OXFORD

JNCI J Natl Cancer Inst (2022) 114(5): djab229

https://doi.org/10.1093/jnci/djab229
First published online January 3, 2022
Commentary

Molecular Biologic and Epidemiologic Insights for Preventability of Colorectal Cancer

Edward Giovannucci, MD, ScD*

Departments of Nutrition and Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

*Correspondence to: Edward Giovannucci, ScD, Harvard T.H. Chan School of Public Health, 665 Huntington Ave, Boston, MA 02115, USA (e-mail: egiovann@hsph.harvard.edu).

COMMENTARY

Abstract

The etiology of colorectal cancer (CRC) has been informed from both a molecular biology perspective, which concerns the study of the nature, timing, and consequences of mutations in driver genes, and epidemiology, which focuses on identifying risk factors for cancer. For the most part, these fields have developed independently, and it is thus important to consider them in a more integrated manner. The molecular mutational perspective has stressed the importance of mutations due to replication of adult stem cells, and the molecular fingerprint of most CRCs does not suggest the importance of direct carcinogens. Epidemiology has identified numerous modifiable risk factors that account for most CRCs, most of which are not direct mutagens. The distribution of CRCs across the large bowel is not uniform, which is possibly caused by regional differences in the microbiota. Some risk factors are likely to act through or interact with the microbiota. The mutational perspective informs when risk factors may begin to operate in life and when they may cease to operate. Evidence from the mutational model and epidemiology supports that CRC risk factors begin early in life and may contribute to the risk of early-onset CRC. Later in carcinogenesis, there may be a "point of no return" when sufficient mutations have accumulated, and some risk factors do not affect cancer risk. This period may be at least 5-15 years for some risk factors. A more precise knowledge of timing of risk factor to cancer is required to inform preventive efforts.

From a molecular mutational perspective, cancer is understood primarily as a disease characterized by the progression of mutations in driver genes that provide the affected cells a growth advantage over their neighboring cells. Thus, the molecular biology of mutations concerns the study of the nature, timing, and consequences of mutations in driver genes (1). Colorectal cancer (CRC) is a malignancy for which molecular understanding has advanced greatly. Epidemiology, focusing on identifying risk factors for CRC, has also advanced substantially (2). For the most part, the study of mutations and that of risk factors for CRC have developed independently. Risk factors could possibly increase the likelihood of selection of driver gene mutations but through processes that are not directly mutagenic. It is thus important to consider conclusions from mutational studies and epidemiologic studies in a more integrated manner, as is attempted herein.

The Mutational Perspective

Although cancer cells typically have numerous mutations, Knudson (3) conjectured several decades ago that only a small number such as 2 or 3 driver mutations are sufficient. Multistage carcinogenesis modeling supported the idea that only 2 to 3 driver mutations were necessary to explain the dynamics of CRC incidence (4-7). In recent decades, a preferential sequence of driver mutations has been confirmed, at least for a substantial proportion of CRC, as initially shown by Fearon and Vogelstein (1). Specifically, *APC* mutations tend to occur early, followed by *KRAS* mutations, loss of 17p and *TP53*, and *SMAD4*; whole-genome duplications occur after several driver mutations have accumulated, and chromosomal gains and losses tending to occur late in carcinogenesis (8). Another major but less common pathway is the serrated pathway, typically involving *BRAF* mutations, CpG island methylator phenotype (CIMP) -high, microsatellite-instable (high) tumors.

Focusing on mutational events, a recent mathematical model by Lahouel et al. (3) considers the process of tumor evolution including different types of fitness advantages for driver genes and considerations of the number of cells that can be supported and sustained, given the resource limitations of the local environment ("carrying capacity"). In this model, a

Received: August 9, 2021; **Revised:** September 30, 2021; **Accepted:** December 16, 2021

proliferating mass of cells is considered to be a CRC when all of the required driver mutations have occurred and the cancer has attained a detectable size. By varying only 2 tissue-specific parameters—the number of adult stem cells in the tissue and the cell division frequency—the model recapitulated a substantial proportion of the observed cancer incidence for CRC, including the risk of CRC rising exponentially with increasing age, the frequent presence of still benign (adenomas) neoplastic lesions in the large bowel comparable with the cancer in size, and the predicted earlier ages of onset for inherited conditions (Lynch and familial adenomatous polyposis). In addition, according to the model predictions, the final driver mutation occurs in CRC frequently after the detection size has been reached, which has favorable implications for screening. In contrast to previous models for which subsequent hits take increasingly shorter times, in this model the first driver mutation typically requires a shorter time than subsequent hits.

