Body size, physical activity, genetic variants in the insulin-like growth factor pathway and colorectal cancer risk

Carcinogenesis, 2015, Vol. 36, No. 9, 971–981

doi:10.1093/carcin/bgv077
Advance Access publication May 29, 2015
Original Manuscript

ORIGINAL MANUSCRIPT

Body size, physical activity, genetic variants in the insulin-like growth factor pathway and colorectal cancer risk

C.C.J.M.Simons, L.J.Schouten, R.Godschalk[1], M.van Engeland[2], P.A.van den Brandt, F.J.van Schooten and M.P.Weijenberg*

Department of Epidemiology, GROW—School for Oncology and Developmental Biology and [1]Department of Toxicology, NUTRIM—School for Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, The Netherlands and [2]Department of Pathology, GROW—School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

*To whom correspondence should be addressed. Tel: +31 43 3882358; Fax: +31 43 3884128; Email: mp.weijenberg@maaastrichtuniversity.nl

ABSTRACT

Insulin-like growth factors (IGFs) have been associated with growth, body size, physical activity and colorectal cancer (CRC). We hypothesized that variants in IGF-related genes increase the CRC susceptibility associated with a larger body size and a lack of physical activity. We assessed this in The Netherlands Cohort Study. Participants (n = 120 852) completed a baseline questionnaire on diet and cancer. ~75% returned toenail clippings. Using a case-cohort approach and 16.3 years of follow-up, toenail DNA from 3768 subcohort members and 2580 CRC cases was genotyped. We aggregated unfavorable alleles (potentially increasing CRC risk) for 18 single nucleotide polymorphisms in 8 genes into a sum score. The sum score (in tertiles) and an *IGF1* 19-CA repeat polymorphism (19/19, 19/non-19 and non-19/non-19 repeats) in combination with body size (mostly in tertiles) and (non-)occupational physical activity (>12, 8–12 and <8 kJ/min in the job and >90, >60–90, >30–60 and ≤30 min/day) were analyzed by Cox regression. Increasingly higher hazard ratios (HRs) for CRC were observed for a larger adult body mass index, larger trouser size and tallness in the presence of more unfavorable alleles in men. HRs (95% confidence intervals) for joint effects were 1.55 (1.06–2.25), 1.78 (1.29–2.46) and 1.48 (1.01–2.17), respectively. In women, variant repeat alleles halved CRC risk irrespective of body size and physical activity. Almost no interactions tested significant. To conclude, a larger body size was a CRC risk factor in men in the presence of an accumulation of unfavorable alleles in IGF-related genes, but interactions were generally nonsignificant.

Introduction

Colorectal cancer (CRC) has emerged as a complex disease as indicated by that risk factors differ in men and women and for cancers occurring in different subsites in the colorectum. Body fatness, tallness and a lack of physical activity are risk factors for CRC (1). Associations are clearer in men than in women and with respect to the colon as compared with the rectum (1). This heterogeneity in associations may be due to that different mechanisms are at play in men and women or in relation to different subsites. The study of gene–environment interactions (GxE) will help identify the mechanisms through which body size and physical activity influence CRC risk.

A mechanism of interest in the context of body size, physical activity and CRC is the insulin-like growth factor (IGF) pathway. The IGF pathway is involved in normal growth and putatively tumorigenesis (2). Key players in the IGF pathway are IGF-1, the main growth factor in adult life and IGF binding proteins, which regulate IGF-1 availability and which can act as tumor suppressors locally (2). Involvement of the IGF pathway in linking body fatness and physical activity to colorectal tumorigenesis

Received: November 3, 2014; **Revised:** May 20, 2015; **Accepted:** May 22, 2015

Abbreviations

BMI	Body mass index
CRC	colorectal cancer
GxE	gene–environment interactions
HRs	hazard ratios
HRT	hormone replacement therapy
IGFs	insulin-like growth factors;
RERI	relative excess risk due to interaction
SNPs	single nucleotide polymorphisms
NLCS	The Netherlands Cohort Study

is plausible considering several lines of observational evidence. First of all, obesity has been associated with blood levels of IGF-1 (inverse U-shaped relationship) and IGFBP-1 and -2 (inverse relationship) (3). Second, higher levels of IGF-1 and lower levels of IGFBPs have been associated with an increased CRC risk (3). Third, genetic variants in IGF-related genes have been associated with CRC risk (4–24). Included here are studies on genetic variants in genes encoding for adiponectin, its receptors and peroxisome proliferator-activated receptor gamma in relation to CRC risk. These adipokines are worth considering because these have been associated with glucose and lipid homeostasis, insulin resistance and compensatory hyperinsulinaemia, and thereby potentially influence IGF-1 levels and CRC risk (25–27). Finally, type 2 diabetics have been shown to be at an increased CRC risk, which might be in part explained by hyperinsulinaemia stimulating the production of IGF-1 (28).

Only few studies have investigated joint effects of genetic variants in IGF-related genes with body size and physical activity on CRC risk. Those that were conducted yielded inconsistent results (23,29–31). Using data from The Netherlands Cohort Study (NLCS), we investigated GxE interactions between body size and physical activity and single nucleotide polymorphisms (SNPs) in IGF-related genes by means of a genetic sum score. The genetic sum score optimized power and allowed for the quantification of sex- and subsite-specific risks. We also studied joint effects of an *IGF1* 19-CA repeat polymorphism with body size and physical activity, because this polymorphism was associated with CRC risk in the literature, though inconsistently (7,21–23). This study builds on previous studies within the NLCS, which have shown a larger body size, a lack of physical activity (32,33) and a higher genetic sum score to be CRC risk factors, particularly in men, whereas variant *IGF1* 19-CA repeat alleles were found to decrease CRC risk in women (manuscript submitted for publication).

Materials and methods

Study population and design

The NLCS includes 120 852 men and women who were between 55 and 69 years old at baseline in 1986, when completing a self-administered questionnaire on diet and cancer. Participants originate from the general population in The Netherlands and were sampled via the municipal population registries. The NLCS has been described in detail previously (34). The baseline questionnaire included a semi-quantitative 150-item food frequency questionnaire, which was found to rank individuals adequately according to dietary intake when compared with a 9-day dietary record (35), and was shown a good indicator of intake for at least 5 years (36). Along with returning the questionnaire, participants were asked to return toenail clippings by way of an enclosed envelope. Approximately 90 000 participants provided toenail clippings. Toenail DNA isolation is performed according to the DNA extraction protocol of Cline *et al.* (37). with some adjustments (38). The NLCS was approved by the review boards of the TNO Nutrition and Food Research Institute (Zeist, The Netherlands) and Maastricht University (Maastricht, The Netherlands).

DNA isolation, the processing of questionnaires and the follow-up are performed using a case-cohort approach. This approach entails that a subcohort (n = 5000), which was randomly selected immediately after baseline, is followed up through linkage to the Central Bureau of Genealogy and municipal registries to estimate the accumulated person-time at risk (~100% completeness). Participants who reported a history of cancer (other than skin cancer) were excluded from follow-up, leaving 4774 subcohort members. The whole cohort is followed up for incident cancer cases through linkage to the population-based cancer registry and PALGA (The Netherlands pathology database) (>96% completeness) (39,40). After 16.3 years, there were 3440 incident CRC cases. Toenail clippings were available for 3768 subcohort members (78.9%) and 2580 CRC cases (75.0%), of which 114 CRC cases in the subcohort. The subcohort is representative of the total cohort, and so the 114 subcohort CRC cases were included in both counts, leaving a total of 6234 unique individuals with toenail samples for genotyping.

