and its application to PheWAS. Am J Hum Genet 101:37-49, 2017
30. DA Silva FC, Wernhoff P, Dominguez-Barrera C, et al: Update on hereditary colorectal cancer. Anticancer Res 36:4399-4405, 2016
31. Yamano T, Hamanaka M, Babaya A, et al: Management strategies in Lynch syndrome and familial adenomatous polyposis: A national healthcare survey in Japan. Cancer Sci 108:243-249, 2017
32. Howlader N, Noone AM, Krapcho M, et al (eds): SEER Cancer Statistics Review, 1975-2012. Bethesda, MD, National Cancer Institute, 2014
33. Guindalini RS, Win AK, Gulden C, et al: Mutation spectrum and risk of colorectal cancer in African American families with Lynch syndrome. Gastroenterology 149:1446-1453, 2015
34. Ramsoekh D, Wagner A, van Leerdam ME, et al: A high incidence of *MSH6* mutations in Amsterdam criteria II–negative families tested in a diagnostic setting. Gut 57:1539-1544, 2008
35. Baglietto L, Lindor NM, Dowty JG, et al: Risks of Lynch syndrome cancers for *MSH6* mutation carriers. J Natl Cancer Inst 102:193-201, 2010
36. Talseth-Palmer BA, McPhillips M, Groombridge C, et al: *MSH6* and *PMS2* mutation positive Australian Lynch syndrome families: Novel mutations, cancer risk and age of diagnosis of colorectal cancer. Hered Cancer Clin Pract 8:5, 2010
37. Castellsagué E, Foulkes WD: Lynch syndrome: Five unanswered questions. Clin Genet 87:503-506, 2015
38. Ten Broeke SW, Suerink M, Nielsen M: Response to Roberts et al. 2018: Is breast cancer truly caused by *MSH6* and *PMS2* variants or is it simply due to a high prevalence of these variants in the population? Genet Med 21:256-257, 2019
39. Roberts ME, Jackson SA, Susswein LR, et al: *MSH6* and *PMS2* germ-line pathogenic variants implicated in Lynch syndrome are associated with breast cancer. Genet Med 20:1167-1174, 2018
40. Couch FJ, Shimelis H, Hu C, et al: Associations between cancer predisposition testing panel genes and breast cancer. JAMA Oncol 3:1190-1196, 2017
41. Møller P, Seppälä TT, Bernstein I, et al: Cancer risk and survival in *path_MMR* carriers by gene and gender up to 75 years of age: A report from the Prospective Lynch Syndrome Database. Gut 67:1306-1316, 2018
42. Stoll J, Rosenthal E, Cummings S, et al: No evidence of increased risk of breast cancer in women with Lynch syndrome identified by multigene panel testing. Precision Oncol 4:51-60, 2020
43. Hall M, Olopade OI: Confronting genetic testing disparities: Knowledge is power. JAMA 293:1783-1785, 2005
44. Huo D, Olopade OI: Genetic testing in diverse populations: Are researchers doing enough to get out the correct message? JAMA 298:2910-2911, 2007
45. Giardiello FM, Allen JI, Axilbund JE, et al: Guidelines on genetic evaluation and management of Lynch syndrome: A consensus statement by the US Multi-Society Task Force on colorectal cancer. Gastroenterology 147:502-526, 2014
46. Murray MF, Evans JP, Khoury MJ: DNA-based population screening: Potential suitability and important knowledge gaps. JAMA 323:307-308, 2019
47. Smith MJ, Urquhart JE, Harkness EF, et al: The contribution of whole gene deletions and large rearrangements to the mutation spectrum in inherited tumor predisposing syndromes. Hum Mutat 37:250-256, 2016
48. De Vos M, Hayward BE, Picton S, et al: Novel *PMS2* pseudogenes can conceal recessive mutations causing a distinctive childhood cancer syndrome. Am J Hum Genet 74:954-964, 2004