The Epidemiologic (Risk Factor) Perspective

CRC incidence rates vary dramatically, up to 45-fold, across populations (2). Although genetic factors such as germline *MLH1* and *APC* mutations contribute to CRC incidence, the majority of CRC is sporadic and believed largely attributable to modifiable environmental risk factors. The 45-fold gradient in risk may be somewhat exaggerated because of underascertainment of CRC in some registries, which may make CRC incidence artificially low in some populations. Nonetheless, if limiting to countries with reliable case ascertainment, the variation in incidence is tenfold. Consistently, incidence rates of CRC increase in a population up to tenfold after economic development and "westernization," as in Japan after 1950 (2). These changes can be dramatic in magnitude and occur relatively quickly. The incidence of rectal cancer diagnosed in US adults younger than 55 years doubling in just 2 decades (9) is a recent example.

The CRC risk factors generally accepted as likely causal include obesity, physical inactivity, certain poor dietary patterns, red and processed meat consumption, alcohol drinking, and tobacco use. Protective factors include dietary fiber, dairy or calcium intake, and use of aspirin (2). Other factors, including probiotics (yogurt), are suggestively protective (10). Unlike the case for a dominant risk factor such as tobacco is for lung cancer, for CRC there are multiple factors that may each contribute approximately 5%-15% of the total. Evidence using different methods suggests that, based on the accepted risk factors in composite, at least 60%-65% of CRC is preventable (11,12). This figure may even be an underestimate because many studies are based on a measure of exposure, typically only once in adulthood; if early life or lifelong exposure is relevant, a single adult measure would provide an underestimate. For example, using multiple measures for body mass index indicated that the proportion of CRC attributable to body mass index greater than 22.5 kg/m^2 is approximately 30% (13), which contrasts to much lower estimates of CRC because of excess adiposity, such as 5% (12). Mendelian randomization studies, which may better estimate "lifelong exposure" to excess adiposity, further support that a single measure of adiposity is an underestimate (14). From the epidemiologic perspective, CRC is potentially a highly preventable cancer.

Nature and Cause of Mutations

A recent model, the EHR model, posits 3 classes of mutations in cancer etiology: those due to random mistakes during normal DNA replication (R), those from environment and lifestyle factors (E), and those related to heredity (H) (15,16). The observation that estimated lifetime adult stem cell replications across tissues correlates well with lifetime cancer risk in the tissue indicates that R mutations are important for some cancers. In support of this, cell-intrinsic mutational processes, such as deamination-induced mutagenesis (C: G to T: A transitions at CpG sites) in rapidly cycling stem cells, contribute substantially to the point mutation load for CRC in driver genes (17). This pattern may suggest that most mutations in driver genes in CRC are not a result of a direct carcinogen, which would leave another molecular signature. A contrasting example is that driver mutations in the same genes for liver cancer have a completely different molecular fingerprint (17).

The EHR model does not directly inform on what the relevant risk factors for CRC are, but it does suggest that most of the mutations (R) in driver genes will not be due to directly mutagenic agents. The estimated fraction of cancer-causing mutations attributable to R mutations varies greatly by cancer type—for example, 33% in lung cancer, which has smoking as a predominant risk factor, and 71% for CRC (15). In fact, many of the risk or protective factors for CRC identified by epidemiology (eg, aspirin, adiposity, physical activity, and certain dietary patterns, fiber, calcium) do not have a clear mutagenic effect. Adiposity, physical activity, and diet in part might act through inflammation and hyperinsulinemia (18). Alcohol can produce acetaldehyde, a potential mutagen, though nonmutagenic mechanisms have been also described (2). Tobacco carcinogens can be directly mutagenic but can also impair DNA methylation (19). Recently, a novel alkylating mutational signature, which is associated with red and processed meat consumption and distal tumor location and predicted to target *KRAS*, was identified (20). Nonetheless, the bulk of external risk factors for CRC are not obviously mutagenic, consistent with the molecular perspective of R mutations being dominant for CRC.