Variant selection and genotyping

Gene and SNP selection. We selected genes encoding for factors in or regulatory to the IGF pathway and genes encoding for adiponectin, adiponectin receptors and peroxisome proliferator-activated receptor gamma. We searched the literature for SNPs in these genes. We required that SNPs had been significantly associated with a selected endpoint at least twice or with more than 1 selected endpoint (an exception was made for missense variants). Endpoints included CRC risk, relevant traits (i.e. obesity, insulin resistance or blood levels of IGF pathway-related factors), type 2 diabetes mellitus risk and the risk of other obesity-related cancers [cancers of the oesophagus, pancreas, gallbladder, breast (in postmenopausal women), endometrium and kidney (41)]. Our literature-based strategy avoided overfitting of the cumulative model due to potential false-positive findings in a single dataset. A more elaborate description of our SNP selection strategy and prioritization is available in Supplemental Material, available at *Carcinogenesis* Online. Prioritization was necessary, because the iPLEX™ assay for the SEQUENOM® MassARRAY® platform (Sequenom, Hamburg, Germany) allows high-throughput genotyping of a maximum of 40 SNPs at once. Not all SNPs can be combined due to sequence incompatibilities between the sequences flanking the SNPs. In total, 25 SNPs in 9 genes could be included in the assay.

SNP genotyping. The protocol for genotyping on the SEQUENOM® MassARRAY® platform has been described previously (42) and was carried out using 100 ng of toenail DNA of 6234 subcohort members and CRC cases, pipetted into 384-well plates. Included were duplicate samples for a random selection of 314 samples and 436 water controls. Twenty-four out of the 25 SNPs in the assay were successfully genotyped. Genotyping of SNP rs35767 failed as only the C-allele was found. Four samples were excluded because our laboratory technicians noted a possible contamination. The reproducibility of genotypes was 98.8% or higher for the different SNPs. All SNPs had call rates of 92.6% or higher, except SNP rs4773082, which had a call rate of 83.6%. All SNPs adhered to Hardy Weinberg equilibrium in subcohort members, except SNP rs1342387 (P value = 0.02, Supplementary Table 1, available at *Carcinogenesis* Online). Since one of these 24 SNPs may be expected to deviate from Hardy Weinberg equilibrium on the basis of chance alone when considering that our study population, although large, is a random sample of the base population, and since all SNPs were genotyped simultaneously, rendering genotyping errors unlikely, we did not exclude this SNP. Exclusion of samples with irreproducible results (n = 1), and samples with a call rate less than 95% (n = 532, 8.5%) resulted in 5697 samples for further analysis.

Genetic sum score. For 18 of the 24 SNPs, the literature was unequivocal as regards to which allele was the unfavorable allele (Supplementary Table 1, available at *Carcinogenesis* Online). Alleles were considered 'unfavorable' if associated with CRC endpoints, type 2 diabetes mellitus or other obesity-related cancers in a risk-increasing manner, or if associated with overweight, obesity, insulin resistance or blood levels of IGF pathway-related factors in a manner that may increase CRC risk. Unfavorable alleles were aggregated into a genetic sum score, which was categorized into tertiles as based on the distribution in the subcohort. The rationale for using the genetic sum score was that it integrates information across genes, which is important because there may be functional compensation between genes (43) and gene–gene interactions (44). 134 subcohort

members and 120 CRC cases could not be categorized due to missing SNP data (one SNP was missing at most). Furthermore, exclusion of participants with inconsistent/incomplete baseline questionnaires left 3069 subcohort members and 2154 CRC cases across tertiles.

Genotyping of the IGF1 19-CA repeat polymorphism. The *IGF1* 19-CA repeat polymorphism was genotyped by PCR amplification and subsequent analysis of the PCR products' length using the 96-capillary ABI 3730xl DNA Analyzer. The PCR was carried out using 100 ng of DNA, 10.75 µl MilliQ, 2.5 µl 10× PCR buffer, 0.875 µl of 50 mM MgCL2, 2 × 0.125 µl of Primer predilution-mix (10 times diluted), 0.5 µl of 10 mM deoxynucleoside triphosphate mix and 0.125 µl of Platinum *Taq* polymerase (Life Technologies, Bleiswijk, The Netherlands). The primers (forward: 5′-ACCACTCTGGGAGAAGGGTA-3′; reverse: 5′-GCTAGCCAGCTGGTGTTATT-3′) were fluorescently labelled with 6-FAM (blue), NED (yellow) and PET (red), which enabled the simultaneous analysis of three samples in a single run on the ABI 3730xl DNA Analyzer. The protocol was carried out in the dark because of the light-sensitivity of the fluorescent labels. The PCR reactions were performed using the following cycles: 94°C for 10 min, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 30 s, followed by 72°C for 10 min and 4°C for 30 min. The analysis included 314 duplicate samples and 436 water controls. The reproducibility of the *IGF1* 19-CA repeat analysis was 93.6%. Genotyping was successful for 70.7% of samples.

IGF1 19-CA repeat categorization. The *IGF1* 19-CA repeat polymorphism was categorized according to Rosen *et al.* (45)., distinguishing between individuals homozygous for the wild-type allele (19/19 CA repeats), heterozygous individuals (19/non-19 CA repeats) and individuals carrying two variant alleles (non-19/non-19 CA repeats). The *IGF1* 19-CA repeat polymorphism was not in Hardy Weinberg equilibrium in the subcohort when taking into account the multiallelic character of this locus (*P* value < 0.001). However, it must be noted that deviations may arise due to the presence of rare alleles and genotypes, which is the case in our population. We, therefore, did not exclude this polymorphism from further analysis. We analyzed this variant separately from the genetic sum score for two reasons. Primarily, the *IGF1* 19-CA repeat polymorphism may be a conceptually different variant than a SNP, meaning the assumption that all variants in the genetic sum score have a similar weight may not hold for this variant in particular. Second, previous studies showed increased (7,22,23) and decreased CRC risks (21) for variant repeat alleles, rendering the unfavorable allele unknown. Exclusion of participants with inconsistent/incomplete questionnaires left 2134 subcohort members and 1833 CRC cases in categories of the *IGF1* 19-CA repeat polymorphism.

Body size

Information derived from the baseline questionnaire indicative of body size included adult body mass index (BMI; weight divided by height squared, kg/m^2), trouser/skirt size (Dutch clothing sizes), height (cm) and BMI at age 20. All variables were categorized into sex-specific tertiles as based on the distribution in the subcohort, except trouser/skirt size, which was categorized as <median sex-specific sizes and ≥median sex-specific sizes. Trouser/skirt size correlated well with hip and waist circumferences in a subset of weight-stable NLCS participants and was associated with endometrial and renal cell cancer risk in a fashion as would be expected for waist circumference (46). When adjusted for BMI, waist circumference is thought to reflect abdominal fatness.

Physical activity

In the NLCS, occupational physical activity in men and non-occupational physical activity in women were indicative of long-term physical activity (this difference exists, because not many women held jobs or only briefly and in the distant past) (33). Information on occupational physical activity was derived from an individual's self-reported longest held job. The categorization used distinguishes between jobs with an occupational energy expenditure of <8, 8–12 and >12 kJ/min (47). Non-occupational physical activity in minutes per day was a sum measure of several activities: daily walking/cycling (min/day), weekly recreational walking/cycling, weekly engagement in gardening/doing odd jobs and weekly participation in sports/gymnastics (categories: never, 1, 1–2 and >2 h/week).

