Metabolic factors and the risk of colorectal cancer by KRAS and BRAF mutation status

Robin Myte, Björn Gylling, Jenny Häggeström, Christel Häggeström, Carl Zingmark, Anna Löfgren Burström, Richard Palmqvist and Bethany Van Guelpen

1Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden
2Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden
3Department of Statistics, Umeå School of Business and Economics, Umeå University, Umeå, Sweden
4Department of Biobank Research, Umeå University, Umeå, Sweden
5Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden
6Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden
7Department of Radiation Sciences, Oncology, Umeå University, SE-901 87 Umeå, Sweden,

Factors related to energy metabolism and the metabolic syndrome, such as higher body mass index (BMI), blood glucose, or blood lipids, and blood pressure, are associated with an increased risk of colorectal cancer (CRC). However, CRC is a heterogeneous disease, developing through distinct pathways with differences in molecular characteristics and prognosis, and possibly also in risk factors. For subtypes defined by KRAS and BRAF mutation status, BMI is the only metabolic factor previously studied, with inconsistent findings. We investigated whether associations between BMI, blood glucose, blood lipids, and blood pressure and CRC risk differ by tumor KRAS and BRAF mutation status in 117,687 participants from two population-based cohorts within the Northern Sweden Health and Disease Study (NSHDS). Hazard ratios (HRs) for overall CRC and CRC subtypes by metabolic factors were estimated with Cox proportional hazards regression, using multiple imputation to handle missing exposure and tumor data. During a median follow-up of 15.6 years, we acquired 1,250 prospective CRC cases, of which 766 cases had complete baseline and molecular tumor data. Consistent with previous evidence, higher BMI, total cholesterol, triglyceride levels, and blood pressure were associated with an increased risk of overall CRC (HRs per 1 standard deviation increase: 1.07 to 1.12). These associations were similar regardless of CRC subtype by KRAS and BRAF mutation status (all PHeterogeneity > 0.05). The same was true for subtypes based on microsatellite instability status. Poor metabolic health may therefore be a universal mechanism for colorectal cancer, acting across multiple developmental pathways.

Introduction

Colorectal cancer (CRC) is a heterogeneous disease, developing through distinct pathways and demonstrating large differences in molecular and clinical characteristics and prognosis.1,2 Molecular classification of tumors using markers such as KRAS and BRAF mutation status is already being used to guide therapy for CRC, notably the resistance to anti-EGFR therapy in KRAS-mutated tumors.3 CRC may, therefore, more accurately be described as a group of separate diseases with potential differences in etiology and risk factors.4,6

Heterogeneity in the relationship between risk factors and CRC subtypes could have implications for CRC prevention, for example, to personalize recommendations for lifestyle change, pharmacoprevention or screening. In this manner, precision medicine might impact not only disease therapy and prognostics but also public health, with the potential to reduce CRC incidence. Moreover, investigating risk factors in relation to subtypes of cancer with homogenous pathogenic mechanisms, as opposed to all subtypes combined, can provide valuable insights into the etiological mechanisms behind the

Key words: metabolic factors, colorectal cancer, KRAS, BRAF, microsatellite instability, risk factors

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

Conflict of interest: There are no financial disclosures or conflicts of interest.

Grant sponsor: Cancer Research Foundation in Northern Sweden; Grant sponsor: Swedish Cancer Society; Grant numbers: 2012/0501, 2014/780, 2017/581; Grant sponsor: Faculty of Medicine at Umeå University; Grant sponsor: Lions Cancer Research Foundation; Grant sponsor: Regional agreement between Umeå University and Västerbotten County Council (ALF)

DOI: 10.1002/ijc.32104

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

History: Received 26 Jul 2018; Accepted 18 Dec 2018; Online 7 Jan 2019

Correspondence to: Robin Myte, Department of Radiation Sciences, Oncology, Umeå University, SE-901 87 Umeå, Sweden,
Tel.: +46-90-785-2838, E-mail: robin.myte@umu.se

Int. J. Cancer: 145, 327–337 (2019) © 2019 The Authors. International Journal of Cancer published by John Wiley & Sons Ltd on behalf of UICC
**What's new?**

Factors related to the metabolic syndrome, such as higher body mass index, blood glucose, blood lipids, and blood pressure, are associated with an increased risk of colorectal cancer. However, whether metabolic factors have different roles in different molecular subtypes of colorectal cancer remains an open question. In this population-based, prospective cohort study of more than 115,000 people, the answer appears to be no. Metabolic factors were associated with colorectal cancer risk consistently across subtypes based on KRAS and BRAF mutations and microsatellite instability. Poor metabolic health may be a universal mechanism for colorectal cancer, acting across multiple developmental pathways.

associations. This field of study has been coined molecular pathological epidemiology, and has led to several key findings, such as the subtype-specific association between regular aspirin use and BRAF-wild type CRC.

