

# 1 Ability of a polygenic risk score to refine colorectal cancer 2 risk in Lynch syndrome

## 3 Authors

4 Nuria Dueñas<sup>\*1,2,3</sup>, Hannah Klinkhammer<sup>\*4,5</sup>, Nuria Bonifaci<sup>\*2,3</sup>, Isabel Spier<sup>6,7,8</sup>,  
5 Andreas Mayr<sup>4</sup>, Emadeldin Hassanin<sup>5</sup>, Anna Díez-Villanueva<sup>9,10,11</sup>, Victor  
6 Moreno<sup>9,10,11,12</sup>, Marta Pineda<sup>1,2,3</sup>, Carlo Maj<sup>5</sup>, Gabriel Capellá<sup>1,2,3,8</sup>, Stefan Aretz<sup>\*\*\*6,7,8</sup>,  
7 Joan Brunet<sup>\*\*\*1,2,3,8,13</sup>

8 1. Hereditary Cancer Program, Catalan Institute of Oncology, Hospitalet de Llobregat, Barcelona, Spain. [nduenas@iconcologia.net](mailto:nduenas@iconcologia.net),  
9 [mpineda@iconcologia.net](mailto:mpineda@iconcologia.net), [gcapella@idibell.cat](mailto:gcapella@idibell.cat), [jbrunet@iconcologia.net](mailto:jbrunet@iconcologia.net)

10 2. Program in Molecular Mechanisms and Experimental Therapy in Oncology (ONCOBELL), Bellvitge Biomedical Research Institute  
11 (IDIBELL), Hospitalet de Llobregat, Barcelona, Spain. [nduenas@iconcologia.net](mailto:nduenas@iconcologia.net), [nbonifaci@idibell.cat](mailto:nbonifaci@idibell.cat), [mpineda@iconcologia.net](mailto:mpineda@iconcologia.net),  
12 [gcapella@idibell.cat](mailto:gcapella@idibell.cat), [jbrunet@iconcologia.net](mailto:jbrunet@iconcologia.net)

13 3. Biomedical Research Centre Network for Oncology (CIBERONC), Instituto Salud Carlos III, Madrid, Spain.  
14 [nduenas@iconcologia.net](mailto:nduenas@iconcologia.net), [mpineda@iconcologia.net](mailto:mpineda@iconcologia.net), [gcapella@idibell.cat](mailto:gcapella@idibell.cat), [jbrunet@iconcologia.net](mailto:jbrunet@iconcologia.net)

15 4. Institute for Medical Biometry, Informatics and Epidemiology, Medical Faculty, University of Bonn, Germany.  
16 [klinkhammer@imbie.uni-bonn.de](mailto:klinkhammer@imbie.uni-bonn.de), [amayr@uni-bonn.de](mailto:amayr@uni-bonn.de)

17 5. Institute for Genomic Statistics and Bioinformatics, Medical Faculty, University of Bonn, Germany. [klinkhammer@imbie.uni-bonn.de](mailto:klinkhammer@imbie.uni-bonn.de),  
18 [emadsf1992@gmail.com](mailto:emadsf1992@gmail.com), [cmaj@uni-bonn.de](mailto:cmaj@uni-bonn.de)

19 6. Institute of Human Genetics, Medical Faculty, University of Bonn, Germany. [isabel.spier@uni-bonn.de](mailto:isabel.spier@uni-bonn.de), [stefan.aretz@uni-bonn.de](mailto:stefan.aretz@uni-bonn.de),  
20 [stefan.aretz@uni-bonn.de](mailto:stefan.aretz@uni-bonn.de),

21 7. National Center for Hereditary Tumor Syndromes, University Hospital Bonn, Germany. [isabel.spier@uni-bonn.de](mailto:isabel.spier@uni-bonn.de),  
22 [stefan.aretz@uni-bonn.de](mailto:stefan.aretz@uni-bonn.de)

23 8. European Reference Network on Genetic Tumour Risk Syndromes (ERN GENTURIS). [isabel.spier@uni-bonn.de](mailto:isabel.spier@uni-bonn.de),  
24 [gcapella@idibell.cat](mailto:gcapella@idibell.cat), [mpineda@iconcologia.net](mailto:mpineda@iconcologia.net), [stefan.aretz@uni-bonn.de](mailto:stefan.aretz@uni-bonn.de), [jbrunet@iconcologia.net](mailto:jbrunet@iconcologia.net)

25 9. Oncology Data Analytics Program (ODAP), Catalan Institute of Oncology, Hospitalet de Llobregat, Barcelona, Spain.  
26 [adiez@iconcologia.net](mailto:adiez@iconcologia.net), [v.moreno@iconcologia.net](mailto:v.moreno@iconcologia.net)

27 10. Colorectal Cancer Group (ONCOBELL), Bellvitge Biomedical Research Institute (IDIBELL), Hospitalet de Llobregat, Barcelona,  
28 Spain. [adiez@iconcologia.net](mailto:adiez@iconcologia.net), [v.moreno@iconcologia.net](mailto:v.moreno@iconcologia.net)

29 11. Biomedical Research Centre Network for Epidemiology and Public Health (CIBERESP), Instituto Salud Carlos III, Madrid, Spain.  
30 [adiez@iconcologia.net](mailto:adiez@iconcologia.net), [v.moreno@iconcologia.net](mailto:v.moreno@iconcologia.net)

31 12. Department of Clinical Sciences, Faculty of Medicine and Health Sciences and Universitat de Barcelona Institute of Complex  
32 Systems (UBICS), University of Barcelona, Barcelona, Spain. [v.moreno@iconcologia.net](mailto:v.moreno@iconcologia.net)

33 13. Hereditary Cancer Program, Catalan Institute of Oncology-IDBIGI, Girona, Spain. [jbrunet@iconcologia.net](mailto:jbrunet@iconcologia.net)

34 \* These authors contributed equally to this work and share first authorship

35 \*\* These authors contributed equally to this work and share last authorship

36 † These authors act as corresponding authors

37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59

**ABSTRACT:**

**Background:** Polygenic risk scores (PRS) have been used to stratify colorectal cancer (CRC) risk in the general population, whereas its role in Lynch syndrome (LS), the most common type of hereditary CRC, is still conflicting. We aimed to assess the ability of PRS to refine CRC risk prediction in European-descendant LS individuals.

**Methods:** 1,465 LS individuals (557 *MLH1*, 517 *MSH2/EPCAM*, 299 *MSH6*, and 92 *PMS2*) and 5,656 CRC-free population-based controls from two independent cohorts were included. A 91-Single Nucleotide Polymorphism PRS was applied. A Cox proportional hazard regression model with “family” as a random effect and a logistic regression analysis, followed by a meta-analysis combining both cohorts were conducted.

**Results:** Overall, we did not observe a statistically significant association between PRS and CRC risk in the entire cohort. Nevertheless, PRS was significantly associated with a slightly increased risk of CRC or advanced adenoma (AA), in those with CRC diagnosed < 50 years, and in individuals with multiple CRCs or AAs diagnosed < 60 years.

**Conclusion:** The PRS may slightly influence CRC risk in LS individuals, in particular in more extreme phenotypes such as early-onset disease. However, the study design and recruitment strategy strongly influence the results of PRS studies. A separate analysis by genes and its combination with other genetic and non-genetic risk factors will help refine its role as a risk modifier in LS.