Statistical analysis

Sex- and subsite-specific hazard ratios (HRs) and 95% confidence intervals for CRC were estimated using Cox regression for all linear combinations of categories of genetic variants and body size and physical activity. In this study on joint effects, we coded variables such that higher combined categories were expected to increase risk. This entailed that we reversed the coding of the physical activity variables. We refrained from recoding the *IGF1* 19-CA repeat polymorphism, because the literature was unclear about which allele was risk-increasing.

Each Cox model yielded between 6 and 12 HRs (depending on the number of combined categories). We will highlight three HRs in our description of the results: those indicative of "genetic effects," "effects of body size and physical activity" and "joint effects." "Genetic effects" are indicated by HRs comparing the highest versus lowest tertile of the genetic sum score and hazard rations comparing variant versus wild-type *IGF1* 19-CA repeat alleles in the presence of a small body size or a high level of physical activity. "Effects of body size or physical activity" are indicated by HRs comparing individuals with a large versus small body size and HRs comparing a low versus high physical activity level in the presence of few unfavorable alleles or the wild-type *IGF1* 19-CA repeat allele. "Joint effects" are indicated by HRs comparing individuals in highest versus lowest combined categories. These HRs were also used for calculating the relative excess risk due to interaction (RERI). This RERI was derived from the formula $RERI = RR_{11} - RR_{10} - RR_{01} + 1$ (48). In this formula, RR_{11}, RR_{10} and RR_{01} correspond to the relative risk (or HR) observed for the joint effect, the genetic effect and the effect of body size or physical activity, as described above. Corresponding 95% bias-corrected confidence intervals were estimated by bootstrapping (*n* bootstrap samples = 1000) (49). The RERI is a measure for additive interaction, that is, departure from additivity of effects on a HR scale. In addition, we assessed multiplicative interactions using the Wald test.

To account for the additional variance introduced by sampling the subcohort from the entire cohort, standard errors were estimated using the robust Huber–White sandwich estimator (50). The proportional hazards assumption was tested using the scaled Schoenfeld residuals and by visually inspecting the -log-log-transformed hazard curves (there were no apparent violations). Models were adjusted for predefined potential confounders [age (years), smoking status (never, ex, current), alcohol intake (0, 0.1–29 ≥30 g/d), total energy intake (kcal/d), processed meat intake (g/d), meat intake (g/d) and first-degree family history of CRC (yes/no)]. In addition, all models, except models for physical activity, were adjusted for physical activity; models for trouser/skirt size and physical activity were adjusted for adult BMI (kg/m^2); and models for height were adjusted for weight (kg). To check for the influence of preclinical disease, a sensitivity analysis was conducted in which the first 2 years of follow-up were excluded (with no essential changes in results).

Cox regression analyses were conducted using Stata version 12 (Stata Corp., College Station, TX). Graphical plots were produced using R version 2.15.1 (the R Foundation for Statistical Computing). Statistical significance was indicated by a *P* value <0.05 for two-sided testing. We did not correct for multiple testing, because our study was hypothesis-based and our use of a genetic sum score significantly reduced the number of tests that had to be performed.

Results

Baseline characteristics

Table 1 shows the distribution of subcohort members and CRC cases across categories of the genetic sum score, the *IGF1* 19-CA repeat polymorphism, adult BMI, trouser/skirt size, height, BMI at age 20, physical activity and potential confounders. Comparison between subcohort members and CRC cases in men and women most clearly showed a difference in the percentage of individuals with a family history of CRC.

The genetic sum score and body size and physical activity

Men. A pattern of increasing CRC risks was observed across combined categories between the genetic sum score and adult BMI, trouser size, height, BMI at age 20 and physical activity in men (Table 2). HRs indicative of joint effects reached statistical significance for combinations including adult BMI, trouser size and height. HRs (95% CIs) for CRC comparing the highest with

Table 1. Baseline characteristics of male and female subcohort members and CRC cases in the Netherlands Cohort Study (1986–2002)

	Male subcohort		Male CRC cases		Female subcohort		Female CRC cases	
	N (%)	Mean (SD)	*N* (%)	Mean (SD)	*N* (%)	Mean (SD)	*N* (%)	Mean (SD)
Genetic variants								
Genetic sum score[a]								
Tertile 1	588 (37.7)		417 (32.1)		563 (37.3)		279 (32.6)	
Tertile 2	603 (38.7)		514 (39.6)		552 (36.6)		343 (40.1)	
Tertile 3	369 (23.7)		367 (28.3)		394 (26.1)		234 (27.3)	
IGF1 CA repeat								
19/19	414 (36.8)		402 (36.4)		293 (29.0)		264 (36.2)	
19/non-19	452 (40.2)		430 (39.0)		366 (36.2)		290 (39.8)	
Non-19/non-19	258 (23.0)		272 (24.6)		351 (34.8)		175 (24.0)	
Body size								
Adult BMI, kg/m^2[b]		24.9 (2.6)		25.2 (2.6)		25.1 (3.6)		24.8 (3.4)
Adult trouser/skirt size								
<Median, sex-specific	490 (38.6)		364 (33.1)		493 (44.4)		371 (44.6)	
≥Median	780 (61.4)		736 (66.9)		618 (55.6)		461 (55.4)	
BMI at age 20, kg/m^2[c]		21.7 (2.4)		21.8 (2.3)		21.4 (2.7)		21.5 (2.8)
Height, cm[d]		177 (6.5)		177 (6.8)		165 (6.1)		166 (6.2)
Physical activity								
Occupational physical activity								
<8 kJ/min	834 (60.3)		733 (60.5)					
8–12	362 (26.2)		308 (25.4)					
>12	186 (13.5)		171 (14.1)					
Non-occupational physical activity								
≤30 min/day					320 (22.0)		225 (26.7)	
>30–60					478 (32.9)		248 (29.4)	
>60–90					337 (23.2)		199 (23.6)	
>90					318 (21.9)		172 (20.4)	
Potential confounders								
Age, years		61.2 (4.2)		61.8 (4.1)		61.4 (4.2)		62.2 (4.0)
Smoking status								
Never	176 (12.7)		124 (10.2)		830 (57.1)		507 (60.1)	
Ex	748 (54.1)		726 (59.9)		316 (21.7)		182 (21.6)	
Current	458 (33.1)		362 (29.9)		307 (21.1)		155 (18.4)	
Alcohol intake								
0 g/day	191 (13.8)		142 (11.7)		467 (32.1)		268 (31.8)	
0.1–29	974 (70.5)		852 (70.3)		934 (64.3)		534 (63.3)	
≥30	217 (15.7)		218 (18.0)		52 (3.6)		42 (5.0)	
Family history of CRC								
No	1303 (94.3)		1092 (90.1)		1363 (93.8)		760 (90.1)	
Yes	79 (5.7)		120 (9.9)		90 (6.2)		84 (10.0)	
Meat intake, g/day		105.3 (42.9)		105.3 (41.2)		93.2 (39.9)		92.5 (40.1)
Processed meat intake, g/day		16.8 (17.1)		17.9 (17.6)		10.9 (12.3)		11.1 (11.6)
Total energy intake, kcal/day		2160 (483)		2163 (485)		1,687 (391)		1687 (372)

Abbreviations: BMI, body mass index; CRC, colorectal cancer; IGF1, insulin-like growth factor 1.

[a]Genetic sum score of unfavorable alleles in the IGF pathway. The range in tertiles of the literature-based genetic sum score was 6–14, 15–18 and 19–29 unfavorable alleles. The theoretical maximum was 36.

[b]The adult BMI range in sex-specific tertiles was 13.6–23.9, 23.9–25.9 and 25.9–39.7 kg/m^2 in men; the range in sex-specific tertiles was 14.5–23.5, 23.4–26.1 and 26.0–41.6 kg/m^2 in women.