Although the EHR model assumes E, H, and R mutations are independent, some exposures may act by increasing or decreasing cell division. Adiposity, physical inactivity, and certain dietary patterns may increase levels of growth factors, such as insulin and insulin-like growth factors, which enhance cell proliferation and inhibit apoptosis (18). Weight loss studies in humans have shown a marked reduction in colorectal tissue expression of Ki-67 by 44% (21) and a 39% reduction in whole-crypt labeling index and a 57% lower in upper crypt labeling (22). In rats, a reduction in colonic crypt proliferation appears to mediate the effect of caloric restriction on lowering incidence of intestinal tumors (23). Aspirin and nonsteroidal anti-inflammatory drugs, which causally reduce risk of CRC, inhibit prostaglandin-endoperoxide synthase-2 (also known as cyclooxygenase-2) overexpression, an enhancer of cell proliferation. Calcium intake, associated with lower risk of CRC, may inhibit proliferation in the large intestine by activating the calcium sensing receptor (24). Calcium intake appears to be associated with lower risk of CRC only when the calcium sensing receptor is expressed (25).

Timing of Mutations in Driver Genes During the Life Course and CRC Risk

Molecular and modeling studies can inform when events in carcinogenesis (particularly driver mutations) are occurring during the life course. According to the Lahouel et al. (3) model discussed above, the first mutation in a driver gene for an eventual CRC tends to occur at approximately age 14 years on average,

and the second and third hits can take progressively longer (eg, 20 and 30 years, respectively). The average age for CRC from this model can be inferred to be around the mid- to late 60s, which is compatible with population cancer rates. This timing of events is also supported by epigenetic studies, which can inform on tissue aging. Combining molecular data of DNA methylation drift with CRC incidence data (26), it was predicted that adenoma that eventually developed into CRC started early in life, most likely before age 20 years (27). Because CRC is typically diagnosed in middle or old age, it is not intuitive for many that the causative factors can act as early as childhood and adolescence, and possibly in utero (28).

Knowledge of the timing during the life course when risk or protective factors are operative for CRC is critical to inform primary prevention. Limited epidemiologic data indicate that the standard CRC risk factors begin operating in childhood, adolescence, and early adulthood, including ionizing radiation, tobacco, obesity, physical inactivity, alcohol, western dietary patterns, excessive sugar, and low calcium or dairy (29-36). Most of these risk factors are observed as adult risk factors for CRC. Yet because behaviors tend to correlate over time, it is critical for studies to confirm whether adult effects are directly causal or merely correlated with early-life causal events. It is not obvious that a risk factor would have a similar effect in the preneoplastic large bowel mucosa (before the initial mutation) and the subsequent neoplasm. For example, some mutations, such as in the *APC* gene, tend to occur early, so any risk factor that preferentially affects mutation of the *APC* gene could possibly be operative only early in life. Interestingly, for obesity, an increased risk is observed at early ages such as in adolescence but seems to extend to later life only in men and not women (21).

The molecular perspective could suggest that after a certain mutational load, the neoplasm may become insensitive to certain exposures. In epidemiologic analyses, time lag analysis could demonstrate a time period after which an exposure is no longer active; for example, for a period of risk in the year 2000, a 10-year lag would suggest that exposure before 1990 affects risk but after 1990 does not influence risk of CRC. For CRC, analyses suggest the following time lags: folate, at least 12-16 years (37,38); yogurt, at least 16-20 years (39); calcium, at least 12-16 years (40); processed red meat (41), at least 4-8 years; and aspirin, approximately 10 years for sporadic CRC (42) and 3-5 years for Lynch Syndrome patients (43). Interestingly, recent evidence suggests that at earlier stages (>10 years before CRC diagnosis) a lower dose of aspirin may be sufficient, but at later stages, within the past 5-10 years, a higher dose may be required to reduce risk of CRC (42). A long time lag could represent that an exposure acts early in carcinogenesis and/or that it takes a long time for a cancer to develop. It is even possible that some risk factors may have opposing actions early and later in carcinogenesis, as has been suggested for folate (44,45).

Aging increases the likelihood of an individual carrying prevalent neoplasms with multiple driver mutations, perhaps making interventions less effective. Of note, many randomized interventions for primary prevention tend to be done at older ages to maximize case numbers for statistical power and with relatively short follow-up time for feasibility. Such trials may be prone to miss true effects. For example, in the Linxian General Population Nutrition Intervention Trial, a benefit of total cancer (mostly gastrointestinal) mortality was seen for antioxidants only in those younger than 55 years, leading the authors to postulate benefit primarily earlier in the course of carcinogenesis and a "point of no return," beyond which supplementation with vitamins is ineffective (46). At older ages (>70 years), aspirin appears to lose effectiveness against CRC (47,48); molecularly, it is possible that most CRCs that occur in people at this age already have amassed the necessary driver mutations.