60

61

62

63 **KEY MESSAGES:**

64 WHAT IS ALREADY KNOWN ON THIS TOPIC?

- 65 - Great variability in the incidence of CRC has been described in LS individuals,  
66 even within the same family.
- 67 - Polygenic risk scores (PRS) can help stratify colorectal cancer risk and, thus,  
68 adjust surveillance or treatment procedures.

69 WHAT THIS STUDY ADDS

- 70 - PRS performed on family-based registries slightly influences CRC risk in  
71 subgroups of LS individuals, even though with weak effects.
- 72 - Our study showed a weak association of PRS with multiple and young CRC  
73 cases, pointing to a possible risk-modifying role in extreme phenotypes.

74 HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- 75 - Gene-based PRS analysis and its combination with other genetic and non-  
76 genetic factors may contribute to refining cancer risk in LS patients.

77

## 78 INTRODUCTION

79 Colorectal cancer (CRC) is the third most incident cancer overall and the second  
80 leading cause of cancer-related death worldwide. Incidence rates are four times higher  
81 in the Global North, associated with lifestyle and dietary risk factors<sup>1</sup>.

82 About 5% of CRC is considered hereditary due to highly penetrant pathogenic germline  
83 variants in cancer-predisposing genes<sup>2,3</sup>. The main cause of hereditary CRC is Lynch  
84 Syndrome (LS), with an estimated carrier frequency in the general population of around  
85 1:279<sup>4</sup>. It is characterised as an autosomal dominant inherited defect in any of the  
86 mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*) or *EPCAM* gene  
87 deletions, resulting in silencing of the *MSH2* gene in epithelial tissues<sup>5</sup>. Median CRC  
88 cumulative incidences at 75 years show an important variability according to mutated  
89 gene and gender: 48/57%, 47/51%, and 18/20% for male and female carriers of  
90 mutations in *MLH1*, *MSH2*, and *MSH6*, respectively, and 10% for both genders in  
91 carriers of mutations in *PMS2*<sup>6</sup>. Differences in CRC risk have also been identified based  
92 on the ethnic or geographical origin of carriers, with lower risks reported for European  
93 vs. American and Australasian individuals<sup>7</sup>. Moreover, LS carriers have an increased  
94 risk of developing multiple CRCs, CRC at a younger age, and other LS-associated  
95 cancers such as endometrial (EC) or ovarian cancer<sup>6</sup>.

96 In LS, as in other hereditary cancer predisposition syndromes characterised by  
97 incomplete penetrance, one of the main challenges is to identify which risk-modifying  
98 factors may modulate the expression of the cancer syndrome<sup>7,8</sup>. In recent years,  
99 multiple, common, low-penetrance CRC risk variants have been identified through  
100 genome-wide association studies (GWAS)<sup>9-11</sup>. Each risk allele individually confers a

101 small risk, but their combined effect as a polygenic risk score (PRS) exhibits significant  
102 risks of developing CRC in the general population. Being in the highest PRS percentiles  
103 was shown to increase the risk of CRC two- to seven-fold<sup>10,12-16</sup>. Moreover, PRS might  
104 be particularly relevant in patients with a more extreme, i.e., severe, phenotype: a study  
105 performed in individuals diagnosed with CRC before 50 years of age (early-onset  
106 disease) demonstrated the existence of an interaction between PRS and CRC risk, with  
107 an odds ratio (OR) of 3.73 (3.28-4.24) in the highest PRS quartile<sup>17</sup>. Another study on  
108 familial CRC (individuals who fulfil Amsterdam or Bethesda criteria without a pathogenic  
109 germline MMR variant) identified an increased CRC risk in individuals in the highest 5%  
110 of the PRS distribution, with an OR of 4.89 (2.37-10.07)<sup>18</sup>.

111 To date, the modulating effect of PRS on CRC risk in LS individuals is still controversial.  
112 Two studies on a population-based repository from the UK Biobank (UKBB) including 76  
113 and 388 LS carriers, respectively, reported that PRS may strongly influence CRC  
114 risk<sup>16,19</sup>; however, another analysis of the clinic-based registry of the Colon Cancer  
115 Family Registry (CCFR), including 826 European-descendant LS individuals, found no  
116 evidence of association, irrespective of sex or mutated gene<sup>20</sup>.

117 Our objective was to evaluate whether differences in CRC penetrance in European-  
118 descendant LS individuals can, in part, be explained by the accumulation of low-risk  
119 CRC alleles using a validated set of 91 SNPs for PRS analysis.

120

## 121 **METHODS**

### 122 **Study participants**

#### 123 LS individuals:

124 A total of 1,465 European-descendant individuals with genetically confirmed LS (557  
125 *MLH1*, 517 *MSH2/EPCAM*, 299 *MSH6*, and 92 *PMS2*) from two independent cohorts  
126 were included: 918 LS individuals (353 families) identified at the Catalan Institute of  
127 Oncology (ICO; Spain) and 547 LS individuals (392 families) from the University  
128 Hospital of Bonn (UKB; Germany). Patients were recruited based on the fulfilment of  
129 Bethesda or Amsterdam criteria or via an EC and CRC-based LS screening programme  
130 (since 2016 at the ICO)<sup>21</sup>. Patients included were affected index patients and affected or  
131 unaffected carriers among the relatives identified through cascade testing. In the ICO  
132 LS cohort, there was a lower percentage of pathogenic *MSH2* variant carriers (mainly  
133 due to the existence of *MLH1* founder mutations in the ICO series) and a higher  
134 percentage of pathogenic *MSH6* and *PMS2* variant carriers (mainly identified through  
135 an EC/CRC-based LS screening) when compared with the UKB LS cohort. In addition,  
136 the ICO cohort included a higher proportion of non-index individuals. There were no  
137 significant differences in the distribution of affected genes between early-onset cases  
138 and the entire cohort (**Table 1**).

139 All patients gave informed consent and the internal Ethics Committee approved this  
140 study.

#### 141 Non-LS individuals:

142 A total of 5,656 unselected CRC-free individuals from the same population were  
143 included in the analysis (CRC-free population controls): 1,642 individuals from Spain  
144 and 4,014 from Germany. The controls from Spain included individuals from the  
145 CRCGEN study and individuals participating in a population-based CRC screening  
146 program, most of whom had a positive faecal immunochemical test (FIT) result and a

147 colonoscopy without cancer or advanced adenoma, as described elsewhere<sup>22</sup>. The  
 148 German controls were drawn from the population-based Heinz Nixdorf RECALL (HNR)  
 149 study (Risk Factors, Evaluation of Coronary Calcification, and Lifestyle) as described  
 150 recently<sup>23</sup> (**Table 1**).