[c]The BMI range at age 20 in sex-specific tertiles was 11.3–20.8, 20.7–22.6 and 22.6–33.1 kg/m^2 in men; the range in sex-specific tertiles was 11.2–20.2, 20.1–22.5 and 22.6–33.1 kg/m^2 in women.

[d]The height range in sex-specific tertiles was 150–173, 174–179 and 180–200 cm in men; the range in sex-specific tertiles was 140–163, 164–168 and 169–185 cm in women.

the lowest tertile of the genetic sum score in the presence of a small body size (genetic effects), comparing a large with a small body size in the lowest tertile of the genetic sum score (effects of body size), and comparing highest with lowest combined categories (joint effects) were 1.20 (0.83, 1.75), 0.98 (0.69, 1.40) and 1.55 (1.06, 2.25) for BMI; 1.56 (1.08, 2.25), 1.34 (0.98, 1.83) and 1.78 (1.29, 2.46) for trouser size; and 1.30 (0.89, 1.91), 1.30 (0.90, 1.87) and 1.48 (1.01, 2.17) for height, respectively. In addition, a significantly increased HR for CRC comparing the highest with the lowest tertile of the genetic sum score was observed in the presence of a high physical activity level (HR = 1.83, 95% CI: 1.00, 3.35). Other estimated HRs were not significant.

Subsite-specific analyses showed a similar pattern as seen for CRC in relation to colon but not rectal cancer in men (Figure 1 and Supplementary Table 2, available at *Carcinogenesis*

Table 2. HRs and 95% confidence intervals for colorectal cancer in relation to combinations of tertiles of the genetic sum score of unfavorable alleles in the IGF pathway and categories of adult BMI, adult trouser size, height, BMI at age 20 and physical activity in men in the Netherlands Cohort Study

		Adult BMI, kg/m^2													
		T1-sex-specific				T2				T3					
		PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	2,234	123	1	Reference	2,565	117	0.76	(0.54, 1.07)	2,154	126	0.98	(0.69, 1.40)		
	T2	2,520	131	0.89	(0.64, 1.25)	2,057	155	1.27	(0.91, 1.78)	2,346	173	1.25	(0.90, 1.75)		
	T3	1,438	101	1.20	(0.83, 1.75)	1,709	102	1.03	(0.72, 1.49)	1,303	129	1.55	(1.06, 2.25)	0.36	(–0.37, 1.01)
		Adult trouser size													
		<Median-sex-specific				≥Median									
		PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)					RERI	(95% CI)[b]
Genetic sum score	T1	2,757	113	1	Reference	3,604	216	1.34	(0.98, 1.83)						
	T2	2,373	130	1.35	(0.96, 1.90)	3,990	286	1.61	(1.19, 2.18)						
	T3	1,541	101	1.56	(1.08, 2.25)	2,556	203	1.78	(1.29, 2.46)					–0.11	(–0.93, 0.50)
		Height, cm													
		T1-sex-specific				T2				T3					
		PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	2,467	119	1	Reference	2,269	111	0.90	(0.64, 1.28)	1,201	136	1.30	(0.90, 1.87)		
	T2	2,269	152	1.37	(0.98, 1.91)	2,197	145	1.35	(0.97, 1.90)	2,457	162	1.22	(0.87, 1.72)		
	T3	1,493	96	1.30	(0.89, 1.91)	1,507	121	1.55	(1.07, 2.23)	1,449	115	1.48	(1.01, 2.17)	–0.12	(–0.88, 0.57)
		BMI at age 20, kg/m^2													
		T1-sex-specific				T2				T3					
		PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	1,912	109	1	Reference	2,126	101	0.87	(0.59, 1.26)	1,201	103	1.03	(0.70, 1.50)		
	T2	2,126	124	1.00	(0.70, 1.43)	1,892	118	1.11	(0.77, 1.61)	1,670	129	1.29	(0.88, 1.88)		
	T3	1,201	100	1.35	(0.90, 2.01)	1,117	80	1.24	(0.82, 1.88)	1,212	94	1.39	(0.93, 2.08)	0.01	(–0.78, 0.76)
		Occupational physical activity, kJ/min													
		>12				8–12				<8					
		PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	PY	N cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	939	46	1	Reference	1,643	92	1.16	(0.72, 1.87)	4,372	228	1.03	(0.67, 1.58)		
	T2	912	64	1.35	(0.79, 2.30)	1,777	116	1.33	(0.84, 2.13)	4,233	279	1.31	(0.86, 2.00)		
	T3	497	48	1.83	(1.00, 3.35)	1,245	85	1.33	(0.81, 2.19)	2,708	199	1.44	(0.93, 2.23)	–0.42	(–2.05, 0.42)

Abbreviations: BMI, body mass index; PY, person-years at risk; RERI, relative excess risk due to interaction; T, tertile. Note on multiplicative interactions: none of the multiplicative interactions tested using the Wald test were statistically significant, except for the interaction between the genetic sum score and BMI in relation to colon cancer risk. The RERI is a measure of additive interaction.

[a]All models were adjusted for age, smoking status, alcohol intake, processed meat intake, meat intake, total energy intake and first-degree family history of colorectal cancer. In addition, all models, except models for occupational physical activity, were adjusted for occupational physical activity; models for adult trouser size and occupational physical activity were adjusted for BMI; and models for height were adjusted for weight.

[b]Bias-corrected confidence interval.

Online). Across subsites, most strongly increased risks were observed in relation to the distal colon for joint effects of the genetic sum score with trouser size and height (HR = 2.03, 95% CI: 1.27, 3.24 and HR = 2.05, 95% CI: 1.16, 3.63, respectively) (Supplementary Table 2, available at *Carcinogenesis* Online). The HR indicative of a joint effect with physical activity on colon cancer also reached statistical significance. However, HRs comparing the highest with the lowest tertile of the genetic sum score in the presence of a high physical activity level indicated a similarly, non-significantly increased colon cancer risk and a significantly increased distal colon cancer risk (HR = 1.76, 95% CI: 0.88, 3.49; HR = 2.39, 95% CI: 1.01, 5.65, respectively). No association with (distal) colon cancer was observed when comparing a low with a high physical activity level in the lowest tertile of the genetic sum score. This suggests that the observed joint effect with physical activity may be reflecting the genetic effect.

Women. A risk pattern was not clear across combined categories of the genetic sum score and adult BMI, trouser/skirt size, height, BMI at age 20 and physical activity in women (Table 3). Most HRs for CRC were not statistically significant, except for combinations including height and physical activity. HRs (95% CIs) for CRC comparing the highest with the lowest tertile of the genetic sum score in the presence of a small body size or a high physical activity level (genetic effects), comparing a large with a small body size or a high with a low physical activity level in the lowest tertile of the genetic sum score (effects of body size) and comparing highest with lowest combined categories (joint effects) were 1.65 (1.08, 2.50), 1.83 (1.22, 2.73) and 1.80 (1.15, 2.80) for height; and 1.25 (0.74, 2.11), 1.62 (1.03, 2.54) and 1.88 (1.15, 3.07) for physical activity.

Subsite-specific analyses in women were generally consistent with overall results (Figure 2 and Supplementary Table 3, available at *Carcinogenesis* Online). The HR indicative of a joint effect of the genetic sum score with physical activity was strongly increased in relation to rectal cancer, but the confidence interval was wide (HR = 4.22, 95% CI: 1.70, 10.47) (Panel B of Figure 2). This suggests that this HR was unstable, warranting caution when interpreting the joint effect of the genetic sum score with physical activity on CRC, which may be reflecting results for rectal cancer.