Despite the potential of subtype-specific analyses to yield important information on the link between body fatness (an established CRC risk factor), metabolic health, and CRC development, data are limited. Body fatness or poor metabolic health may either cause specific molecular errors in specific developmental pathways of CRC, or contribute to a colorectal microenvironment favorable for tumors with specific molecular features to grow, for instance, through chronic low-grade inflammation. For CRC subtypes defined by anatomical site, high body fatness measured as high body mass index (BMI) is mainly associated with an increased risk of colon cancer, with small differences between the right or left colon. However, there is no conclusive evidence for a subtype-specific association for BMI by either KRAS or BRAF-mutation status, or other molecular characteristics of the tumor, such as infiltrating immune cells or microsatellite instability (MSI) status. One study reported a subtype-specific association with KRAS-mutated CRC risk for low plasma levels of adiponectin, an adipokine inversely associated with body fatness and the metabolic syndrome. Metabolic factors related to the metabolic syndrome other than body fatness, such as blood pressure and blood lipids, are associated with CRC risk, independently of BMI. While the evidence for causality for these factors is limited compared to for body fatness, there is genetic evidence supporting a causal, BMI-independent, role for total cholesterol. To our knowledge, no previous study has investigated whether metabolic factors related to the metabolic syndrome demonstrate subtype-specific relationships with CRC risk.

The aim of our study was to investigate factors involved in energy metabolism and the metabolic syndrome, including BMI, fasting blood glucose, oral glucose tolerance test results, blood lipids, and blood pressure, in relation to the risk of molecular CRC subtypes defined by KRAS and BRAF mutation status, using data from two large, population-based cohorts in northern Sweden. In secondary analysis, we also investigated subtypes by MSI status.

**Materials and Methods**

**Study population**

This was a cohort study of two prospective, population-based cohorts within the Northern Sweden Health and Disease Study (NSHDS): the Västerbotten Intervention Programme (VIP) and the northern Swedish MONitoring of Trends and Determinants in CArdiovascular Disease (MONICA) project. Both cohorts have been described in detail elsewhere. In short, the VIP, initiated in 1986 and ongoing, is a health screening intervention in which all residents of Västerbotten County are invited to a general health exam at 10 year intervals at ages 40, 50 and 60 years. Participants fill out an extensive questionnaire on health and lifestyle and may also donate a blood sample for biobanking purposes. The participation rate is approximately 70% (50% in early years), with no indication of any major selection. MONICA consists of randomly selected 25–74 year olds living in northern Sweden (Västerbotten and Norrbotten County) invited to participate in six cross-sectional health surveys between 1986 and 2014 (average participation rate 74%).

At the final date of entrance to our study (January 19, 2016), the NSHDS included 119,738 participants with 183,699 observations before exclusions (see Fig. 1 for study design). Missing data for day or month were replaced with Day 15 for day and June for month. Participants diagnosed with cancer other than non-melanoma skin cancer before study entrance were excluded, with missing data on height or weight, or with implausible measurements (height < 130 or > 210 cm, weight < 35 kg, BMI < 15 or > 70 kg/m², fasting plasma glucose <1 mmol/L, glucose tolerance test <1 or > 35 mmol/L, total cholesterol <0.5 or > 15 mmol/L, triglyceride levels <0.15 or > 20 mmol/L, systolic blood pressure < 20 or > 300 mm Hg, diastolic blood pressure < 20 or > 250 mm Hg). For participants who attended more than one health exam over time, we used the observation with the longest follow-up. After exclusions, the study population included 117,687 participants (VIP: 108107 participants, MONICA: 9580 participants). The study protocol was approved by the Regional Ethical Review Board in Umeå, Sweden. All participants gave a written informed consent for all collection for research purposes.

**Metabolic factors**

All measurements were made by a health professional. Height and weight were measured in light clothing without shoes. A capillary blood sample was drawn after an overnight fast, and again 2 h after a 75 g oral glucose load (to measure glucose tolerance, 80% had fasted longer than 8 h). Plasma glucose concentrations were analyzed with a Reflotron bench-top
analyzer (Roche Diagnostics) until 2004, and from 2004 with a Hemocue bench-top analyzer (Quest Diagnostics). Total cholesterol and triglyceride levels in the VIP were analyzed with a Reflotron bench-top analyzer until 2009, and from 2009 with an enzymatic method at the clinical chemistry laboratory at the nearest hospital. Systolic and diastolic blood pressure were measured once with a mercury sphygmomanometer after a 5-min rest in the supine position until 2009, and from 2009 in the sitting position. To harmonize measurements before and after 2009 in the VIP, total cholesterol, triglyceride, and blood pressure measurements were calibrated using formulas presented in Supporting Information Table S1. Glucose, lipid, and blood pressure measurements in the MONICA cohort used the same methods as after 2009 in the VIP. We adjusted lipid and blood pressure measurements for lipid-lowering or antihypertensive medication usage by adding estimated constants (total cholesterol +1.347 mmol/L, triglyceride levels +0.208 mmol/L, systolic blood pressure + 15 mm Hg, and diastolic blood pressure +10 mm Hg).27,28