	Lynch syndrome			Population controls		
	All series	ICO	UKB	All series	Spain	Germany
<b>Total n of individuals</b>	<b>1465</b> <b>(100%)</b>	918 (62.7%)	547 (37.3%)	<b>5656</b> <b>(100%)</b>	1642 (29.0%)	4014 (70.1%)
<b>Age</b>						
Mean age at censor (range)	<b>45.6 (12-93)</b>	47.9 (16-93)	41.6 (12-86)	<b>71.0 (24-94)</b>	62.4 (24-92)	74.5 (49-94)
<b>Gender</b>						
Male	<b>707</b> <b>(48.3%)</b>	409 (44.6%)	298 (54.5%)	<b>2813</b> <b>(49.7%)</b>	835 (50.9%)	1978 (49.3%)
Female	<b>758</b> <b>(51.7%)</b>	509 (55.4%)	249 (45.5%)	<b>2843</b> <b>(50.3%)</b>	807 (49.2%)	2036 (50.7%)
<b>Mutated gene</b>						
<i>MLH1</i>	<b>557</b> <b>(38.0%)</b>	346 (37.7%)	211 (38.6%)	-	-	-
<i>MSH2/EPCAM</i>	<b>517</b> <b>(35.3%)</b>	247 (26.9%)	270 (49.4%)	-	-	-
<i>MSH6</i>	<b>299</b> <b>(20.4%)</b>	250 (27.2%)	49 (9.0%)	-	-	-
<i>PMS2</i>	<b>92</b> <b>(6.3%)</b>	75 (8.2%)	17 (3.1%)	-	-	-
<b>Index case</b>						
Yes	<b>590</b> <b>(40.3%)</b>	193 (21.0%)	397 (72.6%)	-	-	-
No	<b>875</b> <b>(59.7%)</b>	725 (79.0%)	150 (27.4%)	-	-	-

151 **TABLE 1: Main characteristics of the population studied.**  
 152 *ICO*: Catalan Institute of Oncology, Spain; *UKB*: University Hospital of Bonn, Germany  
 153

154 **Data collection**

155 Clinical data included demographic, personal and oncologic history, and follow-up  
 156 carried out from birth to 06/2021. In LS individuals, histories of colorectal polyps or other

157 LS-related cancers were also collected. Data supporting the results were stored in local  
158 databases at both centres.

### 159 **SNP selection**

160 The selected SNPs (n=95) and associated risks were obtained from the meta-analysis  
161 of CRC risk alleles performed by Huyghe *et al.*,<sup>10</sup> (**Tab. S1**) and were commonly used to  
162 study sporadic CRC risk at the initiation of the study<sup>16,19</sup>. Individual CRC risk-associated  
163 SNPs reached independent genome-wide significance ( $p < 5 \times 10^{-8}$ ) in a large-scale  
164 GWAS.

### 165 **Genotyping**

166 ICO blood DNA samples were genotyped with the Illumina Global Screening Array-24  
167 (GSA) v2.0 and v3.0 ([https://emea.illumina.com/science/consortia/human-](https://emea.illumina.com/science/consortia/human-consortia/global-screening-consortium.html)  
168 [consortia/global-screening-consortium.html](https://emea.illumina.com/science/consortia/human-consortia/global-screening-consortium.html)) and UKB samples with GSA v3.0. Of note,  
169 48% of the ICO population of CRC-free individuals were previously included in the  
170 meta-analysis by Huyghe *et al.*,<sup>10</sup> however, they corresponded to ~1% of the total  
171 number of cases and controls in the analysis. Details regarding quality control  
172 procedures and correlation between arrays have been described previously<sup>18,22</sup>.

173 Non-European-descendant individuals were excluded from the analysis. To assess  
174 ethnicity, Spanish samples were compared with 1,397 HapMap samples, while German  
175 samples were compared with 1k genome samples. Classification into different ethnicity  
176 groups was performed by selecting ancestry-informative marker SNPs (AIM SNPs) and  
177 using a principal components analysis (PCA) approach.

### 178 **Imputation**



179 Thirteen and eighteen of the 95 variants of interest were included in the Illumina GSA-  
180 24 v2.0 and v3.0, respectively. Variants not directly genotyped by the corresponding  
181 arrays were imputed in the ICO with the Michigan Imputation Server (HRC version  
182 r1.1.2016 panel)<sup>24</sup> and in the UKB with a comparable pipeline based on the  
183 bioinformatic tools bcftools, minimac, and vcftools, using GRCh37 as the reference  
184 genome (1000 Genomes, phase 3, v5)<sup>25</sup>. Missing variants and variants with an  
185 imputation quality ( $r^2$ ) <0.3 (considering all genotyped samples) were not included in the  
186 final PRS analysis, which resulted in the exclusion of rs6058093, rs35470271,  
187 rs145364999, and rs755229494 (**Tab. S1**).

#### 188 **PRS calculation**

189 For each participant, PRS was computed using the PLINK score function<sup>26</sup>, based on  
190 the 91 quality-controlled CRC risk alleles (coded as 0, 1, or 2) and effect sizes as  
191 reported by Huyghe *et al.*, (PRS) and averaged over the number of observed variants  
192 per individual<sup>10</sup> (wPRS). To ease interpretation, wPRS values were rescaled (rwPRS) to  
193 indicate risk per allele (using the ratio of non-averaged PRS and wPRS values in  
194 controls as a scaling factor) as previously reported<sup>18</sup>.

#### 195 **Study events**

196 Two events were considered: i) CRC and ii) advanced adenoma (AA) (adenoma with  
197 significant villous features (>25%), size  $\geq 1.0$  cm, high-grade dysplasia, or early invasive  
198 cancer).

199 Two subgroups were defined for the primary analysis: affected individuals (CRC and  
200 CRC or AA LS individuals) and unaffected individuals (CRC-free or CRC-free and AA-  
201 free LS individuals). For the subanalysis of multiple CRCs, three subgroups were

202 defined: multiple events (multiple CRC and multiple CRC or AA LS individuals), single  
203 event (single CRC and single CRC or AA LS individuals) and no-event (CRC-free and  
204 CRC-free and AA-free LS individuals). CRC-free population controls were only  
205 compared to CRC or multiple CRC LS individuals when considering CRC as a study  
206 event as no reliable information was available regarding AA in this population. (**Tab. S2,**  
207 **Tab. S3, and Fig. S1**).

## 208 **Statistical methods**

209 Statistical analyses and graphical representations were conducted with R version 4.0.5.  
210 For the primary analysis, the association of rwPRS with CRC and CRC or AA risk was  
211 tested by considering time to CRC (years since birth to event of study) using a Cox  
212 proportional hazard regression model with family as a random effect (frailty model).  
213 Observations in the control cohort were right censored at the age of last contact and  
214 CRC diagnosis (yes/no) was used as an event variable. The date of the first  
215 polypectomy for adenoma was used as a time-dependent variable. Additionally, sex,  
216 birth cohort (<1940, 1940-49, 1950-59, 1960-69, 1970-79, >1980), and other LS-related  
217 cancers were included as covariates.

218 For the subanalysis of multiple CRCs, the association of rwPRS with multiple CRCs or  
219 AAs was tested using a mixed effects logistic regression, including age, sex, birth  
220 cohort, polypectomy before the second CRC, the occurrence of other cancers, and  
221 family (random effect) as covariates.