The IGF1 19-CA repeat polymorphism and body size and physical activity

Men. No risk pattern was observed across combined categories of the *IGF1* 19-CA repeat polymorphism and adult BMI, trouser size, height, BMI at age 20 and physical activity in men (Panel A of Figure 3, Supplementary Table 4, available at *Carcinogenesis* Online). None of the estimated HRs for CRC were statistically significant, except for the HR indicative of a joint effect of variant *IGF1* 19-CA repeat alleles with adult BMI (HR = 1.66, 95% CI: 1.04, 2.65) (purple bullets in Panel A of Figure 3). In subsite-specific analyses, this HR was not statistically significant.

Women. We also observed no risk pattern across combined categories of the *IGF1* 19-CA repeat polymorphism and adult BMI, trouser size, height, BMI at age 20 and physical activity in women (Panel B of Figure 3 and Supplementary Table 5, available at *Carcinogenesis* Online). Decreased CRC risks were observed for joint effects of variant repeat alleles with a high BMI and ≥median trouser/skirt size (purple bullets in Panel B of Figure 3), but even stronger decreased CRC risks were evident for variant repeat alleles in the presence of a low BMI, <median trouser/skirt size, short height, low BMI at age 20 and high level of physical activity (blue bullets in Panel B of Figure 3). The latter HRs for CRC were 0.56 (95% CI: 0.36, 0.87), 0.51 (95% CI: 0.35, 0.75), 0.62 (95% CI: 0.40, 0.96), 0.42 (95% CI: 0.25, 0.71) and 0.41 (95% CI: 0.23, 0.76), respectively. Subsite-specific analyses for colon but not rectal cancer showed similar results as for CRC.

GxE interaction tests

Most gene–environment interactions were not statistically significant when tested on an additive scale using the RERIs or when tested on a multiplicative scale (Supplementary Tables 2–5, available at *Carcinogenesis* Online). Exceptions were the multiplicative interaction between the genetic sum score and adult BMI in relation to colon cancer risk in men; the multiplicative interaction between the genetic sum score and physical activity in relation to CRC risk in women (*P* values < 0.05); and the RERI for combinations of the *IGF1* 19-CA repeat polymorphism with trouser size and BMI at age 20 in relation to rectal cancer in men

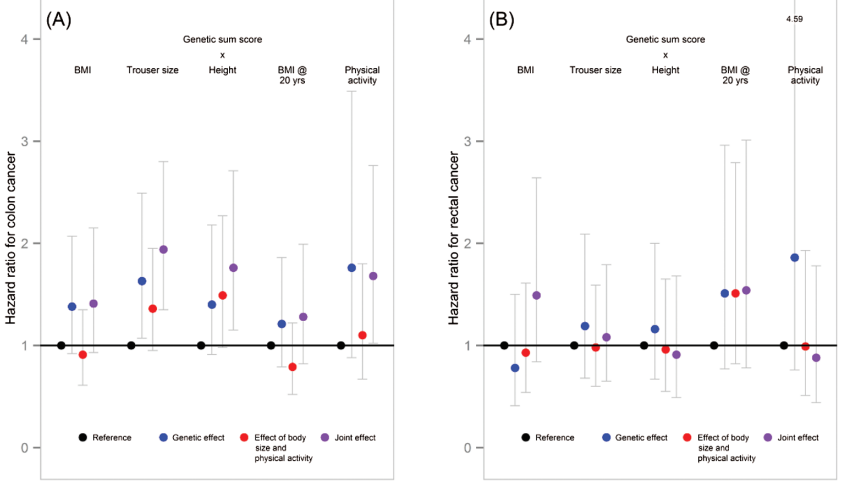

Figure 1. Plots show HRs and 95% confidence intervals for (**A**) colon cancer and (**B**) rectal cancer in men. Blue bullets show HRs indicative of genetic effects comparing the highest tertile of the genetic sum score with the lowest in the presence of a small body size or high physical activity level. Red bullets show HRs indicative of body size and physical activity effects comparing highest categories of body size with lowest and comparing a low physical activity level with a high level in the lowest tertile of the genetic sum score. Purple bullets show HRs indicative of joint effects comparing highest combined categories with lowest.

Table 3. HRs and 95% confidence intervals for colorectal cancer in relation to combinations of tertiles of the genetic sum score of unfavorable alleles in the IGF pathway and categories of adult BMI, adult trouser/skirt size, height, BMI at age 20 and physical activity in women in The Netherlands Cohort Study

		Adult BMI, kg/m^2													
		T1-sex-specific				T2				T3					
		PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	2,406	130	1	Reference	2,528	128	0.98	(0.67, 1.43)	2,727	136	0.80	(0.55, 1.17)		
	T2	2,621	144	1.27	(0.88, 1.82)	2,711	168	0.99	(0.69, 1.42)	2,210	183	1.20	(0.82, 1.74)		
	T3	1,823	112	1.15	(0.77, 1.71)	1,849	108	1.06	(0.71, 1.58)	1,852	134	0.98	(0.65, 1.47)	−0.01	(−0.65, 0.53)
		Adult trouser/skirt size													
		<Median-sex-specific				≥Median									
		PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)					RERI	(95% CI)[b]
Genetic sum score	T1	3,235	118	1	Reference	4,363	236	0.99	(0.69, 1.41)						
	T2	3,401	143	1.16	(0.83, 1.61)	4,076	306	1.29	(0.92, 1.82)						
	T3	2,234	109	1.34	(0.94, 1.91)	3,193	216	0.97	(0.68, 1.40)					−0.35	(−1.07, 0.17)
		Height, cm													
		T1-sex-specific				T2				T3					
		PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	2,857	127	1	Reference	2,716	120	1.63	(1.12, 2.37)	2,088	147	1.83	(1.22, 2.73)		
	T2	2,513	164	1.42	(0.96, 2.09)	2,972	154	2.01	(1.40, 2.88)	2,057	177	1.96	(1.32, 2.90)		
	T3	1,767	103	1.65	(1.08, 2.50)	2,314	132	1.59	(1.07, 2.36)	1,443	119	1.80	(1.15, 2.80)	−0.62	(−1.73, 0.40)
		BMI at age 20, kg/m^2													
		T1-sex-specific				T2				T3					
		PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	2,352	118	1	Reference	2,317	107	1.26	(0.84, 1.89)	2,317	113	1.35	(0.90, 2.01)		
	T2	2,413	135	1.45	(0.99, 2.14)	2,155	130	1.70	(1.15, 2.52)	2,246	139	1.40	(0.94, 2.09)		
	T3	1,544	106	1.48	(0.96, 2.28)	1,841	86	1.31	(0.86, 2.00)	1,741	99	1.18	(0.76, 1.83)	−0.66	(−1.67, 0.07)
		Non-occupational physical activity, min/day													
		>90				>60–90				>30–60				≤30					
		PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	PY	*N* cases	HR[a]	(95% CI)	RERI	(95% CI)[b]
Genetic sum score	T1	1,632	130	1	Reference	1,817	98	1.15	(0.72, 1.83)	2,527	108	0.76	(0.48, 1.20)	1,686	58	1.62	(1.03, 2.54)		
	T2	1,859	170	1.12	(0.71, 1.78)	1,681	96	1.45	(0.92, 2.28)	2,254	157	1.52	(0.99, 2.33)	1,749	72	1.25	(0.79, 1.98)		
	T3	1,045	117	1.25	(0.74, 2.11)	1,314	69	1.21	(0.73, 2.00)	2,121	110	0.95	(0.59, 1.50)	1,044	58	1.88	(1.15, 3.07)	−0.03	(−1.35, 0.93)

Abbreviations: BMI, body mass index; PY, person-years at risk; RERI, relative excess risk due to interaction; T, tertile. Note on multiplicative interactions: none of the multiplicative interactions tested using the Wald test were statistically significant, except for the interaction between the genetic sum score and non-occupational physical activity in relation to colorectal cancer risk. The RERI is a measure of additive interaction.