TABLE A1. ICD Codes for LS-Related and Other Cancers

Cancer Type	ICD-10 Codes	ICD-9 Codes
Colorectal	C18 (malignant neoplasm of colon)	153* (malignant neoplasm of colon)
	C19 (malignant neoplasm of rectosigmoid junction)	154.1 (malignant neoplasm of the rectum)
	C20 (malignant neoplasm of the rectum)	154.0 (malignant neoplasm of rectosigmoid junction)
	C21 (malignant neoplasm of anus and anal canal)	154.3 (malignant neoplasm of anus, unspecified site)
	D01.0 (carcinoma in situ of colon)	230.3 (carcinoma in situ of colon)
	D01.1 (carcinoma in situ of rectosigmoid junction)	230.4 (carcinoma in situ of rectum)
	D01.2 (carcinoma in situ of rectum)	230.7 (carcinoma in situ of other and unspecified parts of intestine)
	D01.3 (carcinoma in situ of anus and anal canal)	211.3 (benign neoplasm of colon)
	D01.40 (carcinoma in situ of unspecified part of intestine)	211.4 (benign neoplasm of rectum and anal canal)
	D12 (benign neoplasm of colon, rectum, anus and anal canal)	V12.72 (personal history of colonic polyps)
	K63.5 (polyp of colon)	V10.06 (personal history of malignant neoplasm of rectum, rectosigmoid junction, and anus)
	Z86.010 (personal history of colonic polyps)	V10.05 (personal history of malignant neoplasm of large intestine)
	Z85.04* (personal history of malignant neoplasm of rectum, rectosigmoid junction, and anus)	V18.51 (family history of colonic polyps)
	Z80.0 (family history of malignant neoplasm of digestive organs)	V16.0 (family history of malignant neoplasm of gastrointestinal tract)
	Z83.71 (family history of colonic polyps)	
Endometrial	C55 (malignant neoplasm of uterus, part unspecified)	179 (malignant neoplasm of uterus, part unspecified)
	C54 (malignant neoplasm of corpus uteri, including endometrium)	182 (malignant neoplasm of body of uterus)
	D07.0 (carcinoma in situ of endometrium)	233.2 (carcinoma in situ of other and unspecified parts of uterus)
	D25.0 (submucous leiomyoma of uterus)	218.0 (submucous leiomyoma of uterus)
	D25.1 (intramural leiomyoma of uterus)	218.1 (intramural leiomyoma of uterus)
	D25.2 (subserosal leiomyoma of uterus)	218.2 (subserous leiomyoma of uterus)
	Z85.42 Personal history of malignant neoplasm of other parts of uterus	V10.42 (personal history of malignant neoplasm of other parts of uterus)
	Z80.49 Family history of malignant neoplasm of other genital organs	V16.40 Family history of malignant neoplasm of genital organ, unspecified
		V16.49 Family history of malignant neoplasm of other genital organs

(Continued on following page)

TABLE A1. ICD Codes for LS-Related and Other Cancers (Continued)