222 Results from both cohorts (ICO and UKB) were combined and analysed via a fixed-  
223 effect meta-analysis and the inverse-variance method. The combined rwPRS effect was  
224 estimated as the weighted average of the estimates of the individual studies and

225 weights were derived as the inverse of the variance of the individual effect estimate.  
226 The population was stratified according to rwPRS tertiles using the medium category as  
227 a reference. Additionally, to test for heterogeneity, Cochran's Q was computed on the  
228 derived estimates and a  $\chi^2$ -test with one degree of freedom was performed. Results  
229 with  $p$ -values  $<0.05$  in the test for heterogeneity were not considered. The meta-  
230 analysis was conducted via R package meta<sup>27</sup>. To correct for multiple testing, analyses  
231 were grouped by study event and control group, and  $p$ -values inside these groups were  
232 corrected via false discovery rate (FDR) correction<sup>28</sup>. Only results with  $p$ -values  $<0.05$   
233 after FDR correction ( $p$ -FDR) were considered statistically significant.

234

## 235 RESULTS

236 No differences in PRS distribution were observed when comparing CRC-free LS  
237 individuals and CRC-free population controls in any of the cohorts studied (**Fig. S2**).

### 238 Primary analysis

#### 239 CRC as the study event

240 A statistically significant association of rwPRS with CRC risk was found in LS carriers  
241 under 50 years of age compared with CRC-free LS individuals (HR=1.022 [1.007-  
242 1.038],  $p$ -FDR=0.01). We found a tendency for an association of rwPRS with CRC risk  
243 in the entire cohort and *MSH6* variant carriers. We found no statistically significant  
244 association of rwPRS with CRC risk when comparing CRC LS to CRC-free population  
245 individuals (**Table 2 and Tab. S4**).

246 Additionally, rwPRS tended to be associated with higher CRC risk in *MSH2/EPCAM*  
247 (tertile low: HR=0.716 [0.505-1.016],  $p$ -FDR= 0.53 vs. tertile high: HR=1.058 [0.769-

248 1.455]  $p$ -FDR=0.96) and *MSH6* variant carriers (tertile low: HR=0.617 [0.299-1.271],  $p$ -  
249 FDR=0.53 vs. tertile high: HR=1.594 [0.929-2.735],  $p$ -FDR=0.53), however, results were  
250 not statistically significant (**Figure 1**).

#### 251 CRC or AA as study events

252 A statistically significant association of rwPRS with CRC or AA risk was observed in the  
253 entire cohort (HR=1.019 [1.005-1.032],  $p$ -FDR=0.03) and in LS carriers under 50 years  
254 of age (HR=1.022 [1.006-1.038],  $p$ -FDR=0.006). We observed a tendency for an  
255 association of rwPRS with CRC or AA risk in *MSH2/EPCAM* and *MSH6* carriers (**Table**  
256 **2 and Tab. S5**).

257 Even though no statistically significant associations were observed (**Figure 2**), rwPRS  
258 tended to be associated with a higher risk of CRC and AA in *MSH6* variant carriers  
259 (tertile low: HR=0.669 [0.322-1.393],  $p$ -FDR=0.57 vs. tertile high: HR=2.015 [1.169-  
260 3.471],  $p$ -FDR=0.39).

261

262

Subgroup	Cases	Controls	Number of events	Number of controls	HR (range)	p-value	FDR-corrected p-value
Entire cohort	CRC						
	LS	CRC-free LS	701	727	1.016 (1.003-1.030)	0.019	0.056
<i>MLH1</i>	CRC LS	CRC-free LS	302	237	1.006 (0.986-1.026)	0.579	0.579
<i>MSH2/EPCAM</i>	CRC LS	CRC-free LS	275	239	1.016 (0.994-1.038)	0.154	0.305
<i>MSH6</i>	CRC LS	CRC-free LS	89	195	1.052 (1.012-1.092)	0.010	0.056
<i>PMS2</i>	CRC LS	CRC-free LS	35	56	1.037 (0.962-1.119)	0.339	0.407
Age <50y	CRC LS	CRC-free LS	501	727	<b>1.019 (1.003-1.035)</b>	0.019	<b>0.019</b>
Entire cohort	CRC LS	CRC-free population controls	701	5579	1.012 (0.999-1.024)	0.069	0.347
<i>MLH1</i>	CRC LS	CRC-free population controls	302	5579	1.007 (0.988-1.027)	0.470	0.564
<i>MSH2/EPCAM</i>	CRC LS	CRC-free population controls	275	5579	1.014 (0.994-1.034)	0.173	0.347
<i>MSH6</i>	CRC LS	CRC-free population controls	89	5579	1.015 (0.980-1.051)	0.400	0.564
<i>PMS2</i>	CRC LS	CRC-free population controls	35	5579	1.002 (0.928-1.081)	0.965	0.965
Age <50y	CRC LS	CRC-free population controls	501	5579	1.011 (0.996-1.025)	0.162	0.247
Entire cohort	CRC or AA LS	CRC-free and AA-free LS	733	695	<b>1.019 (1.005-1.032)</b>	0.005	<b>0.033</b>
<i>MLH1</i>	CRC or AA LS	CRC-free and AA-free LS	314	224	1.009 (0.992-1.025)	0.309	0.370
<i>MSH2/EPCAM</i>	CRC or AA LS	CRC-free and AA-free LS	292	222	1.016 (1.001-1.031)	0.037	0.056
<i>MSH6</i>	CRC or AA LS	CRC-free and AA-free LS	92	193	1.026 (1.004-1.049)	0.022	0.056
<i>PMS2</i>	CRC or AA LS	CRC-free and AA-free LS	35	56	1.010 (0.985-1.035)	0.447	0.447
Age <50y	CRC or AA LS	CRC-free and AA-free LS	524	695	<b>1.022 (1.006-1.038)</b>	0.007	<b>0.006</b>

**TABLE 2. Meta-analysis results of the association between Polygenic risk score (rwPRS) and colorectal cancer (CRC) risk or CRC and advanced adenoma (AA) risk.**

CRC LS: LS individuals previously diagnosed of CRC; *CRC-free* LS: LS individuals with no previous CRC diagnosis; CRC-free population control: individuals from the population without previous diagnosis of CRC; *CRC or AA* LS: LS individuals with a previous diagnosis of CRC or AA, whichever occurred first; *CRC-free and AA-free* LS: LS individuals not diagnosed with CRC or AA; *HR*: Hazard ratio; *FDR* False discovery rate; *age <50y*: cases with CRC <50 years of age.

263  
264  
265  
266  
267  
268

269 **Subanalysis: multiple CRCs**

270 Multiple CRCs as the study event

271 No statistically significant association of rwPRS with multiple CRC risk was observed  
272 (irrespective of the gene involved) when comparing single-CRC LS cases, CRC-free LS  
273 individuals, or CRC-free population controls (**Table 3 and Tab. S6**). These analyses  
274 could not be performed in *MSH6* or *PMS2* carriers due to the low sample sizes.

275 Multiple CRCs or AAs as study events

276 A significant association of rwPRS with multiple CRC or AA risk was observed in LS  
277 individuals under 60 years when comparing with single-CRC or AA LS (HR=1.057  
278 [1.010-1.100],  $p$ -FDR=0.04) and CRC-free and AA-free LS (HR=1.043 [1.008-1.079],  $p$ -  
279 FDR=0.03). A tendency was observed for an association of rwPRS with multiple CRC or  
280 AA risk in the entire cohort and *MLH1* carriers (**Table 3 and Tab. S7**). These analyses  
281 could not be performed in *MSH6* and *PMS2* carriers due to the low sample size.