[a]All models were adjusted for age, smoking status, alcohol intake, meat intake, processed meat intake, total energy intake and first-degree family history of colorectal cancer. In addition, all models, except models for non-occupational physical activity, were adjusted for non-occupational physical activity; models for adult trouser/skirt size and non-occupational physical activity were adjusted for adult BMI; and models for height were adjusted for weight.

[b]Bias-corrected confidence interval.

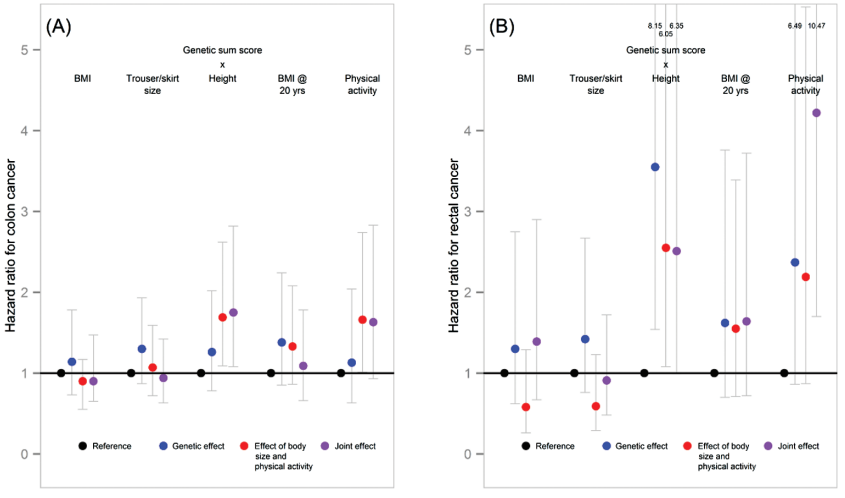

Figure 2. Plots show HRs and 95% confidence intervals for (**A**) colon cancer and (**B**) rectal cancer in women. Blue bullets show HRs indicative of a genetic effect comparing the highest tertile of the genetic sum score with the lowest in the presence of a small body size or high physical activity level. Red bullets shows HR indicative of body size and physical activity effects comparing highest categories of body size with lowest and comparing a low physical activity level with a high level in the lowest tertile of the genetic sum score. Purple bullets show HRs indicative of a joint effect comparing highest combined categories with lowest.

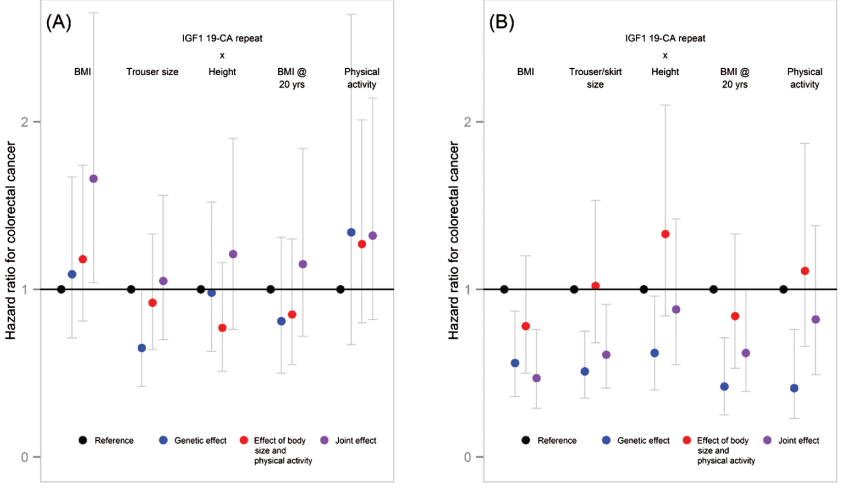

Figure 3. Plots show HRs and 95% confidence intervals for colorectal cancer in (**A**) men and (**B**) women. Blue bullets show HRs indicative of genetic effects comparing variant repeat alleles with wild-type alleles in the presence of a small body size or high physical activity level. Red bullets show HRs indicative of effects of body size and physical activity comparing highest categories of body size with lowest and comparing a low physical activity level with a high level in wild-type repeat allele carriers. Purple bullets show HRs indicative of joint effects comparing highest categories of body size in variant repeat allele carriers with lowest categories of body size in wild-type repeat allele carriers and comparing a low physical activity level in variant repeat allele carriers with a high physical activity level in wild-type allele carriers.

(RERI = 0.63, 95% CI: 0.03, 1.14 and RERI = 1.05, 95% CI: 0.10, 2.09, respectively).

Discussion

We hypothesized that the IGF pathway is a biological mechanism through which body size and physical activity may influence CRC risk, and therefore assessed joint effects of genetic variants in IGF-related genes with body size and physical activity. To our knowledge, this is the most comprehensive study on this topic. We observed that a larger body size was a risk factor for colon but not rectal cancer in men in the presence of an accumulation of unfavorable alleles in IGF-related genes as indicated by a higher genetic sum score. Risk was significantly increased with 50–110%. We also found halved colon cancer risks in women for variant *IGF1* 19-CA repeat alleles, but these were irrespective of body size and physical activity. Most interaction tests did not reach significance.

Previous studies showed interactions between physical activity and SNP rs2665802 in *GH1* (29), and between BMI and SNPs rs2289046 in *IRS2* (31), rs1063538 in *ADIPOQ* (30) and rs1539355 in *ADIPOR1* (30). There were no interactions for SNPs rs6214 in *IGF1* (31), rs3110697 (31) and rs2854744 (23) in *IGFBP3*, rs1801278 in *IRS1* (23), rs1805097 in *IRS2* (23), other tagSNPs in *ADIPOQ* and *ADIPOR1* (30), or the *IGF1* 19-CA repeat polymorphism (23). Interestingly, the observed joint effects in men in our study were present in relation to both proximal and distal colon cancer risk, whereas marginal associations based on NLCS data mainly showed associations between body size and distal colon cancer risk (32). This perhaps reiterates that CRC has a complex multifactorial etiology as reflected in that it is a heterogeneous disease. On the other hand, it cannot be excluded that the joint effects in relation to proximal colon cancer risk simply reflected main effects of the genetic sum score (manuscript submitted for publication), especially considering the absence of significant tests for interaction.

That joint effects of the genetic sum score with body size in relation to CRC risk at all subsites were less clear in women may be logical considering that female reproductive hormones might counteract the insulin resistance associated with body fatness. This is suggested by experimental findings showing that estrogens improved insulin sensitivity in obesity-induced mice, influencing CRC progression (51). In addition, the European Prospective Investigation into Cancer observed a positive association between waist/hip ratio and colon cancer risk in postmenopausal women not using hormone replacement therapy (HRT), whereas this association was absent in women using HRT (52). Possibly, therefore, HRT counteracted the insulin resistance associated with abdominal fatness, nullifying CRC risk in women in the European Prospective Investigation into Cancer. This fits with that HRT in itself has been associated with a reduced CRC risk (53). Additional adjustment for HRT use in our data did not change results, but only a small percentage of female subcohort members reported HRT use (12.8%).