Cancer Type	ICD-10 Codes	ICD-9 Codes
Gastric	C16 (malignant neoplasm of stomach)	151 (malignant neoplasm of stomach)
	D01.7 (carcinoma in situ of other specified digestive organs)	230.9 (carcinoma in situ of other and unspecified digestive organs)
	D01.9 (carcinoma in situ of digestive organ, unspecified)	11.1 (benign neoplasm of stomach)
	D13.1 (benign neoplasm of stomach)	V10.04 (personal history of malignant neoplasm of stomach)
	Z85.020 (personal history of malignant carcinoid tumor of stomach)	V10.00 (personal history of malignant neoplasm of gastrointestinal tract, unspecified)
	Z85.028 (personal history of other malignant neoplasm of stomach)	V16.0 (family history of malignant neoplasm of gastrointestinal tract)
	Z85.00 (personal history of malignant neoplasm of unspecified digestive organ)	V18.59 (family history of other digestive disorders)
	Z85.09 (personal history of malignant neoplasm of other digestive organs)	
	Z80.0 (family history of malignant neoplasm of digestive organs)	
	Z83.79 (family history of other diseases of the digestive system)	
Ovarian	C56 (malignant neoplasm of ovary)	183* (malignant neoplasm of ovary and other uterine adnexa)
	C57 (malignant neoplasm of other and unspecified female genital organs)	233.30 (carcinoma in situ of unspecified female genital organs)
	C48 (malignant neoplasm of retroperitoneum and peritoneum)	V10.43 (personal history of malignant neoplasm of ovary)
	D07.30 (carcinoma in situ of unspecified female genital organs)	V16.41 (family history of malignant neoplasm of ovary)
	Z85.43 (personal history of malignant neoplasm of ovary)	
	Z80.41 (family history of malignant neoplasm of ovary)	
Small bowel	C17 (malignant neoplasm of small intestine)	152 (malignant neoplasm of small intestine including duodenum)
	Z85.068 (personal history of other malignant neoplasm of small intestine)	V10.09 (personal history of malignant neoplasm of other gastrointestinal tract)
Hepatobiliary tract	C23 (malignant neoplasm of gallbladder)	156* (malignant neoplasm of gallbladder)
	D01.5 (carcinoma in situ of liver, gallbladder and bile ducts)	230.8 (carcinoma in situ of liver and biliary system)
	C24.9 (biliary tract, unspecified)	

(Continued on following page)

TABLE A1. ICD Codes for LS-Related and Other Cancers (Continued)

Cancer Type	ICD-10 Codes	ICD-9 Codes
Urinary tract	C64, C65, C66, C67, C68 (malignant neoplasms of urinary tract)	189* (malignant neoplasm of kidney and other and unspecified urinary organs)
	D09.0 (carcinoma in situ of bladder)	188* (malignant neoplasm of bladder)
	D09.10 (carcinoma in situ of unspecified urinary organ)	233.7 (carcinoma in situ of bladder)
	D09.19 (carcinoma in situ of other urinary organs)	233.9 (carcinoma in situ of other and unspecified urinary organs)
	Z85.520 (personal history of malignant carcinoid tumor of kidney)	V10.91 (personal history of malignant neuroendocrine tumor)
	Z85.528 (personal history of malignant neoplasm of kidney)	V10.52 (personal history of malignant neoplasm of kidney)
	Z85.53 (personal history of malignant neoplasm of renal pelvis)	V10.53 (personal history of malignant neoplasm of renal pelvis)
	Z85.54 (personal history of malignant neoplasm of ureter)	V10.59 (personal history of malignant neoplasm of other urinary organs)
	Z85.51 (personal history of malignant neoplasm of bladder)	V10.51 (personal history of malignant neoplasm of bladder)
	Z85.59 (personal history of malignant neoplasm of other urinary tract organs)	V16.59 (family history of malignant neoplasm of other urinary organs)
	Z80.59 (family history of malignant neoplasm of other urinary tract organ)	
Brain	C71 (malignant neoplasm of brain)	191 (malignant neoplasm of brain)
	Z85.841 (personal history of malignant neoplasm of brain)	V10.85 (personal history of malignant neoplasm of brain)
Sebaceous gland	C44 (other malignant neoplasms of skin, including sebaceous glands)	173 (other and unspecified malignant neoplasm of skin, including sebaceous glands)
	Z85.828 (personal history of other malignant neoplasm of skin)	V10.83 (personal history of other malignant neoplasm of skin)
Pancreas	C25 (malignant neoplasm of pancreas)	157 (malignant neoplasm of pancreas)
	D13.6 (benign neoplasm of pancreas)	211.6 (benign neoplasm of pancreas, except islets of Langerhans)
	D13.7 (benign neoplasm of endocrine pancreas)	211.7 (benign neoplasm of islets of Langerhans)
	Z85.07 (personal history of malignant neoplasm of pancreas)	V10.09 (personal history of malignant neoplasm of other gastrointestinal tract)
	Z80.0 (family history of malignant neoplasm of digestive organs)	V16.0 (family history of malignant neoplasm of gastrointestinal tract)
Prostate	C61 (malignant neoplasm of prostate)	185 (malignant neoplasm of prostate)
	D07.5 (carcinoma in situ of prostate)	233.4 (carcinoma in situ of prostate)
	Z85.46 (personal history of malignant neoplasm of prostate)	V10.46 (personal history of malignant neoplasm of prostate)
	Z80.42 (family history of malignant neoplasm of prostate)	V16.42 (family history of malignant neoplasm of prostate)
Breast	C50 (malignant neoplasm of breast)	174 (malignant neoplasm of female breast)
	D05.00 (lobular carcinoma in situ of unspecified breast)	233.0 (carcinoma in situ of breast)
	D05.10 (intraductal carcinoma in situ of unspecified breast)	V10.3 (personal history of malignant neoplasm of breast)
	D05.90 (unspecified type of carcinoma in situ of unspecified breast)	V16.3 (family history of malignant neoplasm of breast)
	Z85.3 (personal history of malignant neoplasm of breast)	
	Z80.3 family history of malignant neoplasm of breast)	