Subgroup	Cases	Controls	Number of events	Number of controls	OR (range)	p-value	FDR-corrected p-value
Entire cohort	Multiple CRC LS	Single CRC LS	257	437	1.021 (0.992-1.051)	0.153	0.204
<i>MLH1</i>	Multiple CRC LS	Single CRC LS	112	185	1.044 (0.997-1.094)	0.067	0.204
<i>MSH2/EPCAM</i>	Multiple CRC LS	Single CRC LS	113	160	0.959 (0.907-1.014)	0.144	0.204
Age <60y	Multiple CRC LS	Single CRC LS	206	437	1.022 (0.992-1.053)	0.152	0.303
Entire cohort	Multiple CRC LS	CRC-free LS	257	727	1.025 (0.984-1.068)	0.227	0.454
<i>MLH1</i>	Multiple CRC LS	CRC-free LS	112	237	1.026 (0.970-1.084)	0.370	0.462
<i>MSH2/EPCAM</i>	Multiple CRC LS	CRC-free LS	113	239	1.023 (0.963-1.087)	0.462	0.462
Age <60y	Multiple CRC LS	CRC-free LS	206	727	1.032 (0.996-1.070)	0.084	0.167
Entire cohort	Multiple CRC LS	CRC-free population controls	257	5630	1.022 (0.991-1.055)	0.170	0.380
<i>MLH1</i>	Multiple CRC LS	CRC-free population controls	112	5630	1.027 (0.983-1.073)	0.228	0.380
<i>MSH2/EPCAM</i>	Multiple CRC LS	CRC-free population controls	113	5630	1.002 (0.958-1.047)	0.938	0.938
Age <60y	Multiple CRC LS	CRC-free population controls	206	5630	1.019 (0.987-1.052)	0.240	0.480
Entire cohort	Multiple CRC or AA LS	Single CRC or AA LS	272	454	1.030 (1.000-1.060)	0.035	0.104
<i>MLH1</i>	Multiple CRC or AA LS	Single CRC or AA LS	112	198	1.048 (1.000-1.100)	0.042	0.104
<i>MSH2/EPCAM</i>	Multiple CRC or AA LS	Single CRC or AA LS	123	167	0.992 (0.943-1.040)	0.740	0.740
Age <60y	Multiple CRC or AA LS	Single CRC or AA LS	219	454	<b>1.034 (1.005-1.064)</b>	0.023	<b>0.046</b>
Entire cohort	Multiple CRC or AA LS	CRC-free and AA-free LS	272	695	1.042 (1.010-1.080)	0.015	0.060
<i>MLH1</i>	Multiple CRC or AA LS	CRC-free and AA-free LS	112	224	1.030 (0.974-1.090)	0.303	0.303
<i>MSH2/EPCAM</i>	Multiple CRC or AA LS	CRC-free and AA-free LS	123	222	1.032 (0.975-1.090)	0.277	0.303
Age <60y	Multiple CRC or AA LS	CRC-free and AA-free LS	219	695	<b>1.043 (1.008-1.079)</b>	0.016	<b>0.032</b>

282 **TABLE 3. Meta-analysis results of the association between polygenic risk score (rwPRS) and multiple colorectal cancer**  
283 **(CRC) risk or multiple CRC or advanced adenoma (AA) risk.**

284 *Multiple CRC LS*: LS individuals previously diagnosed with more than one CRC; *Single CRC LS*: LS individuals previously diagnosed  
285 with one CRC; *CRC-free LS*: LS individuals without previous diagnosis of CRC; *CRC-free population*: individuals from the population  
286 without previous diagnosis of CRC; *Multiple CRC or AA LS*: LS individuals previously diagnosed of multiple CRCs, multiple AAs or at  
287 least one CRC and one AA; *Single CRC or AA LS*: LS individuals previously diagnosed of CRC or AA, whichever occurred first;  
288 *CRC-free and AA-free LS*: LS individuals without previous diagnosis of CRC or AA; *OR*: Odds ratio; *FDR*: False discovery rate; *Age*  
289 *<60y*: cases with CRC <60 years of age.

## 290 **DISCUSSION**

291 PRS is regarded as an important addition to the assessment of an individual's genetic  
292 risk in patients with sporadic and hereditary cancers; it can be used to identify  
293 individuals with a CRC risk several times lower or higher than that of the average  
294 population. In this way, its implementation seems to be a promising approach for a more  
295 individualised risk stratification. Several studies described the impact of PRS on  
296 modelling CRC risk in the general population<sup>10,12-16</sup>. In line with this, the risk alleles of  
297 those SNPs were found to accumulate in unexplained familial and early-onset CRC  
298 cases<sup>17,18</sup>. However, the interplay between a PRS based on sporadic CRC-associated  
299 SNPs and LS CRC risk remains controversial.

300 It is well known that among patients with hereditary CRC, in particular Lynch syndrome,  
301 the age of onset and cumulative CRC incidence is very heterogeneous, even within the  
302 same family transmitting the same pathogenic germline variant<sup>6</sup>. The estimated gene-  
303 specific, individual lifetime CRC risks of LS patients with *MLH1* or *MSH2* variants can be  
304 lower than 10% or as high as 90%-100% in a considerable fraction, highlighting relevant  
305 genetic and non-genetic modifiers of CRC risk<sup>7,8</sup>. Initially, a small subset of common  
306 CRC-associated SNPs was analysed in selected LS cohorts<sup>29-32</sup>. More recently, some  
307 studies used a more comprehensive set of around 100 CRC-associated SNPs in large  
308 population-based or familial CRC cohorts with conflicting results<sup>16,19,20</sup>.

309 Herein, we used a large, combined cohort of 1,465 affected and unaffected LS patients  
310 with pathogenic MMR germline variants, recruited at two European centres based on  
311 the fulfilment of clinical criteria (revised Bethesda or Amsterdam criteria) or as a result  
312 of an EC or CRC-based LS screening programme, to evaluate to what extent the



313 polygenic background modulates CRC risk. When we compared LS carriers with CRC  
314 against population-based CRC-free controls (mean age 71 years), we did not observe  
315 any significant effect of PRS on CRC risk, neither in the entire cohort nor in subgroups  
316 (gene-specific groups, early-onset group). Nevertheless, the PRS was associated with a  
317 modestly increased risk of CRC or AA in the entire LS cohort. These results are in line  
318 with the work by Jenkins *et al.*, which is based on a similar study design, recruitment  
319 strategy, and a set of 107 SNPs used for PRS calculation<sup>20</sup>. In that work, 826 European-  
320 descent LS carriers from the Colon Cancer Family Registry (CCFR) were included and  
321 the authors found no statistical evidence of an association between PRS and CRC risk,  
322 irrespective of sex or mutated gene.

323 Regarding the analysis between CRC and CRC-free LS probands, we did not find a  
324 statistically significant association between CRC and PRS in the entire cohort or the  
325 different subgroups except for early-onset LS CRC cases (<50 years) and LS with  
326 multiple CRCs or AAs (< 60 years), where a slightly increased CRC risk was evidenced.  
327 In contrast, two recent studies using UKBB data and the same 95 SNPs for PRS  
328 calculation, demonstrated that the polygenic background substantially influences CRC  
329 risk in LS with ORs ranging from 8 to 118 (estimated effect of PRS), or 4 to 16  
330 (calculated effect of PRS) compared with the median tertile of the CRC-free  
331 population<sup>16,19</sup>. According to these results, PRS would account for parts of the  
332 interindividual variation in CRC risk among LS carriers and might contribute towards a  
333 clinically relevant individualised risk stratification.

