**Short Report**

**Effects of common genetic variants associated with colorectal cancer risk on survival outcomes after diagnosis: A large population-based cohort study**

Yazhou He, Evropi Theodoratou, Xue Li, Farhat V.N. Din, Peter Vaughan-Shaw, Victoria Sviinti, Susan M. Farrington, Harry Campbell, Malcolm G. Dunlop and Maria Timofeeva

1 Cancer Research UK Edinburgh Centre, Medical Research Council Institute of Genetics & Molecular Medicine, Western General Hospital, The University of Edinburgh, Edinburgh, United Kingdom
2 Colon Cancer Genetics Group, Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics & Molecular Medicine, Western General Hospital, The University of Edinburgh, Edinburgh, United Kingdom
3 Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, United Kingdom

**Abstract**

Genome-wide association studies have thus far identified 130 genetic variants linked to colorectal cancer (CRC) risk ($r^2 < 0.2$). Given their implication in disease causation, and thus plausible biologically effects on cancer-relevant biological pathways, we investigated whether these variants are associated with CRC prognosis and also whether they might provide predictive value for survival outcome. We conducted the analysis in a well-characterized population-based study of 5,675 patients after CRC diagnosis in Scotland. None of the genetic risk variants were associated with either overall survival (OS) or CRC-specific survival. Next, we combined the variants in a polygenic risk score, but again we observed no association between survival outcome and overall genetic susceptibility to CRC risk—as defined by common genetic variants (OS: hazard ratio = 1.00, 95% confidence interval = 0.96–1.05). Furthermore, we found no incremental increase in the discriminative performance when adding these genetic variants to the baseline CRC-survival predictive model of age, sex and stage at diagnosis. Given that our study is well-powered (>0.88) to detect effects on survival for 74% of the variants, we conclude that effects of common variants associated with CRC risk which have been identified to date are unlikely to have clinically relevant effect on survival outcomes for patients diagnosed with CRC.

**Introduction**

Globally, colorectal cancer (CRC) is the second leading cause of cancer-related deaths, accounting for 9.2% of all cancer-related deaths (0.8 million CRC deaths in 2018). The strongest known predictor of CRC outcome is stage, but even within one stage, there is considerable heterogeneity in survival. Identification of biomarkers of cancer prognosis can inform clinical management and treatment of disease. Evidence of the family concordance for CRC-

**Additional Supporting Information** may be found in the online version of this article.

**Key words:** common genetic variants, colorectal cancer, survival, cohort study

**Abbreviations:** AJCC: American Joint Committee on Cancer; CI: confidence interval; CRC: colorectal cancer; FDR: false positive rates; HR: hazard ratio; LASSO: least absolute shrinkage and selection operator.; MAF: minor allele frequency; PRS: polygenic risk score; SCR: Scottish Cancer Registry; SOCCS: Study of Colorectal Cancer in Scotland

**Conflict of Interest:** All authors declare no conflicts of interest.

**Grant sponsor:** MRC Clinical Research Training Fellowship; Grant number: MR/M004007/1; Grant sponsor: CRUK Career Development Fellowship; Grant number: C31250/A22804; Grant sponsor: MRC Human Genetics Unit Centre; Grant numbers: U127527198, U127527202; Grant sponsor: Edinburgh CRUK Cancer Research Centre; Grant sponsor: CRUK Program; Grant number: C348/A18927

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

**DOI:** 10.1002/ijc.32550

**History:** Received 3 May 2019; Accepted 17 Jun 2019; Online 4 Jul 2019

**Correspondence to:** Malcolm G. Dunlop, Institute of Genetics and Molecular Medicine, The University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom, Tel./Fax: +44-131-651-8631, E-mail: malcolm.dunlop@igmm.ed.ac.uk; or Maria Timofeeva, Institute of Genetics and Molecular Medicine, The University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom, Tel./Fax: +44-131-651-8687, E-mail: maria.timofeeva@igmm.ed.ac.uk

**Int. J. Cancer: 00, 00–00 (2019) © 2019 The Authors. International Journal of Cancer published by John Wiley & Sons Ltd on behalf of UICC**
pecific survival, together with some suggestions of improved survival in cancer patients with a family history compared to patients without a family history, indicates that genetic signature can affect prognosis of CRC patients after diagnosis. Indeed, improved survival for Lynch syndrome patients with germline rare variations in DNA mismatch repair genes is well documented, suggesting that genetic variants associated with CRC pathogenesis may subsequently affect tumor progression. However, very few studies with sufficient power tested roles of common genetic risk variants in CRC prognosis. Previously published smaller studies examined up to 30 CRC risk genetic loci and detected no or little evidence of associations with survival. Two recent large meta-analyses of genome-wide association studies (GWAS) have identified more than 70 new genetic variants associated with CRC risk. In this analysis, we investigated the association between all previously and newly GWAS-identified common genetic variants and CRC survival.