Abbreviations: ICD, International Classification of Diseases; LS, Lynch syndrome. Asterisk denotes that all subcodes of an ICD code were queried.

TABLE A2. Amsterdam II Criteria and Revised Bethesda Guidelines for Identification of Individuals at Risk for LS

Amsterdam II Criteria	Revised Bethesda Guidelines for Testing Colorectal Tumors for MSI
1. Three or more relatives with histologically verified LS-related cancer[a], one of whom is a first-degree relative of the other 2, and	1. Diagnosis of colorectal cancer in a patient < 50 years of age, or
2. Cancer involving at least 2 generations, and	2. Presence of synchronous, metachronous colorectal cancers or other LS-related cancers[a] regardless of patient age, or
3. One or more cancer cases diagnosed < 50 years of age	3. Diagnosis of colorectal cancer with high frequency of MSI on the basis of histologic findings (Crohn's-like lymphocytic reaction, mucinous or signet-ring cell differentiation, or medullary growth pattern) in a patient < 60 years of age, or
	4. Diagnosis of colorectal cancer in at least 1 first-degree relative with an LS-related cancer, with at least 1 diagnosis occurring < 50 years of age, or
	5. Diagnosis of colorectal cancer in at least 2 first- or second-degree relatives with LS-related cancers regardless of patient age

Abbreviations: LS, Lynch syndrome; MSI, microsatellite instability.

[a]LS-related cancers include cancers of the colorectum, endometrium, stomach, ovary, pancreas, ureter, renal pelvis, biliary tract, brain, small bowel, and sebaceous glands as well as keratoacanthomas.

TABLE A3. Clinical Characteristics of Individuals Harboring Expected Pathogenic Variants in MMR Genes