The general lack of statistically significant additive and multiplicative interactions along with the observation of joint effects of the genetic sum score with body size on colon cancer risk in men is confusing. However, where statistically significant interactions would have strengthened our findings, the absence of these does not argue against a biological interaction (48). This becomes apparent when thinking of the sufficient-component cause model, in which interaction is defined as the participation of two component causes in the same sufficient cause (54). Another reason for the general lack of statistically significant interactions in this study may have been power. It has been described that a four times larger sample size is needed to detect an interaction effect as compared with a marginal effect of similar magnitude (55). Thus, in the presence of small interaction effects, the power to detect a statistically significant interaction may have been limited, even in this large study. It is furthermore important to realize that statistical interaction tests require there to be main effects. This lies in the fact that these tests assess whether the observed joint effect differs from the expected joint effect (56). However, main effects need not be present. In fact, in genetics, true biological epistasis (gene–gene interaction) is described as joint effects of genetic variants that lack an effect on their own (57). The absence of statistically significant main effects in our study may have hampered the detection of statistically significant interactions. In particular, main effects of body fatness may have been lacking due to our narrow BMI range. BMI distributions were concentrated in the normal range at age 20 (18.5–<25 kg/m^2) and in the normal and overweight range in adulthood (18.5–<30 kg/m^2). A mere 0.4% of NLCS subcohort members was obese at age 20 (≥30 kg/m^2), and 6.5% at an adult age (55–69 years).

The public health relevance of investigating joint effects on CRC risk of genetic variants in IGF-related genes with body size and physical activity is that it will contribute to the evidence base underlying preventive strategies for CRC aimed at maintaining a healthy weight. For this, more research is needed to further elucidate the role of genetic variants in IGF-related genes, but also in genes related to other pathways through which body fatness potentially influences colorectal tumorigenesis, such as the PI3K/Akt/mTOR signalling pathway (58) and inflammatory pathways (59). On the basis of such studies, genetic subgroups may be discerned which could particularly benefit from targeted CRC prevention strategies.

Strengths of this study include the prospective design of the NLCS and its long follow-up, yielding large case numbers. Strengths also include our SNP selection strategy and the use of a genetic sum score. We only included SNPs that had been associated with selected endpoints in at least two previous studies or SNPs that were missense SNPs, minimizing the chance of selecting SNPs on the basis of false-positive results. Our genetic sum score of unfavorable alleles furthermore integrated information across IGF-related genes, optimized power and greatly reduced the multiple testing problem. Limitations of this study include the single baseline measurement of body size, physical activity and potential confounders. In addition, numbers did not allow for the assessment of higher-order interactions. For example, we previously reported on an interaction between trouser/skirt size and physical activity in relation to proximal colon cancer risk in women, whereas trouser/skirt size was not marginally associated with CRC endpoints in women (32). That we could not assess higher-order interactions may be another explanation for the absence of joint effects in women. Furthermore, molecular tumor subtypes, which we did not distinguish due to power limitations, may have hampered the detection of GxE interactions, although major molecular tumor subtypes in CRC, such as microsatellite instability and the CpG island methylator phenotype, correlate with the tumor location (60), for which we performed subanalyses.

To conclude, even though a larger body size was a risk factor for CRC, particularly colon cancer, in men in the presence of an accumulation of unfavorable alleles in IGF-related genes, GxE interactions did not test significant. Variant *IGF1* 19-CA repeat alleles decreased CRC risk, particularly colon cancer risk, in women irrespective of body size and physical activity.

Supplementary material

Supplementary Material and Supplementary Tables 1–5 can be found at http://carcin.oxfordjournals.org/

Funding

Dutch Cancer Society (2009-4281 to M.P.W.); Biobanking and Biomolecular Research Infrastructure Netherlands (to M.P.W.); Health Foundation Limburg (to M.P.W.).

Acknowledgements

We are indebted to the participants of this study and wish to thank The Netherlands Cancer Registry and The Netherlands nationwide registry of pathology (PALGA). We also thank Drs A. Volovics and A. Kester for statistical advice; S. van de Crommert, H. Brants, J. Nelissen, C. de Zwart, M. Moll, W. van Dijk and A. Pisters for data management; H. van Montfort, T. van Moergastel, L. van den Bosch, R. Schmeitz and J. Berben for programming assistance; L. Jonkers, J. Goessens, K. Lemmens and S. Lumeij for the laboratory work involved; and the Biobank Maastricht UMC+ for sample storage.
Conflict of Interest Statement: None declared.

References

1. Colorectal cancer | Continuous Update Project | WCRF. http://www.wcrf.org/cancer_research/cup/key_findings/colorectal_cancer.php. (10 May 2013, date last accessed).
2. Baxter, R.C. (2014) IGF binding proteins in cancer: mechanistic and clinical insights. Nat. Rev. Cancer, 14, 329–341.
3. Kaaks, R. et al. (2001) Energy balance and cancer: the role of insulin and insulin-like growth factor-I. Proc. Nutr. Soc., 60, 91–106.
4. Landi, S. et al. (2003) Association of common polymorphisms in inflammatory genes interleukin (IL)6, IL8, tumor necrosis factor alpha, NFKB1, and peroxisome proliferator-activated receptor gamma with colorectal cancer. Cancer Res., 63, 3560–3566.
5. Slattery, M.L. et al. (2004) Associations among IRS1, IRS2, IGF1, and IGFBP3 genetic polymorphisms and colorectal cancer. Cancer Epidemiol. Biomarkers Prev., 13, 1206–1214.