334 The most obvious explanation for the apparently discrepant results between family-  
335 based (Jenkins *et al.*,<sup>20</sup> and the present study) and population-based (Fahed *et al.*,<sup>16</sup>

336 and Hassanin *et al.*,<sup>19</sup>) studies is differences in study design and recruitment strategies.  
337 The LS probands from the two familial CRC registry studies were mainly recruited  
338 based on established clinical criteria, in particular early-onset and familial clustering of  
339 CRC and other LS-related tumours. Consistent with this ascertainment approach, the  
340 vast majority of participants carry pathogenic variants in the highly penetrant *MLH1* and  
341 *MSH2* genes (**Table 4**), which are likely to be less influenced by the genetic  
342 background.

343 In contrast, studies using individuals from a population-based repository (UKBB) show a  
344 different distribution of affected MMR genes, with the vast majority of LS individuals  
345 carrying pathogenic variants in the moderate and low penetrance genes *MSH6* or *PMS2*  
346 (**Table 4**). In a gene-specific analysis, Hassanin *et al.*, found that the modifying effect of  
347 the PRS is inversely related to the penetrance of the MMR gene, with the strongest  
348 effect in *MSH6* and *PMS2* carriers<sup>19</sup>, which are clearly underrepresented in the studies  
349 of Jenkins *et al.*,<sup>20</sup> and the present.

350 This is in line with hereditary breast cancer, where PRS has proven most relevant as a  
351 cancer risk modifier in carriers of pathogenic variants in moderate penetrance genes  
352 such as *CHEK2*, *ATM*, or *PALB2* compared with *BRCA1/2*<sup>33,34</sup>. While it can be expected  
353 that PRS may have a major influence in less penetrant CRC risk genes, we have not  
354 been able to show a significant effect, likely due to the small numbers of *MSH6* and  
355 *PMS2* variant carriers present in our family-based cohorts due to the aforementioned  
356 selection bias.

357

Study	Inclusion criteria	<i>MLH1</i>	<i>MSH2 / EPCAM</i>	<i>MSH6</i>	<i>PMS2</i>	Cases	Controls	All
-------	--------------------	-------------	---------------------	-------------	-------------	-------	----------	-----

<i>Fahed</i> <sup>16</sup>	Population-based	<b>19</b> (25%)	<b>6</b> (8%)	<b>43</b> (57%)	<b>8</b> (11%)	11 (14%)	65 (86%)	<b>76</b>
<i>Hassanin</i> <sup>19</sup>	Population-based	<b>11</b> (3%)	<b>13</b> (3%)	<b>135</b> (35%)	<b>229</b> (59%)	27 (7%)	361 (93%)	<b>388</b>
<i>Jenkins</i> <sup>20</sup>	Family registry	<b>293</b> (36%)	<b>336</b> (41%)	<b>126</b> (15%)	<b>71</b> (9%)	504 (61%)	322 (39%)	<b>826</b>
<i>This study</i>	Family registry	<b>557</b> (38%)	<b>517</b> (35%)	<b>299</b> (20%)	<b>92</b> (6%)	712 (49%)	753 (51%)	<b>1465</b>

358 **TABLE 4. Characteristics of inclusion criteria and distribution of affected MMR genes in**  
 359 **the different studies analysing PRS in LS individuals.**

360

361 Another plausible explanation for the differences observed may be the sample size. In  
 362 this study, we included twice as many LS individuals as Jenkins *et al.*<sup>20</sup> and four and 20  
 363 times more than the UKBB analyses<sup>16,19</sup>. Moreover, the distribution and composition of  
 364 cases and controls differ between family-registry and population-based studies (**Table**  
 365 **4**). In Jenkins *et al.*, and our study, the percentages of affected (LS carriers with CRC)  
 366 and unaffected (CRC-free LS carriers) individuals are similar and some are members of  
 367 the same family. Hence, the CRC-free LS controls are relatives of the cases and, thus,  
 368 they likely share parts of the polygenic background and other risk factors with their  
 369 affected relatives (cases) to a certain extent, which may explain the observed missing  
 370 effect of PRS. In contrast, the UKBB studies include ten times more controls,  
 371 supposedly healthy LS carriers apparently unrelated to the CRC LS cases. In this  
 372 regard, it was shown that both in sporadic CRC<sup>35</sup> and LS CRC<sup>19</sup>, family history and PRS  
 373 are largely independent and provide complementary information about CRC risk.

374 On the other hand, there are differences in the results obtained between Jenkins *et*  
 375 *al.*,<sup>20</sup> and the present study. These differences can be explained by the sample size, as  
 376 discussed above, and methodological differences. LS individuals in Jenkins *et al.*, were  
 377 censored after a polypectomy, while we considered the first polypectomy as a time-

378 dependent variable of CRC risk, as per studies showing a reduction in CRC incidence in  
379 LS individuals undergoing regular colonoscopies<sup>36,37</sup>. However, since alternative  
380 pathways of colorectal carcinogenesis seem to exist in LS carriers, which originate  
381 doubts regarding the risk-reducing impact of colonoscopies, especially in *MLH1* carriers,  
382 future evidence will determine whether it is useful to apply this time-dependent variable  
383 correction in LS individuals<sup>38,39</sup>. Taken together, this and previous PRS studies on LS  
384 demonstrate that the study design and recruitment strategy strongly influence the  
385 results and conclusions of the PRS.

386 Finally, when we analysed extreme phenotypes, such as early-onset CRC (<50y) and  
387 young (<60y) LS cases with multiple CRCs or AAs, a significant, albeit low, association  
388 between PRS and risk was observed, pointing to a possible contribution of PRS to  
389 these higher-risk situations.

390 Some authors questioned whether the same CRC-associated SNPs identified in the  
391 general population and their specific effect sizes can be applied to stratify CRC risk in  
392 LS individuals and whether both specific SNPs and their risk-modifying power may differ  
393 for each mutated gene<sup>8,20,29</sup>. The potential identification of LS risk-modifier SNPs in  
394 large GWASs might contribute to the description of more specific risk-modulating factors  
395 in the future.

396 PRS studies in much larger, international, multicentric, LS cohorts are needed to more  
397 precisely estimate the PRS effect size in LS individuals, especially in those with extreme  
398 phenotypes, to evaluate the relevance of the polygenic background and interplay with  
399 other genetic and non-genetic risk factors. This will enable its eventual application in  
400 routine clinical practice.

401 In summary, this work shows, for the first time in a family-registry LS cohort, that the  
402 PRS can influence the CRC risk in specific subgroups of LS individuals, albeit with very  
403 weak effect sizes, which contrasts with the clearer modulating effect of the PRS in LS  
404 carriers identified in population-based cohorts.