Demographic Data			LS-Associated Variant			BC		CRC		EC		LS-Related Cancer		Any Cancer		LS-Related Cancer Types
Age at Enrollment (years)	Sex	Self-Reported Ancestry	Gene	HGVS.c	ClinVar Assertion (review status*)	Personal	Family	Personal	Family	Personal	Family	Personal	Family	Personal	Family	
61	F	EA	*MLH1*	c.184C>T	P (3*)				Y	Y	Y	Y	Y	Y	Y	CRC (family); EC (personal, family)
42	F	H/L	*MLH1*	c.676C>T	P (3*)		Y		Y				Y		Y	CRC (family)
44	F	H/L	*MLH1*	c.790+1G>A	P (3*)	Y								Y		
44	M	SA	*MLH1*	c.790+1G>A	P (3*)			Y	Y			Y	Y	Y	Y	CRC (personal, family)
55	F	H/L	*MLH1*	c.790+1G>A	P (3*)		Y								Y	
57	M	AA	*MLH1*	c.790+1G>A	P (3*)											
61	F	EA	*MLH1*	c.790+1G>A	P (3*)				Y		Y		Y		Y	CRC (family); EC (family)
63	M	AA	*MLH1*	c.790+1G>A	P (3*)							Y		Y		Gastric (personal)
43	F	H/L	*MLH1*	c.1297G>T	P/LP (2*)											
62	M	AA	*MLH1*	c.1989+1G>A	P (3*)										Y	
74	M	EA	*MLH1*	c.2080G>T	NA								Y	Y	Y	Ovarian (family)
40	F	AA	*MLH1*	c.793C>T	P (3*)						Y		Y		Y	EC (family)
41	M	H/L	*MSH2*	c.478C>T	P (3*)				Y			Y	Y	Y	Y	CRC (family); small bowel (personal); sebaceous (personal)
43	F	H/L	*MSH2*	c.1387-2A>G	NA		Y								Y	
26	F	AA	*MSH2*	c.1697del	P (1*)											
37	F	AA	*MSH2*	c.1697del	P (1*)										Y	
51	F	AA	*MSH2*	c.1697del	P (1*)											
55	F	H/L	*MSH2*	c.1697del	P (1*)		Y							Y	Y	
61	F	Mult	*MSH2*	c.1697del	P (1*)	Y						Y		Y		Fallopian tube (personal)
69	M	Mult	*MSH2*	c.1697del	P (1*)											
73	F	H/L	*MSH2*	c.1697del	P (1*)											
76	F	Mult	*MSH2*	c.1697del	P (1*)					Y		Y		Y	Y	EC (personal)
40	M	Mult	*MSH2*	c.1906G>C	P (3*)				Y		Y		Y		Y	CRC (family); EC (family)
69	M	EA	*MSH2*	c.1906G>C	P (3*)				Y				Y	Y	Y	CRC (family); gastric (family)

(Continued on following page)

TABLE A3. Clinical Characteristics of Individuals Harboring Expected Pathogenic Variants in MMR Genes (Continued)

Demographic Data			LS-Associated Variant			BC		CRC		EC		LS-Related Cancer		Any Cancer		LS-Related Cancer Types
Age at Enrollment (years)	Sex	Self-Reported Ancestry	Gene	HGVS.c	ClinVar Assertion (review status*)	Personal	Family	Personal	Family	Personal	Family	Personal	Family	Personal	Family	
42	F	EA	*MSH2*	c.2632G>T	NA								Y		Y	Ovarian (family)
58	F	EA	*MSH6*	c.260+2T>A	LP (0*)					Y		Y		Y		EC (personal)
69	F	AA	*MSH6*	c.892C>T	P (3*)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	CRC (personal, family); EC (personal); pancreatic (family)
60	F	H/L	*MSH6*	c.1100A>G	P (3*)	Y								Y	Y	
58	F	AA	*MSH6*	c.1350_1351delAT	P (3*)							Y		Y		Brain (personal)
65	M	H/L	*MSH6*	c.1705_1706delTT	P (2*)			Y				Y		Y		CRC (personal)
47	F	AA	*MSH6*	c.1842delC	P (2*)											
71	F	Mult	*MSH6*	c.2079dupA	P (2*)	Y				Y		Y		Y		EC (personal)
40	M	SA	*MSH6*	c.3226C>T	LP (3*)									Y		
57	F	H/L	*MSH6*	c.3226C>T	LP (3*)											
75	M	AA	*MSH6*	c.3261del	P (3*)									Y		
30	F	AA	*MSH6*	c.3261dupC	P (3*)											
45	F	AA	*MSH6*	c.3261dupC	P (3*)										Y	
44	F	H/L	*MSH6*	c.3311_3312delTT	P (3*)		Y		Y				Y		Y	CRC (family); renal (family)
42	F	AA	*MSH6*	c.3732_3735dup	NA											
77	M	EA	*MSH6*	c.3959_3962delCAAG	P (3*)										Y	
38	M	EA	*MSH6*	c.3984_3987dup	P (3*)											
24	F	M	*PMS2*	c.2444C>T	LP (3*)											
52	F	AA	*PMS2*	c.2444C>T; c.2331dup	LP (3*); NA											
52	M	AA	*PMS2*	c.2444C>T; c.2331dup	LP (3*); NA											
59	M	AA	*PMS2*	c.2444C>T	LP (3*)											
30	F	AA	*PMS2*	c.2404C>T	P (3*)											