To facilitate comparisons between metabolic factors, they were analyzed as continuous z-transformed variables (scaled to mean 0 and standard deviation (SD) 1). z-Transformations were made separately by sex and cohort for BMI and blood pressure, and by sex, cohort, and fasting status (above or below 8 h) for glucose, glucose tolerance, cholesterol, and triglycerides. Because of a skewed distribution, the glucose, glucose tolerance, and triglycerides variables were log-transformed before z-transformation. We observed expected correlations between BMI, glucose, triglycerides, and blood pressure (r = 0.2 to 0.4, Fig. 2). Three groups of more correlated metabolic factors were observed: fasting glucose and glucose tolerance (r = 0.3), triglycerides and cholesterol (r = 0.3), and systolic blood pressure and diastolic blood pressure (r = 0.7).

Follow-up
Participants were followed starting from 1 year after study entry (baseline), to reduce the risk of reverse causality, until May 31, 2016 by linkage to the virtually complete Swedish national registries, for whichever occurred first of either cancer diagnosis other than non-melanoma skin cancer (Cancer Registry of Northern Sweden), death (Swedish Cause of Death Registry), or migration (Swedish Registry of Total Population and Population Changes). Participants diagnosed with cancer in the period between study entry and start of follow-up (1 year later) were excluded (Fig. 1). Participants diagnosed with colorectal adenocarcinoma were identified using ICD-10 C18.0 and C18.2–18.9 for colon, C19.9 and C20.9 for rectum. CRC diagnoses were verified, and data on anatomical tumor site and tumor stage were collected, by linkage to the Swedish National Quality Registry for Colorectal Cancer or by medical record assessment by a gastrointestinal pathologist. A total of 24 cases (2%) lacked data on tumor site, and 102 cases (8%) had missing data on stage (mainly due to incomplete clinical TNM staging).

**Tumor tissue analysis**
DNA was extracted and purified with a Qiagen QIAamp DNA FFPE Tissue Kit from formalin-fixed, paraffin-embedded tumor tissue collected during routine clinical practice at the Department of Clinical Pathology, Umeå University Hospital, Umeå, Sweden. In total, 916 of 1,250 cases (73%) had available tumor tissue.

**KRAS** was analyzed by sequencing the activating mutations in codon 12 and 13 using Big Dye v. 3.1, according to the manufacture protocol (Applied Biosystems, Life Technologies, Foster city, CA, USA).29 The mutational status of **BRAF^V600E** was analyzed using TaqMan allelic discrimination assay (reagents from Applied Biosystems)30 or digital droplet PCR (reagents from Bio-Rad Laboratories, Hercules, CA, USA). As **KRAS** and **BRAF** are considered mutually exclusive within cell clones, CRC cases were classified as **KRAS**-mutated, **BRAF**-mutated, or **KRAS**/**BRAF** wild-type. Cases with mutations in both **KRAS** and **BRAF** (n = 6, 2%, in the complete data set) were censored in the analysis.

Microsatellite instability (MSI) status was assessed by immunohistochemical analysis of MLH1, MSH2, MSH6, and PMS2 for cases diagnosed up until 2009 (samples lacking tumor cells with nuclear staining for any of the mismatch repair proteins were categorized as MSI).31 For cases...
diagnosed 2009–2016, a DNA-based method was used (Promega MSI Analysis System, Version 1.2, Madison, WI). A subset of cases (n = 70) were analyzed with both methods, with 100% concordance.

A total of 144 cases had inconclusive KRAS or BRAF mutation status, and 145 cases inconclusive MSI status, due mostly to insufficient tumor DNA in the available tissue sample. Thus, in total, 766 cases had complete KRAS and BRAF mutation status data. The probability of missing tumor data depends on clinical characteristics such as tumor site and stage, and potentially other observable characteristics, so called missing at random. In our study, cases with missing KRAS and BRAF mutation status data were more often had more recent CRC diagnosis, distal tumors, advanced stage tumors, slightly higher education level at baseline, were current or ex-smokers at baseline, and less physically inactive at baseline, compared to cases with available data (Supporting Information Table S2). We used this information to impute data in the remaining cases, as described in the statistical analysis section.