Risk Factors for Precursors of CRC

The vast majority of CRC develops either through the conventional adenoma-carcinoma continuum or the serrated pathway (49,50). The accessibility of these precursors for CRC through colonoscopy has helped create a parallel epidemiology of CRC precursors. The risk factors for CRC2 are consistently risk factors for overall adenomas, high-risk (eg, large, dysplastic) adenomas, or serrated polyps. Among 13 factors summarized as associated with CRC risk in a recent review (2), all have been associated with adenomas or serrated polyps (51). The relative importance of each risk factor may differ by type of lesion. For example, tobacco and alcohol appear relatively more important for serrated polyps, and dietary factors are more strongly related to conventional adenomas (51). Thus, risk factors for cancer are reflected in their precursors, and the magnitude of their association with CRC and high-risk adenoma is approximately the same (52). That the epidemiology of CRC and adenoma (especially high-risk adenoma) is largely interchangeable further supports that common CRC risk factors are operative early in carcinogenesis.

Because adenomas and serrated polyps are useful endpoints to study the effect of exposures in observational studies and randomized interventions, it is important to consider them from the mutational perspective. In the model by Lahouel et al. (3), the first driver mutation occurs on average at age 14 years, and by age 25-30 years, almost all would have occurred for CRCs that are eventually diagnosed. Based on one estimate, only an extremely small proportion of microneoplasms in crypts (carrying an initial driver mutation) transform into a macroscopic adenoma (<1 in 375 000) or CRC (<1 in 3 million) within the following few decades (53). Macroscopic adenomas may represent lesions with 1 or 2 driver mutations, or perhaps all 3 mutations (especially for large, dysplastic adenomas, often called carcinoma-in-situ) but for which the cancer has not manifested yet as invasion through the basement membrane, the clinical definition for cancer. Although not all advanced adenomas become malignant, the natural history and risk factors of advanced adenomas and CRCs would be predicted to largely overlap, as is indeed observed.

Heterogeneity of CRC by Subtype and Anatomic Site

The large bowel is often treated as a homogenous organ, yet the probability of a CRC developing varies greatly depending on the location within the bowel. Expressed as relative incidence of CRC per unit of area, the relative incidence rises from increasing distance from the transverse colon to the cecum sevenfold and to the rectum 21-fold (54). Furthermore, if we consider molecular subtypes of CRC as distinct pathways for CRC, the differential between the rectum and transverse colon to develop CRC of the "conventional pathway" (nonmicrosatellite-instable-high, non-CIMP-high) is approximately 30-fold (55). Despite the rectum's much higher propensity to develop CRC than the transverse colon per unit surface area, the rectum has lower proliferation based on labeling index than the colon (56). In addition, risk factors for CRC are differential (at least for magnitude of effect) for distinct CRC molecular subtypes and across segments of the large bowel (57-59).

Differences due to the content of the blood supply (eg, carcinogens, growth factors) do not appear plausible to explain 30-fold variation in CRC susceptibility across segments of the large intestine. There are marked differences in the composition of the stool as it travels across the bowel (60,61), resulting in local differences in many features, such as water content, oxygenation, pH, nutrient availability, inflammation and immunity, carcinogen concentration, and composition of the microbiome (62). The presumed mechanisms for some risk factors for CRC involve microbiome effects (eg, fiber, yogurt) or luminal effects (eg, calcium). Whole grain and cereal fiber intake had an increasingly stronger inverse association from cecum to rectum, and calcium intake had a stronger inverse association in the distal compared with the proximal colon (58). Bacterial content may potentially influence the conversion of alcohol to acetaldehyde in the mucosa (63,64). Alkylating damage is more common in distal colon compared with proximal colon tumors (20). Presumably, alkylating damage is induced by N-nitroso-compounds, which are microbial metabolic products of heme iron or meat nitrites and nitrates (65,66). In addition, some bacteria may be directly oncogenic. For example, the proportion of *F. nucleatum* –high CRCs gradually increased statistically significantly from rectal cancers (2.5%) to cecal cancers (11%) (67). *F. nucleatum* appears to inhibit the immune response (68), which is particularly relevant for proximal CRC.

Are the EHR Model and Epidemiologic Perspectives Compatible?