405 **ACKNOWLEDGEMENTS:** We thank the participating patients and families and all members of  
406 the Units of Genetic Counseling and Genetic Diagnostic of the Hereditary Cancer Program of  
407 the Catalan Institute of Oncology (ICO-IDIBELL) and the Institute of Human Genetics of the  
408 University Hospital Bonn as well as the BufaLynch association for their support and funding of  
409 ICO's Lynch Syndrome Database. We thank Gemma Aiza for technical support. The authors  
410 would also like to acknowledge the Department of Medicine at the Universitat Autònoma de  
411 Barcelona and the CERCA Program/Generalitat de Catalunya for institutional support.

412 **COMPETING INTERESTS:** The authors declare no conflicts of interest.

413 **FUNDING:** This research has been funded by the Instituto de Salud Carlos III and co-funded by  
414 the European Social Fund—ESF investing in your future (grant CM19/00099), the Catalan-  
415 Balearic Society of Oncology (2018 grant of the Catalan-Balearic Society of Oncology), the  
416 European Union's Horizon 2020 research and innovation program under the EJP RD COFUND-  
417 EJP n° 825575, the Spanish Ministry of Economy and Competitiveness and the Spanish  
418 Ministry of Science and Innovation, co-funded by FEDER funds -a way to build Europe-  
419 (grants SAF2015-68016-R and PID2019-111254RB-I00), CIBERONC (CB16/12/00234) and the  
420 Government of Catalonia (SGR\_01112). ADV and VM are part of group 55 of CIBERESP. The  
421 GSA genotyping was performed at the Spanish National Cancer Research Centre, in the  
422 Human Genotyping lab, a member of CeGen, PRB3, and is supported by grant PT17/0019, of  
423 the PE I+D+i 2013-2016, funded by ISCIII and ERDF. ADV was supported by PERIS contract  
424 SLT017/20/000042. This research was partially funded by public grants from the Spanish  
425 Association Against Cancer (AECC) Scientific Foundation—grant GCTRA18022MORE—and  
426 Spanish Ministry for Economy and Competitivity, Instituto de Salud Carlos III, co-funded by  
427 FEDER funds –a way to build Europe (FIS PI14-00613).

428 **AUTHOR CONTRIBUTIONS:**

429 Núria Dueñas – Study concept and design, Analysis and interpretation of data, Drafting of the  
430 manuscript, Critical revision of the manuscript for important intellectual content

431 Hannah Klinkhammer, Nuria Bonifaci, Anna Diez-Villanueva – Study concept and design,  
432 Statistical analysis, Analysis and interpretation of data, Drafting of the manuscript, Critical  
433 revision of the manuscript for important intellectual content  
434 Andreas Mayr, Emadeldin Hassanin, Carlo Maj – Study concept and design, Statistical analysis,  
435 Analysis and interpretation of data, Critical revision of the manuscript for important intellectual  
436 content  
437 Anna Díez-Villanueva, Victor Moreno – Study concept and design, Critical revision of the  
438 manuscript for important intellectual content  
439 Isabel Spier, Marta Pineda, Gabriel Capellá, Stefan Aretz, Joan Brunet – Study concept and  
440 design, Analysis and interpretation of data, Critical revision of the manuscript for important  
441 intellectual content, Study supervision  
442  
443  
444  
445

## 446 REFERENCES

- 447 1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN  
448 Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.  
449 *CA Cancer J Clin.* 2021;71(3):209-249. doi:10.3322/caac.21660
- 450 2. Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of  
451 cancer among 9.6 million individuals in the Swedish Family-Cancer Database. *Int*  
452 *J cancer.* 2002;99(2):260-266. doi:10.1002/ijc.10332
- 453 3. Moreira L, Balaguer F, Lindor N, et al. Identification of Lynch syndrome among  
454 patients with colorectal cancer. *JAMA - J Am Med Assoc.* 2012;308(15):1555-  
455 1565. doi:10.1001/jama.2012.13088
- 456 4. Win AK, Jenkins MA, Dowty JG, et al. Prevalence and penetrance of major genes  
457 and polygenes for colorectal cancer. *Cancer Epidemiol Biomarkers Prev.*  
458 2017;26(3):404-412. doi:10.1158/1055-9965.EPI-16-0693
- 459 5. Hampel H, de la Chapelle A. How do we approach the goal of identifying  
460 everybody with Lynch syndrome? *Fam Cancer.* 2013;12(2):313-317.  
461 doi:10.1007/s10689-013-9611-5
- 462 6. Dominguez-Valentin M, Sampson JR, Seppälä TT, et al. Cancer risks by gene,  
463 age, and gender in 6350 carriers of pathogenic mismatch repair variants: findings  
464 from the Prospective Lynch Syndrome Database. *Genet Med.* 2020;22(1):15-25.  
465 doi:10.1038/s41436-019-0596-9
- 466 7. Win AK, Dowty JG, Reece JC, et al. Variation in the risk of colorectal cancer in  
467 families with Lynch syndrome: a retrospective cohort study. *Lancet Oncol.*  
468 2021;22(7):1014-1022. doi:10.1016/S1470-2045(21)00189-3
- 469 8. Scott RJ. Modifier genes and Lynch syndrome: some considerations. *Hered*  
470 *Cancer Clin Pract.* 2022;20(1):35. doi:10.1186/s13053-022-00240-2
- 471 9. Law PJ, Timofeeva M, Fernandez-Rozadilla C, et al. Association analyses identify  
472 31 new risk loci for colorectal cancer susceptibility. *Nat Commun.*  
473 2019;10(1):2154. doi:10.1038/s41467-019-09775-w
- 474 10. Huyghe JR, Bien SA, Harrison TA, et al. Discovery of common and rare genetic  
475 risk variants for colorectal cancer. *Nat Genet.* 2019;51(1):76-87.  
476 doi:10.1038/s41588-018-0286-6
- 477 11. Fernandez-Rozadilla C, Timofeeva M, Chen Z, et al. Deciphering colorectal  
478 cancer genetics through multi-omic analysis of 100,204 cases and 154,587  
479 controls of European and east Asian ancestries. *Nat Genet.* Published online  
480 December 20, 2022. doi:10.1038/s41588-022-01222-9
- 481 12. Hsu L, Jeon J, Brenner H, et al. A model to determine colorectal cancer risk using  
482 common genetic susceptibility loci. *Gastroenterology.* 2015;148(7):1330-  
483 1339.e14. doi:10.1053/j.gastro.2015.02.010
- 484 13. Jenkins MA, Makalic E, Dowty JG, et al. Quantifying the utility of single nucleotide  
485 polymorphisms to guide colorectal cancer screening. *Futur Oncol.*  
486 2016;12(4):503-513. doi:10.2217/fon.15.303
- 487 14. Frampton MJE, Law P, Litchfield K, et al. Implications of polygenic risk for  
488 personalised colorectal cancer screening. *Ann Oncol.* 2016;27(3):429-434.  
489 doi:10.1093/annonc/mdv540
- 490 15. Thomas M, Sakoda LC, Hoffmeister M, et al. Genome-wide Modeling of Polygenic