**Materials and Methods**

We included 5,675 CRC cases (detailed patient selection in Supporting Information Fig. S1) with genome-wide genotyping data and data on age at diagnosis, sex and American Joint Committee on Cancer (AJCC) stage information from a population-based case-control study (Study of Colorectal Cancer in Scotland, SOC CCS; 1999-current). Ethics approval was obtained from the MultiCentre Research Ethics committee for Scotland (approval number MREC/01/0/5) and other committees (presented elsewhere). A total of 130 genetic variants identified by previous GWAS studies and were genotyped or imputed (25/130 variants were directly genotyped). For correlated variants (linkage disequilibrium correlation), whichever came ascertainment from the Scottish Cancer Registry (SCR), and patients without a family history, indicates that genetic signature can affect prognosis of CRC patients after diagnosis. Indeed, improved survival for Lynch syndrome patients with germline rare variations in DNA mismatch repair genes is well documented, suggesting that genetic variants associated with CRC pathogenesis may subsequently affect tumor progression. However, very few studies with sufficient power tested roles of common genetic risk variants in CRC prognosis. Previously published smaller studies examined up to 30 CRC risk genetic loci and detected no or little evidence of associations with survival. Two recent large meta-analyses of genome-wide association studies (GWAS) have identified more than 70 new genetic variants associated with CRC risk. In this analysis, we investigated the association between all previously and newly GWAS-identified common genetic variants and CRC survival.

**What’s new?**

Genetic variants associated with colorectal cancer (CRC) risk potentially also exert effects on disease prognosis, though little is known about the heritability of CRC survival. In this population-based study, genetic variants of CRC susceptibility were investigated for their relevance in CRC survival based on analyses of variants identified in genome-wide association studies (GWASs). Nonsignificant associations were detected between small numbers of genetic variants and overall survival and CRC-specific survival. Overall, the findings do not support the existence of prognostic effects of common CRC risk variants. Rather, CRC survival may have distinct genetic determinants, warranting separate investigation by GWAS.
Data availability
The data that support the findings of our study are available upon reasonable request from the corresponding authors. The data are not publicly available due to privacy or ethical restrictions.

Results and Discussion
The basic characteristics of included CRC patients are summarized in Table 1. In total, 1,918 patients (34%) died during follow-up. With 5,675 CRC cases, our study had 88% power to detect a hazard ratio of 1.20 on OS for 97/130 (74%) variants (MAF > 0.15) at the significance level of 0.0005 (a power curve for other effects is shown in Fig. 1). Overall, we observed 14 genetic variants associated with OS and 10 with CSS at nominal statistical significance \((p < 0.05)\) with six variants \((\text{rs10994860}, \text{rs12143541}, \text{rs3217810}, \text{rs34405347}, \text{rs6065668}, \text{rs847208})\) being associated with both OS and CSS. However, none of the variants remained statistically significant after Bonferroni or FDR correction. The summary results for variants with nominal significance are presented in Table 2. Stratified analyses by stage, sex and tumor site did not identify any statistically significant associations after multiple-testing correction either (Supporting Information Tables S2–S4). With regard to overall genetic susceptibility to CRC, no statistically significant association was observed between the PRS and OS or CSS (Table 2). The LASSO regression model selected six variants for OS in addition to age at diagnosis, sex and AJCC stage to minimize prediction error in the train set. However, nearly no incremental predictive improvement was observed compared to the model without genetic variants in the test set (C-statistic for OS: 0.73282 vs. 0.73277, U-statistic test: \(p = 0.322\)). Similar results were found for CSS (Supporting information Table S5).