The EHR model, which indicated that 71% of mutations for CRC are random (R) mutations, generated controversy because of the apparent incompatibility with epidemiology, which suggests that the majority of CRC is preventable. Yet, these perspectives may not be inconsistent (69). Many identified CRC risk and protective factors are not obviously mutagenic. Rather, they may involve mechanisms such as growth factors (eg, insulin-like growth factors), hormones, nutrients related to DNA synthesis and repair, and inflammation, among others, as encompassed in a broader carcinogenic model, such as that from Hanahan and Weinberg (70,71). In fact, as summarized above, some CRC risk (or protective) factors may modulate cell proliferation rates, suggesting that E and R mutations are not mutually exclusive as suggested by the EHR model.

In addition to potentially causing solely random replication errors, the degree of cell proliferation likely multiplies the effect of risk factors. In particular, tissues with a high cell turnover (such as the large intestine) tend to have a higher cancer incidence (eg, 0.5% lifetime risk even in a low-risk population) compared with tissues with low proliferation rates (eg, 0.005% lifetime risk) (15,16). Yet, it is likely that external factors amplify this risk. For example, a tenfold elevation in risk because of external risk factors would increase the lifetime risk from 0.5%, typical for a low-incidence CRC population, to 5%, typical for a high-incidence CRC population. As argued previously, the major types and most abundant cancers in each population appear to be in organs that exhibit relatively high stem cell division and that have prevalent risk factors (72). In this regard, it is notable that although the small intestine is 4 times as long as the large intestine and encompasses 90% of the absorptive surface area of the gastrointestinal tract, occurrence of small intestine adenocarcinoma is approximately 50 times rarer than large intestine adenocarcinoma. Because the gut microbiome influences the tumor yield and location in mice genetically prone for polyposis (73), the much greater microbial density in large intestine (10^{11} cells/g feces) compared with the small intestine (10^{3}-10^{8} cells/g feces) (74) may in part explain the propensity for CRC rather than small intestine adenocarcinoma in humans.

Conclusion

The mutational and epidemiologic approaches can inform each other on important future directions. Epidemiology has identified numerous risk and protective factors that putatively enhance the likelihood of a driver mutation to occur or contributes to a selection advantage for cells with driver mutations. However, most risk factors are unlikely to be directly mutagenic. The mutational model reinforces the importance of early-life risk factors for CRC. Interestingly, the estimated time of the first driver mutation coincides with emerging evidence that risk factors operate at childhood, adolescence, and early adulthood, and increasing prevalence of risk factors during this time may possibly account for the increased incidence of CRC at younger ages in some populations.

Various lines of research efforts are needed for a more complete understanding of CRC. The mutational models need to account for risk factors that are not directly mutagenic and operate through processes such as enhanced proliferation or inflammation. They need to incorporate that cancer susceptibility of different parts of the colorectum varies profoundly. Genetic epidemiology can continue to complement traditional epidemiology in identifying genetic risk factors. More studies linking risk factors to specific molecular features in CRC tissue ("molecular pathologic epidemiology") will help integrate the epidemiologic and molecular approaches. Knowing when risk factors act is critical in designing and interpreting interventions to reduce risk of CRC. Specifically, innovations are needed to examine how risk factors operate over the life course, particularly risk factors likely to be operative early in life. Better integrating the epidemiology of CRC precursors, the appropriate targets for screening, with the epidemiology of CRC might also improve our understanding of the natural history of CRC.

Funding

None.

Notes

Role of the funder: Not applicable.

Disclosures: Author has no disclosures to submit.

Author contribution: EG: conceptualization; writing, original draft; writing, editing and revision.

Data Availability

No new data were generated or analyzed in support of this research.

References

1. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61(5):759–767.
2. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol*. 2019; 16(12):713–732.