- 491 Risk Score in Colorectal Cancer Risk. *Am J Hum Genet.* 2020;107(3):432-444.  
492 doi:10.1016/j.ajhg.2020.07.006
- 493 16. Fahed AC, Wang M, Homburger JR, et al. Polygenic background modifies  
494 penetrance of monogenic variants for tier 1 genomic conditions. *Nat Commun.*  
495 2020;11(1). doi:10.1038/s41467-020-17374-3
- 496 17. Archambault AN, Su Y-R, Jeon J, et al. Cumulative Burden of Colorectal Cancer-  
497 Associated Genetic Variants Is More Strongly Associated With Early-Onset vs  
498 Late-Onset Cancer. *Gastroenterology.* 2020;158(5):1274-1286.e12.  
499 doi:10.1053/j.gastro.2019.12.012
- 500 18. Mur P, Bonifaci N, Díez-Villanueva A, et al. Non-lynch familial and early-onset  
501 colorectal cancer explained by accumulation of low-risk genetic variants. *Cancers*  
502 *(Basel).* 2021;13(15). doi:10.3390/cancers13153857
- 503 19. Hassanin E, Spier I, Bobbili DR, et al. Clinically relevant combined effect of  
504 polygenic background, rare pathogenic germline variants, and family history on  
505 colorectal cancer incidence. *BMC Med Genomics.* 2023;16(1):42.  
506 doi:10.1186/s12920-023-01469-z
- 507 20. Jenkins MA, Buchanan DD, Lai J, et al. Assessment of a Polygenic Risk Score for  
508 Colorectal Cancer to Predict Risk of Lynch Syndrome Colorectal Cancer. *JNCI*  
509 *cancer Spectr.* 2021;5(2):pkab022. doi:10.1093/jncics/pkab022
- 510 21. Dueñas N, Navarro M, Sanjuán X, et al. Lessons learnt from the implementation  
511 of a colorectal cancer screening programme for lynch syndrome in a tertiary  
512 public hospital. *Cancer Epidemiol.* 2023;82. doi:10.1016/j.canep.2022.102291
- 513 22. Obón-Santacana M, Díez-Villanueva A, Alonso MH, et al. Polygenic risk score  
514 across distinct colorectal cancer screening outcomes: from premalignant polyps to  
515 colorectal cancer. *BMC Med.* 2021;19(1). doi:10.1186/s12916-021-02134-x
- 516 23. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent  
517 atherosclerotic disease and established and novel risk factors for predicting  
518 myocardial infarction and cardiac death in healthy middle-aged subjects:  
519 Rationale and design of the Heinz Nixdorf RECALL study. *Am Heart J.*  
520 2002;144(2):212-218. doi:10.1067/mhj.2002.123579
- 521 24. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service  
522 and methods. *Nat Genet.* 2016;48(10):1284-1287. doi:10.1038/ng.3656
- 523 25. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic  
524 variation. *Nature.* 2015;526(7571):68-74. doi:10.1038/nature15393
- 525 26. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome  
526 association and population-based linkage analyses. *Am J Hum Genet.*  
527 2007;81(3):559-575. doi:10.1086/519795
- 528 27. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: A  
529 practical tutorial. *Evid Based Ment Health.* 2019;22(4):153-160.  
530 doi:10.1136/ebmental-2019-300117
- 531 28. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and  
532 Powerful Approach to Multiple Testing. *J R Stat Soc Ser B.* 1995;57(1):289-300.  
533 doi:https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
- 534 29. Win AK, Hopper JL, Buchanan DD, et al. Are the common genetic variants  
535 associated with colorectal cancer risk for DNA mismatch repair gene mutation  
536 carriers? *Eur J Cancer.* 2013;49(7):1578-1587. doi:10.1016/j.ejca.2013.01.029

- 537 30. ten Broeke SW, Elsayed FA, Pagan L, et al. SNP association study in PMS2-  
538 associated Lynch syndrome. *Fam Cancer*. 2018;17(4):507-515.  
539 doi:10.1007/s10689-017-0061-3
- 540 31. Talseth-Palmer BA, Wijnen JT, Brenne IS, et al. Combined analysis of three lynch  
541 syndrome cohorts confirms the modifying effects of 8q23.3 and 11q23.1 in MLH1  
542 mutation carriers. *Int J Cancer*. 2013;132(7):1556-1564. doi:10.1002/ijc.27843
- 543 32. Houlle S, Charbonnier F, Houivet E, et al. Evaluation of Lynch syndrome modifier  
544 genes in 748 MMR mutation carriers. *Eur J Hum Genet*. 2011;19(8):887-892.  
545 doi:10.1038/ejhg.2011.44
- 546 33. Gallagher S, Hughes E, Wagner S, et al. Association of a Polygenic Risk Score  
547 with Breast Cancer among Women Carriers of High- And Moderate-Risk Breast  
548 Cancer Genes. *JAMA Netw Open*. 2020;3(7).  
549 doi:10.1001/jamanetworkopen.2020.8501
- 550 34. Hassanin E, May P, Aldisi R, et al. Breast and prostate cancer risk: The interplay  
551 of polygenic risk, rare pathogenic germline variants, and family history. *Genet  
552 Med*. 2022;24(3):576-585. doi:10.1016/j.gim.2021.11.009
- 553 35. Mars N, Lindbohm J V., della Briotta Parolo P, et al. Systematic comparison of  
554 family history and polygenic risk across 24 common diseases. *Am J Hum Genet*.  
555 Published online December 2022. doi:10.1016/j.ajhg.2022.10.009
- 556 36. Järvinen HJ, Aarnio M, Mustonen H, et al. Controlled 15-year trial on screening  
557 for colorectal cancer in families with hereditary nonpolyposis colorectal cancer.  
558 *Gastroenterology*. 2000;118(5):829-834. doi:10.1016/S0016-5085(00)70168-5
- 559 37. Kalady MF, Lipman J, McGannon E, Church JM. Risk of colonic Neoplasia after  
560 proctectomy for rectal cancer in hereditary Nonpolyposis Colorectal cancer. *Ann  
561 Surg*. 2012;255(6):1121-1125. doi:10.1097/SLA.0b013e3182565c0b
- 562 38. Ahadova A, Gallon R, Gebert J, et al. Three molecular pathways model colorectal  
563 carcinogenesis in Lynch syndrome. *Int J Cancer*. 2018;143(1):139-150.  
564 doi:10.1002/ijc.31300
- 565 39. Seppälä T, Pylvänäinen K, Evans DG, et al. Colorectal cancer incidence in  
566 path\_MLH1 carriers subjected to different follow-up protocols: a Prospective  
567 Lynch Syndrome Database report. *Hered Cancer Clin Pract*. 2017;15:18.  
568 doi:10.1186/s13053-017-0078-5  
569

570 **FIGURE LEGENDS**

571 **FIGURE 1. Interplay of each mutated germline gene and polygenic risk score**  
572 **(*rwPRS*) for colorectal cancer (CRC) risk.** CRC risk for each mutated gene, stratified  
573 according to *rwPRS* tertiles using the intermediate *rwPRS* tertile as the reference group.  
574 95% confidence intervals are indicated by vertical lines. PMS2 carriers are not included  
575 due to the low sample size.

576

577 **FIGURE 2. Interplay of each mutated germline gene and polygenic risk score**  
578 **(*rwPRS*) for colorectal cancer (CRC) or advanced adenoma (AA) risk.** Risk of CRC  
579 or AA for each mutated gene, stratified according to *rwPRS* tertiles using the  
580 intermediate *rwPRS* tertile as the reference group. 95% confidence intervals are  
581 indicated by vertical lines. PMS2 carriers are not included due to the low sample size.