We used a concept of statistical significance as decision criteria to define if the risk variants have an effect on survival. This concept has been criticized in the literature as subjective and commonly misused.\(^{24,25}\) We additionally looked into
direction of effects to overcome limitations of statistical significance. We hypothesized that in case of no effects, the CRC risk variants will cause improvement or impairment of OS and CSS in equal proportion. Ambiguous AT and CG genetic variants were excluded and out of 124 tested variants only 52 risk variants were associated with worse OS (proportion of risk variants associated with worse OS 41%, 95% CI = 33–51%) and 58 were associated with decrease in CSS (proportion of risk variants associated with worse CSS 47%, 95% CI = 38–56%). Though not reaching suggested significance level (p ≤ 0.05), these results are consistent with directions of effects observed in previous studies.3,4

None of the nominally significant genetic variants have known detrimental clinically relevant effects on gene function (Supporting Information Table S6). Rs3087967, which is located 3’UTP of C11orf53, is known to be associated with higher COLCA2 and C11orf53 expression in colon transverse tissue for C allele.5,20 However, little is known about COLCA2 and C11orf53 functions. Rs4759277, which is associated with worse OS and located within intron region of LRPI, is likely to affect binding of transcription factors and associated with LRPI gene expression in tibial artery and sun-exposed skin.20 Another variant with nominally significant effects on both OS and CSS is rs10994860 variant. It is located 5’UTP of APOBEC1 complementation factor (AICF) and likely to affect binding (regulomeDB score 2b). The same variant has been previously associated with estimated glomerular filtration rate (eGFR), a measure of the kidney’s filtration ability in serum.26

This is the first study capturing all currently identified CRC risk genetic loci (n = 92) and investigating their associations with survival outcomes in a population-based study. Our results indicate that overall genetic CRC susceptibility measured by GWAS-identified variants is not statistically significantly associated with survival after CRC diagnosis. With regard to each variant, our study, which had acceptable power for HR > 1.2, found multiple variants (14 for OS and 10 for CSS) associated with survival outcomes at p < 0.05, although the significance fails to survive correction for multiple testing. Similarly, previous studies identified some CRC-risk variants27,28 that might be associated with survival after CRC diagnosis, these findings were not immune to false-positive results from multiple testing. A widely studied variant, rs9929218 lies in the intron of CDH1 gene encoding E-cadherin, the loss of function of which can cause tumor progression and metastasis.29 Previous studies reported that the CRC-risk decreasing allele (A) is statistically significantly associated with poor survival outcomes.27,28 However, in contrast, we observed a potentially favorable effect (though not statistically significant after multiple-testing correction) of the A allele on OS in our study (i.e., the direction of CRC risk and prognosis are consistent in our study). Smith et al. reported that the A allele is significantly associated with poor response to chemotheraphy, implying a