3. Knudson AG. Two genetic hits (more or less) to cancer. *Nat Rev Cancer*. 2001; 1(2):157–162.
4. Moolgavkar SH, Luebeck EG. Multistage carcinogenesis: population-based model for colon cancer. *J Natl Cancer Inst*. 1992;84(8):610–618.
5. Luebeck EG, Moolgavkar SH. Multistage carcinogenesis and the incidence of colorectal cancer. *Proc Natl Acad Sci USA*. 2002;99(23):15095–15100.
6. Meza R, Jeon J, Moolgavkar SH, Luebeck EG. Age-specific incidence of cancer: phases, transitions, and biological implications. *Proc Natl Acad Sci USA*. 2008; 105(42):16284–16289.
7. Luebeck EG, Curtius K, Jeon J, Hazelton WD. Impact of tumor progression on cancer incidence curves. *Cancer Res*. 2013;73(3):1086–1096.
8. Gerstung M, Jolly C, Leshchiner I, et al.; PCAWG Consortium. The evolutionary history of 2,658 cancers. *Nature*. 2020;578(7793):122–128.
9. Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal cancer incidence patterns in the United States, 1974-2013. *J Natl Cancer Inst*. 2017;109(8):djw322.
10. Zhang K, Dai H, Liang W, Zhang L, Deng Z. Fermented dairy foods intake and risk of cancer. *Int J Cancer*. 2019;144(9):2099–2108.
11. Kim H, Wang K, Song M, Giovannucci EL. A comparison of methods in estimating population attributable risk for colorectal cancer in the United States. *Int J Cancer*. 2021;148(12):2947–2953.
12. Islami F, Sauer AG, Miller KD, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin*. 2018;68(1):31–54.
13. Thygesen LC, Gronbaek M, Johansen C, Fuchs CS, Willett WC, Giovannucci E. Prospective weight change and colon cancer risk in male US health professionals. *Int J Cancer*. 2008;123(5):1160–1165.
14. Fang Z, Song M, Lee D, Giovannucci EL. The role of Mendelian randomization studies in deciphering the effect of obesity on cancer [published online ahead of print]. *J Natl Cancer Inst*. 2022;114(3):361–371.
15. Tomasetti C, Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*. 2015; 347(6217):78–81.
16. Tomasetti C, Li L, Vogelstein B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science*. 2017;355(6331):1330–1334.
17. Blokzijl F, Ligt J, Jager M, et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature*. 2016;538(7624):260–264.
18. Giovannucci E. A framework to understand diet, physical activity, body weight, and cancer risk. *Cancer Causes Control*. 2018;29(1):1–6.
19. Noreen F, Roosli M, Gaj P, et al. Modulation of age- and cancer- associated DNA methylation change in the healthy colon by aspirin and lifestyle. *J Natl Cancer Inst*. 2014;106(7):dju161.
20. Gurjao C, Zhong R, Haruki K, et al. Discovery and features of an alkylating signature in colorectal cancer. *Cancer Discov*. 2021;11(10):2446–2455.
21. Song M, Hu FB, Spiegelman D, et al. Long-term status and change of body fat distribution, and risk of colorectal cancer: a prospective cohort study. *Int J Epidemiol*. 2016;45(3):871–883.
22. Steinbach G, Heymsfield S, Olansen NE, Tighe A, Holt PR. Effect of caloric restriction on colonic proliferation in obese persons: implications for colon cancer prevention. *Cancer Res*. 1994;54(5):1194–1197.
23. Steinbach G, Kumar SP, Reddy BS, Lipkin M, Holt PR. Effects of caloric restriction and dietary fat on epithelial cell proliferation in rat colon. *Cancer Res*. 1993;53(12):2745–2749.
24. Rogers AC, Hanly AM, Collins D, Baird AW, Winter DC. Review article: loss of the calcium-sensing receptor in colonic epithelium is a key event in the pathogenesis of colon cancer. *Clin Colorectal Cancer*. 2012;11(1):24–30.
25. Yang W, Liu L, Masugi Y, et al. Calcium intake and risk of colorectal cancer according to expression status of calcium-sensing receptor (CASR). *Gut*. 2018; 67(8):1475–1483.
26. Meza R, Jeon J, Renehan AG, Luebeck EG. Colorectal cancer incidence trends in the United States and United Kingdom: evidence of right- to left-sided biological gradients with implications for screening. *Cancer Res*. 2010;70(13): 5419–5429.
27. Luebeck EG, Hazelton WD, Curtius K, et al. Implications of epigenetic drift in colorectal neoplasia. *Cancer Res*. 2019;79(3):495–504.
28. Murphy CC, Cirillo PM, Krigbaum NY, et al. Maternal obesity, pregnancy weight gain, and birth weight and risk of colorectal cancer. *Gut*. 2021; gutjnl-2021-325001.
29. Sugiyama H, Misumi M, Brenner A, et al. Radiation risk of incident colorectal cancer by anatomical site among atomic bomb survivors: 1958-2009. *Int J Cancer*. 2020;146(3):635–645.
30. Giovannucci E, Colditz GA, Stampfer MJ, et al. A prospective study of cigarette smoking and risk of colorectal adenoma and colorectal cancer in U.S. women. *J Natl Cancer Inst*. 1994;86(3):192–199.
31. Hur J, Smith-Warner S, Rimm EB, et al. Alcohol intake in early adulthood and risk of colorectal cancer: three large prospective cohort studies of men and women in the United States. *Eur J Epidemiol*. 2021;36(3):325–333.
32. Yue Y, Hur J, Cao Y, et al. Prospective evaluation of dietary and lifestyle pattern indices with risk of colorectal cancer in a cohort of younger women. *Ann Oncol*. 2021;32(6):778–786.
33. Hur J, Otegbeye E, Joh HK, et al. Sugar-sweetened beverage intake in adulthood and adolescence and risk of early-onset colorectal cancer among women. *Gut*. 2021;70(12):2330–2336.