(Continued on following page)

TABLE A3. Clinical Characteristics of Individuals Harboring Expected Pathogenic Variants in MMR Genes (Continued)

Demographic Data			LS-Associated Variant			BC		CRC		EC		LS-Related Cancer		Any Cancer		LS-Related Cancer Types
Age at Enrollment (years)	Sex	Self-Reported Ancestry	Gene	HGVS.c	ClinVar Assertion (review status*)	Personal	Family	Personal	Family	Personal	Family	Personal	Family	Personal	Family	
46	M	AA	*PMS2*	c.2404C>T	P (3*)											
52	M	Mult	*PMS2*	c.2404C>T	P (3*)											
42	F	EA	*PMS2*	c.1927C>T	P (3*)						Y		Y		Y	EC (family)
56	F	EA	*PMS2*	c.1831dupA	P (3*)				Y				Y	Y	Y	CRC (family)
57	M	AA	*PMS2*	c.1591G>T	P (1*)										Y	
27	F	O	*PMS2*	c.1119_1122delTCAG	NA											
71	M	EA	*PMS2*	c.989-1G>T	P (2*)								Y		Y	Ovarian (family)
67	F	H/L	*PMS2*	c.949C>T	P (3*)											
54	M	AA	*PMS2*	c.943C>T	P (3*)											
77	M	EA	*PMS2*	c.943C>T	P (3*)									Y	Y	
36	M	H/L	*PMS2*	c.903G>T	LP (3*)		Y									
68	M	EA	*PMS2*	c.903G>T	LP (3*)									Y	Y	
70	F	O	*PMS2*	c.825A>G	LP (2*)											
46	M	AA	*PMS2*	c.809C>G	P (2*)											
28	F	ESA	*PMS2*	c.325dupG	P (2*)											
64	F	EA	*PMS2*	c.1A>G	LP (1*)				Y	Y	Y	Y	Y	Y	Y	CRC (family); EC (personal, family)
28	M	O	*PMS2*	c.137G>T	LP (3*)		Y								Y	
52	M	O	*PMS2*	c.137G>T	LP (3*)				Y				Y		Y	CRC (family)
58	F	EA	*PMS2*	c.137G>T	LP (3*)					Y	Y	Y	Y	Y	Y	EC (personal, family)
59	M	H/L	*PMS2*	c.137G>T	LP (3*)		Y								Y	
65	F	O	*PMS2*	c.137G>T	LP (3*)											
76	M	Mult	*PMS2*	c.137G>T	LP (3*)									Y	Y	
72	M	AA	*PMS2*	c.23+1G>T	P/LP (2*)											
35	M	H/L	*PMS2*	c.23+1G>T	LP (3*)								Y		Y	Gastric (family)

Abbreviations: AA, African ancestry; BC, breast cancer; CRC, colorectal cancer; EA, European ancestry; EC, endometrial cancer; ESA East/Southeast Asian ancestry; H/L, Hispanic/Latinx ancestry; LP, likely pathogenic; LS, Lynch syndrome; MMR, mismatch repair; Mult, multiple selected ancestries; NA, not available; O, other ancestry; P, pathogenic; SA, South Asian ancestry.