<table>
<thead>
<tr>
<th>Genetic marker</th>
<th>Gene</th>
<th>EA</th>
<th>MAF</th>
<th>HR* (95% CI)</th>
<th>p value</th>
<th>Pfdr</th>
<th>HR* (95% CI)</th>
<th>p value</th>
<th>Pfdr</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10161980</td>
<td>AL139383.1</td>
<td>G</td>
<td>0.40</td>
<td>1.08 (1.01–1.15)</td>
<td>0.019</td>
<td>0.251</td>
<td>1.06 (0.98–1.14)</td>
<td>0.160</td>
<td>0.832</td>
</tr>
<tr>
<td>rs10994860</td>
<td>A1CF</td>
<td>T</td>
<td>0.17</td>
<td>1.11 (1.02–1.20)</td>
<td>0.016</td>
<td>0.251</td>
<td>1.13 (1.02–1.24)</td>
<td>0.018</td>
<td>0.603</td>
</tr>
<tr>
<td>rs11196171</td>
<td>TCF7L2</td>
<td>G</td>
<td>0.22</td>
<td>0.91 (0.83–0.98)</td>
<td>0.017</td>
<td>0.251</td>
<td>0.93 (0.85–1.03)</td>
<td>0.161</td>
<td>0.832</td>
</tr>
<tr>
<td>rs12143541</td>
<td>TTC22</td>
<td>G</td>
<td>0.15</td>
<td>1.13 (1.04–1.24)</td>
<td>0.006</td>
<td>0.251</td>
<td>1.13 (1.02–1.25)</td>
<td>0.023</td>
<td>0.603</td>
</tr>
<tr>
<td>rs16959063</td>
<td>FMN1</td>
<td>A</td>
<td>0.01</td>
<td>0.72 (0.53–0.97)</td>
<td>0.034</td>
<td>0.380</td>
<td>0.83 (0.59–1.16)</td>
<td>0.273</td>
<td>0.832</td>
</tr>
<tr>
<td>rs174537</td>
<td>MYRF</td>
<td>T</td>
<td>0.34</td>
<td>1.08 (1.01–1.16)</td>
<td>0.019</td>
<td>0.251</td>
<td>1.05 (0.97–1.14)</td>
<td>0.218</td>
<td>0.832</td>
</tr>
<tr>
<td>rs2696839</td>
<td>Intergenic</td>
<td>C</td>
<td>0.48</td>
<td>0.95 (0.90–1.02)</td>
<td>0.143</td>
<td>0.642</td>
<td>0.93 (0.86–1.00)</td>
<td>0.048</td>
<td>0.603</td>
</tr>
<tr>
<td>rs3087967</td>
<td>C11orf53</td>
<td>C</td>
<td>0.33</td>
<td>0.92 (0.86–0.99)</td>
<td>0.017</td>
<td>0.251</td>
<td>0.93 (0.86–1.00)</td>
<td>0.062</td>
<td>0.616</td>
</tr>
<tr>
<td>rs3217810</td>
<td>CCND2</td>
<td>T</td>
<td>0.14</td>
<td>1.13 (1.02–1.25)</td>
<td>0.016</td>
<td>0.251</td>
<td>1.13 (1.00–1.27)</td>
<td>0.044</td>
<td>0.603</td>
</tr>
<tr>
<td>rs3217874</td>
<td>CCND2</td>
<td>T</td>
<td>0.42</td>
<td>1.07 (1.00–1.14)</td>
<td>0.050</td>
<td>0.438</td>
<td>1.06 (0.98–1.14)</td>
<td>0.139</td>
<td>0.823</td>
</tr>
<tr>
<td>rs34405347</td>
<td>Intergenic</td>
<td>G</td>
<td>0.08</td>
<td>0.86 (0.74–0.95)</td>
<td>0.007</td>
<td>0.251</td>
<td>0.85 (0.73–0.99)</td>
<td>0.042</td>
<td>0.603</td>
</tr>
<tr>
<td>rs35509282</td>
<td>Intergenic</td>
<td>A</td>
<td>0.12</td>
<td>0.94 (0.86–1.04)</td>
<td>0.256</td>
<td>0.772</td>
<td>0.88 (0.78–0.99)</td>
<td>0.033</td>
<td>0.603</td>
</tr>
<tr>
<td>rs4759277</td>
<td>LRP1</td>
<td>A</td>
<td>0.38</td>
<td>1.07 (1.00–1.14)</td>
<td>0.051</td>
<td>0.438</td>
<td>1.09 (1.01–1.18)</td>
<td>0.028</td>
<td>0.603</td>
</tr>
<tr>
<td>rs6065668</td>
<td>Intergenic</td>
<td>T</td>
<td>0.28</td>
<td>0.90 (0.83–0.96)</td>
<td>0.003</td>
<td>0.251</td>
<td>0.89 (0.82–0.97)</td>
<td>0.011</td>
<td>0.603</td>
</tr>
<tr>
<td>rs7495132</td>
<td>CRT3</td>
<td>T</td>
<td>0.12</td>
<td>1.03 (0.93–1.13)</td>
<td>0.602</td>
<td>0.888</td>
<td>1.13 (1.01–1.26)</td>
<td>0.032</td>
<td>0.603</td>
</tr>
<tr>
<td>rs847208</td>
<td>LINC01081</td>
<td>A</td>
<td>0.37</td>
<td>0.93 (0.88–1.00)</td>
<td>0.042</td>
<td>0.416</td>
<td>0.92 (0.85–0.99)</td>
<td>0.027</td>
<td>0.603</td>
</tr>
<tr>
<td>rs9537521</td>
<td>AL139383.1</td>
<td>A</td>
<td>0.38</td>
<td>1.10 (1.02–1.17)</td>
<td>0.008</td>
<td>0.251</td>
<td>1.06 (0.98–1.15)</td>
<td>0.153</td>
<td>0.832</td>
</tr>
<tr>
<td>rs9929218</td>
<td>CDH1</td>
<td>A</td>
<td>0.28</td>
<td>0.93 (0.86–0.99)</td>
<td>0.035</td>
<td>0.380</td>
<td>0.96 (0.88–1.04)</td>
<td>0.294</td>
<td>0.836</td>
</tr>
<tr>
<td>PRS</td>
<td>NA</td>
<td></td>
<td></td>
<td>1.00 (0.95–1.04)</td>
<td>0.864</td>
<td>NA</td>
<td>1.03 (0.97–1.08)</td>
<td>0.340</td>
<td>NA</td>
</tr>
</tbody>
</table>