34. Nimptsch K, Lee DH, Zhang X, et al. Dairy intake during adolescence and risk of colorectal adenoma later in life. *Br J Cancer*. 2021;124(6):1160–1168.
35. Rezende LFM, Lee DH, Keum N, et al. Physical activity during adolescence and risk of colorectal adenoma later in life: results from the Nurses' Health Study II. *Br J Cancer*. 2019;121(1):86–94.
36. Nimptsch K, Wu K. Is timing important? The role of diet and lifestyle during early life on colorectal neoplasia. *Curr Colorectal Cancer Rep*. 2018;14(1):1–11.
37. Wang F, Wu K, Li Y, et al. Association of folate intake and colorectal cancer risk in the postfortification era in US women. *Am J Clin Nutr*. 2021;114(1): 49–58.
38. Lee JE, Willett WC, Fuchs CS, et al. Folate intake and risk of colorectal cancer and adenoma: modification by time. *Am J Clin Nutr*. 2011;93(4):817–825.
39. Michels KB, Willett WC, Vaidya R, Zhang X, Giovannucci E. Yogurt consumption and colorectal cancer incidence and mortality in the Nurses' Health Study and the Health Professionals Follow-Up Study. *Am J Clin Nutr*. 2020; 112(6):1566–1575.
40. Zhang X, Keum N, Wu K, et al. Calcium intake and colorectal cancer risk: results from the Nurses' Health Study and Health Professionals Follow-Up study. *Int J Cancer*. 2016;139(10):2232–2242.
41. Bernstein AM, Song M, Zhang X, et al. Processed and unprocessed red meat and risk of colorectal cancer: analysis by tumor location and modification by time. *PLoS One*. 2015;10(8):e0135959.
42. Zhang Y, Chan AT, Meyerhardt JA, Giovannucci EL. Timing of aspirin use in colorectal cancer chemoprevention: a prospective cohort study. *J Natl Cancer Inst*. 2021;113(7):841–851.
43. Burn J, Gerdes AM, Macrae F, et al.; CAPP2 Investigators. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. *Lancet*. 2011;378(9809): 2081–2087.
44. Mason JB. Folate, cancer risk, and the Greek god, Proteus: a tale of two chameleons. *Nutr Rev*. 2009;67(4):206–212.
45. Luebeck EG, Moolgavkar SH, Liu AY, Boynton A, Ulrich CM. Does folic acid supplementation prevent or promote colorectal cancer? Results from model-based predictions. *Cancer Epidemiol Biomarkers Prev*. 2008;17(6):1360–1367.
46. Qiao YL, Dawsey SM, Kamangar F, et al. Total and cancer mortality after supplementation with vitamins and minerals: follow-up of the Linxian general population nutrition intervention trial. *J Natl Cancer Inst*. 2009;101(7):507–518.
47. McNeil JJ, Nelson MR, Woods RL, et al.; ASPREE Investigator Group. Effect of aspirin on all-cause mortality in the healthy elderly. *N Engl J Med*. 2018; 379(16):1519–1528.
48. Guo CG, Ma W, Drew DA, et al. Aspirin use and risk of colorectal cancer among older adults. *JAMA Oncol*. 2021;7(3):428–435.
49. Hawkins N, Norrie M, Cheong K, et al. CpG island methylation in sporadic colorectal cancers and its relationship to microsatellite instability. *Gastroenterology*. 2002;122(5):1376–1387.
50. Spring KJ, Zhao ZZ, Karamatic R, et al. High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. *Gastroenterology*. 2006;131(5):1400–1407.
51. He X, Wu K, Ogino S, Giovannucci EL, Chan AT, Song M. Association between risk factors for colorectal cancer and risk of serrated polyps and conventional adenomas. *Gastroenterology*. 2018;155(2):355–373.
52. Cao Y, Rosner BA, Ma J, et al. Assessing individual risk for high-risk colorectal adenoma at first-time screening colonoscopy. *Int J Cancer*. 2015;137(7): 1719–1728.
53. Lee-Six H, Olafsson S, Ellis P, et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature*. 2019;574(7779):532–537.
54. Sadahiro S, Ohmura T, Saito T, Suzuki S. Relationship between length and surface area of each segment of the large intestine and the incidence of colorectal cancer. *Cancer*. 1991;68(1):84–87.
55. Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut*. 2012;61(6): 847–854.
56. Potten CS, Kellett M, Rew DA, Roberts SA. Proliferation in human gastrointestinal epithelium using bromodeoxyuridine in vino: data for different sites, proximity to a tumour, and polyposis coli. *Gut*. 1992;33(4):524–529.
57. Wang L, He X, Ugai T, et al. Risk factors and incidence of colorectal cancer according to major molecular subtypes. *JNCI Cancer Spectr*. 2020;5(1):pkaa089.
58. Wang L, Lo CH, He X, et al. Risk factor profiles differ for cancers of different regions of the colorectum. *Gastroenterology*. 2020;159(1):241–256.
59. Murphy N, Ward HA, Jenab M, et al. Heterogeneity of colorectal cancer risk factors by anatomical subsite in 10 European countries: a multinational cohort study. *Clin Gastroenterol Hepatol*. 2019;17(7):1323–1331.
60. Hong PY, Croix JA, Greenberg E, Gaskins HR, Mackie RI. Pyrosequencing-based analysis of the mucosal microbiota in healthy individuals reveals ubiquitous bacterial groups and micro-heterogeneity. *PLoS One*. 2011;6(9):e25042.
61. Stearns JC, Lynch MDJ, Senadheera DB, et al. Bacterial biogeography of the human digestive tract. *Sci Rep*. 2011;1:170.
62. Flynn KJ, Ruffin MT, Turgeon DK, Schloss PD. Spatial variation of the native colon microbiota in healthy adults. *Cancer Prev Res (Phila)*. 2018;11(7):393–402.
63. Salaspuro MP. Acetaldehyde, microbes, and cancer of the digestive tract. *Crit Rev Clin Lab Sci*. 2003;40(2):183–208.