6. Jiang, J. et al. (2005) Influence of the C161T but not Pro12Ala polymorphism in the peroxisome proliferator-activated receptor-gamma on colorectal cancer in an Indian population. Cancer Sci., 96, 507–512.
7. Morimoto, L.M. et al. (2005) Insulin-like growth factor polymorphisms and colorectal cancer risk. Cancer Epidemiol. Biomarkers Prev., 14, 1204–1211.
8. Wong, H.L. et al. (2005) A new single nucleotide polymorphism in the insulin-like growth factor I regulatory region associates with colorectal cancer risk in Singapore chinese. Cancer Epidemiol. Biomarkers Prev., 14, 144–151.
9. Koh, W.P. et al. (2006) Peroxisome proliferator-activated receptor (PPAR) gamma gene polymorphisms and colorectal cancer risk among Chinese in Singapore. Carcinogenesis, 27, 1797–1802.
10. Theodoropoulos, G. et al. (2006) Relation between common polymorphisms in genes related to inflammatory response and colorectal cancer. World J. Gastroenterol., 12, 5037–5043.
11. Pechlivanis, S. et al. (2007) Polymorphisms in the insulin like growth factor 1 and IGF binding protein 3 genes and risk of colorectal cancer. Cancer Detect. Prev., 31, 408–416.
12. Kaklamani, V.G. et al. (2008) Variants of the adiponectin (ADIPOQ) and adiponectin receptor 1 (ADIPOR1) genes and colorectal cancer risk. JAMA, 300, 1523–1531.
13. Küry, S. et al. (2008) Low-penetrance alleles predisposing to sporadic colorectal cancers: a French case-controlled genetic association study. BMC Cancer, 8, 326.
14. Wong, H.L. et al. (2008) Insulin-like growth factor-1 promoter polymorphisms and colorectal cancer: a functional genomics approach. Gut, 57, 1090–1096.
15. Feik, E. et al. (2010) Association of IGF1 and IGFBP3 polymorphisms with colorectal polyps and colorectal cancer risk. Cancer Causes Control, 21, 91–97.
16. Tsilidis, K.K. et al. (2009) Association of common polymorphisms in IL10, and in other genes related to inflammatory response and obesity with colorectal cancer. Cancer Causes Control, 20, 1739–1751.
17. Xiang, H. et al. (2009) Association between two functional polymorphisms of insulin-like growth factor binding protein 3 and colorectal cancer risk in a Chinese population. J. Toxicol. Environ. Health. A, 72, 706–711.
18. Gao, C.M. et al. (2010) Relationship between growth hormone 1 genetic polymorphism and susceptibility to colorectal cancer. J. Hum. Genet., 55, 163–166.
19. Xu, W. et al. (2010) PPARgamma polymorphisms and cancer risk: a meta-analysis involving 32,138 subjects. Oncol. Rep., 24, 579–585.
20. He, B. et al. (2011) Effects of genetic variations in the adiponectin pathway genes on the risk of colorectal cancer in the Chinese population. BMC Med. Genet., 12, 94.
21. Keku, T.O. et al. (2012) Genetic variants in IGF-I, IGF-II, IGFBP-3, and adiponectin genes and colon cancer risk in African Americans and Whites. Cancer Causes Control, 23, 1127–1138.
22. Samowitz, W.S. et al. (2006) Polymorphisms in insulin-related genes predispose to specific KRAS2 and TP53 mutations in colon cancer. Mutat. Res., 595, 117–124.
23. Slattery, M.L. et al. (2005) Energy balance, insulin-related genes and risk of colon and rectal cancer. Int. J. Cancer, 115, 148–154.
24. Slattery, M.L. et al. (2009) Colon tumor mutations and epigenetic changes associated with genetic polymorphism: insight into disease pathways. Mutat. Res., 660, 12–21.
25. Bull, A.W. (2003) The role of peroxisome proliferator-activated receptor gamma in colon cancer and inflammatory bowel disease. Arch. Pathol. Lab. Med., 127, 1121–1123.
26. Komninou, D. et al. (2003) Insulin resistance and its contribution to colon carcinogenesis. Exp. Biol. Med. (Maywood), 228, 396–405.
27. Pais, R. et al. (2009) Metabolic syndrome and risk of subsequent colorectal cancer. World J. Gastroenterol., 15, 5141–5148.
28. Deng, L. et al. (2012) Diabetes mellitus and the incidence of colorectal cancer: an updated systematic review and meta-analysis. Dig. Dis. Sci., 57, 1576–1585.
29. Khoury-Shakour, S. et al. (2008) Recreational physical activity modifies the association between a common GH1 polymorphism and colorectal cancer risk. Cancer Epidemiol. Biomarkers Prev., 17, 3314–3318.
30. Liu, L. et al. (2011) Interactions between genetic variants in the adiponectin, adiponectin receptor 1 and environmental factors on the risk of colorectal cancer. PLoS One, 6, e27301.
31. Karimi, K. et al. (2013) Is there an association between variants in candidate insulin pathway genes IGF-I, IGFBP-3, INSR, and IRS2 and risk of colorectal cancer in the Iranian population? Asian Pac. J. Cancer Prev., 14, 5011–5016.
32. Hughes, L.A.E. et al. (2011) Body size and colorectal cancer risk after 16.3 years of follow-up: an analysis from the Netherlands Cohort Study. Am. J. Epidemiol., 174, 1127–1139.
33. Simons, C.C.J.M. et al. (2013) Physical activity, occupational sitting time, and colorectal cancer risk in the Netherlands cohort study. Am. J. Epidemiol., 177, 514–530.
34. van den Brandt, P.A. et al. (1990) A large-scale prospective cohort study on diet and cancer in The Netherlands. J. Clin. Epidemiol., 43, 285–295.
35. Goldbohm, R.A. et al. (1994) Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. Eur. J. Clin. Nutr., 48, 253–265.
36. Goldbohm, R.A. et al. (1995) Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. Eur. J. Clin. Nutr., 49, 420–429.
37. Cline, R.E. et al. (2003) The fingernails of Mary Sullivan: developing reliable methods for selectively isolating endogenous and exogenous DNA from evidence. J. Forensic Sci., 48, 328–333.
38. van Breda, S.G. et al. (2007) Toenails: an easily accessible and long-term stable source of DNA for genetic analyses in large-scale epidemiological studies. Clin. Chem., 53, 1168–1170.
39. Goldbohm, R.A. et al. (1994) Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. Tijdschr Soc Gezondheidsz, 72, 80–84.
40. Casparie, M. et al. (2007) Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell. Oncol., 29, 19–24.
41. World Cancer Research Fund / American Institute for Cancer Research. (2007) Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective. AICR, Washington DC.
42. Geybels, M.S. et al. (2014) Selenoprotein gene variants, toenail selenium levels, and risk for advanced prostate cancer. J. Natl Cancer Inst., 106, dju003.
43. Zhang, J. (2012) Genetic redundancies and their evolutionary maintenance. Adv. Exp. Med. Biol., 751, 279–300.
44. Maher, B. (2008) Personal genomes: the case of the missing heritability. Nature, 456, 18–21.
45. Rosen, C.J. et al. (1998) Association between serum insulin growth factor-I (IGF-I) and a simple sequence repeat in IGF-I gene: implications for genetic studies of bone mineral density. J. Clin. Endocrinol. Metab., 83, 2286–2290.
46. Hughes, L.A. et al. (2009) Self-reported clothing size as a proxy measure for body size. Epidemiology, 20, 673–676.
47. Hettinger, T.H. et al. (1989) Ermittlung des Arbeitsenergieumsatzes bei Dynamisch Muskulaerer Arbeit. Bundesarbeit fuer Arbeitsschutz, Dortmund, Germany.
48. Rothman, K.J. et al. (1998) Modern Epidemiology. Lippincott-Raven Publishers, Philadelphia, PA.
49. Carpenter, J. et al. (2000) Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat. Med., 19, 1141–1164.
50. Barlow, W.E. (1994) Robust variance estimation for the case-cohort design. Biometrics, 50, 1064–1072.
51. Rondini, E.A. et al. (2011) Energy balance modulates colon tumor growth: interactive roles of insulin and estrogen. Mol. Carcinog., 50, 370–382.
52. Pischon, T. et al. (2006) Body size and risk of colon and rectal cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). J. Natl Cancer Inst., 98, 920–931.
53. Grodstein, F. et al. (1999) Postmenopausal hormone therapy and the risk of colorectal cancer: a review and meta-analysis. Am. J. Med., 106, 574–582.
54. Rothman, K.J. et al. (2005) Causation and causal inference in epidemiology. Am. J. Public Health, 95 (suppl. 1), S144–S150.
55. Aschard, H. et al. (2012) Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Hum. Genet., 131, 1591–1613.
56. Greenland, S. (1993) Basic problems in interaction assessment. Environ. Health Perspect., 101 (suppl. 4), 59–66.

57. Gilbert-Diamond, D. et al. (2011) Analysis of gene-gene interactions. Curr. Protoc. Hum. Genet., Chapter 1, Unit 1.14. doi:10.1002/0471142905.hg0114s70.
58. Weijenberg, M.P. et al. (2013) The mTOR pathway and the role of energy balance throughout life in colorectal cancer etiology and prognosis: unravelling mechanisms through a multidimensional molecular epidemiologic approach. Curr. Nutr. Rep., 2, 19–26.
59. Pérez-Hernández, A.I. et al. (2014) Mechanisms linking excess adiposity and carcinogenesis promotion. Front. Endocrinol., 5, 65.
60. Simons, C.C.J.M. et al. (2013) A novel classification of colorectal tumors based on microsatellite instability, the CpG island methylator phenotype and chromosomal instability: implications for prognosis. Ann. Oncol., 24, 2048–2056.