1Hazard ratios are derived from Cox regression models adjusted for age at diagnosis, sex and AJCC stage.
2Minor allele frequency is the minor allele prevalence in SOCCS.
Abbreviations: CI, confidence interval; CRC, colorectal cancer; EA, effect alleles; HR, hazard ratio; MAF, minor allele frequency; NA, not available; Pfdr, p values after correction for false positive rates; PRS, polygenic risk score.
possible gene \times therapy interaction for this variant. However, our study is limited by data on response to chemotherapy being unavailable so that this could not be explored further. Notably, there has been other evidence showing that rs9929218 may modify CRC susceptibility by interacting with other factors such as height and alcohol consumption. Investigation of possible gene–environment interactions should be considered in future efforts with large sample sizes to further dissect the prognostic effect of this variant in CRC patients. Consistently, when looking for overall direction of effects among 124 tested variants we noted little evidence of potentially detrimental effects of CRC risk on survival with only 41 and 47% of risk variants showing association with poor (HR > 1) OS and CSS in our study.

Previously we have shown that SOCCS study is representative of British and Scottish populations and cases from SOCCS cluster tightly with population-based controls from SOCCS and Generation Scotland. The allele frequencies of studied genetic variants are in range expected for European populations (Supporting Information Table S1). However, the results may be not generalizable to populations of cancer patients where substantial differences in allele frequencies and/or treatment and disease management are anticipated.

In conclusion, our study finds that overall genetic susceptibility to CRC captured by known CRC risk variants is not statistically significantly associated with survival outcomes of CRC. However, possible roles of each variant in CRC progression remain to be explored. Our study indicated that the heritable variation of patient survival may have distinct genetic determinants from CRC susceptibility. The previous GWAS on CRC survival with 3,494 cases identified no variants at genome-wide significance (p < 5E–8), meriting future collaborative efforts of aggregating larger CRC cohorts to illuminate genetic structure of survival outcomes for CRC patients.

Acknowledgements
We acknowledge the excellent technical support from Stuart Reid and Marion Walker. We are grateful to Donna Markie and all those who continue to contribute to recruitment, data collection, and data curation for the Study of Colorectal Cancer in Scotland studies. We acknowledge that these studies would not be possible without the patients and surgeons who take part. We acknowledge the expert support on sample preparation from the Genetics Core of the Edinburgh Wellcome Trust Clinical Research Facility. We are grateful for the provision of public data from the GTEx consortium. The data used for the analyses described in this article were obtained from the GTEx Portal on 05/27/2019. This work was supported by CRUK Program grant C348/A18927 (MGD). It was also supported by funding for the infrastructure and staffing of the Edinburgh CRUK Cancer Research Centre. This work was also funded by a grant to MGD as Project Leader with the MRC Human Genetics Unit Centre Grant (U172572702 and U1725727198 from 1/4/18). ET is supported by a CRUK Career Development Fellowship (C31250/A22804). PVS was supported by MRC Clinical Research Training Fellowship (MR/M004007/1), a research fellowship from the Harold Bridges Bequest and by the Melville Trust for the Care and Cure of Cancer. The work received support from COST Action BM1206.

References
20. GTEx. [cited 27 May 2019]. Available at: https://www.gtexportal.org/home/.


Join us for 24 Days of Stem Cells; a premiere virtual event featuring the latest advances in stem cell research.

This year's format will feature a new hour of cutting edge content every weekday starting October 1st. Attend the sessions that are most relevant to your work - at your convenience and at your pace.

During the 24-day long event, you can:
- Access leading scientific presentations from thought leaders around the world
- Watch live training demonstrations from our stem cell experts
- Download key stem cell tools and resources
- Complete weekly challenges to earn points towards certification and prizes

Register today at www.24daysofstemcells.com