64. Seitz HK, Simanowski UA, Garzon FT, et al. Possible role of acetaldehyde in ethanol-related rectal cocarcinogenesis in the rat. *Gastroenterology*. 1990; 98(2):406–413.
65. Bastide NM, Pierre FHF, Corpet DE. Heme iron from meat and risk of colorectal cancer: a meta-analysis and a review of the mechanisms involved. *Cancer Prev Res (Phila)*. 2011;4(2):177–184.
66. Massey RC, Key PE, Mallett AK, Rowland IR. An investigation of the endogenous formation of apparent total N-nitroso compounds in conventional microflora and germ-free rats. *Food Chem Toxicol*. 1988;26(7):595–600.
67. Mima K, Cao Y, Chan AT, et al. Fusobacterium nucleatum in colorectal carcinoma tissue according to tumor location. *Clin Transl Gastroenterol*. 2016;7(11):e200.
68. Hamada T, Zhang X, Mima K, et al. Fusobacterium nucleatum in colorectal cancer relates to immune response differentially by tumor microsatellite instability status. *Cancer Immunol Res*. 2018;6(11):1327–1336.
69. Song M, Vogelstein B, Giovannucci EL, Willett WC, Tomasetti C. Cancer prevention: molecular and epidemiologic consensus. *Science*. 2018;361(6409): 1317–1318.
70. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57–70.
71. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646–674.
72. Giovannucci EL. Are most cancers caused by specific risk factors acting on tissues with high underlying stem cell divisions? *J Natl Cancer Inst*. 2015;108(3): djv343.
73. Son JS, Khair S, Pettet DW III, et al. Altered interactions between the gut microbiome and colonic mucosa precede polyposis in APCMin/+ mice. *PLoS One*. 2015;10(6):e0127985.
74. Eckburg PB, Bik EM, Bernstein CN, et al. Diversity of the human intestinal microbial flora. *Science*. 2005;308(5728):1635–1638.