Statistical analysis
Missing data for metabolic factors, potential confounders, and tumor traits were assumed to be missing at random conditional on observed data, and imputed with multiple imputation by chained equations using the mice R-package. First, 15 data sets were imputed (number of imputation based on the highest % missing among exposures and potential confounders, 15%), in 20 iterations, with a predictive mean matching model for continuous variables and logistic or multinomial logistic regression for categorical variables. Missing data on metabolic factors or potential confounders were modeled using all metabolic factors as well as age, sex (man or woman), cohort (VIP or MONICA), fasting status (above or below 8 h), educational status (elementary school, junior secondary school, upper secondary school, or university education), smoking status (never-, ex-, or current smoker), recreational physical activity (on a scale of 1–5, from never to >3 times/week), occupational physical activity (on a scale of 1–5, from sedentary or standing work to physically strenuous most of the time), alcohol intake from a validated food frequency questionnaire (zero intake, above/below sex and questionnaire-version-specific median of self-reported intake in grams/day), event status, and follow-up time as predictors. For each of the 15 imputed data sets, missing data on stage, site, KRAS and BRAF mutation status, and MSI status in cases were imputed in 40 imputed data sets (based on the highest % missing tumor data, 39%) in 20 iterations, including the same predictors as in the exposure and potential confounder imputation and additionally the clinical and molecular tumor variables and year of and age of diagnosis. Convergence of imputed values were checked graphically, and plausibility of imputations values were checked by graphically comparing distributions with non-missing values. Analyses were run on all imputed data sets and aggregated using Rubin’s rules. To ensure that the choice of method for handling missing tumor data did not affect our results, we also ran analyses using an inverse probability weighting method, with very similar findings (data not shown).

Associations between metabolic factors and background and lifestyle-related variables were estimated by fitting linear regression models for each factor. To investigate the relative importance of each individual predictor, we estimated contributions to total $R^2$ using the lmig algorithm in the relaimpo R-package.

Associations with CRC risk were evaluated by estimating hazard ratios (HRs) per 1 SD increase in metabolic factors using Cox proportional hazards regression, with attained age as the time scale. The proportional hazards assumption was checked by visually examining time-dependent log(HRs) vs. time and in Schoenfeld residual-based tests. No violations were observed. To check for nonlinear associations, continuous variables were modeled using restricted cubic splines.
Table 1. Baseline and follow-up characteristics.

| Variable                              | All participants (n = 117,687) | CRC cases (n = 1,250) | Missing, n (%) |
|---------------------------------------|---------------------------------|-----------------------|----------------|
| **Age groups, n (%)**                 |                                 |                       |                |
| <30                                   | 985 (1)                         | 1 (0)                 |                |
| 30–39                                 | 12,077 (10)                     | 35 (3)                |                |
| 40–49                                 | 47,693 (40)                     | 167 (13)              |                |
| 50–59                                 | 31,741 (27)                     | 437 (35)              |                |
| ≥60                                   | 25,391 (22)                     | 610 (49)              |                |
| **Sex, n (%)**                        |                                 |                       |                |
| Men                                   | 58,493 (50)                     | 652 (52)              |                |
| Women                                 | 59,194 (50)                     | 598 (48)              |                |
| **BMI, kg/m²**                        | 25.3 (15.2–69.9)                | 25.9 (17.4–44.6)      | 0 (0)          |
| **Glucose, mmol/L**                   | 5.3 (1.0–24.6)                  | 5.4 (2.1–17.1)        | 5,831 (5)      |
| **Glucose tolerance, mmol/L**         | 6.4 (1.0–33.8)                  | 6.6 (1.4–23.6)        | 11,639 (10)    |
| **Total cholesterol, mmol/L**         | 5.5 (0.5–15.0)                  | 6.0 (2.6–10.9)        | 788 (1)        |
| **Triglycerides, mmol/L**             | 1.1 (0.1–19.0)                  | 1.4 (0.4–15.2)        | 17,801 (15)    |
| **Systolic BP, mm Hg**                | 124.3 (60.0–265.8)              | 132.0 (85.8–222.9)    | 779 (1)        |
| **Diastolic BP, mm Hg**               | 81.3 (30.0–164.6)               | 84.8 (61.1–132.1)     | 834 (1)        |
| **Smoking status, n (%)**             | 2071 (2)                        |                       |                |
| Never smoker                          | 57,178 (49)                     | 482 (39)              |                |
| Ex-smoker                             | 33,554 (29)                     | 434 (35)              |                |
| Current smoker                        | 24,884 (21)                     | 309 (25)              |                |
| **Occupational PA score, n (%)**      | 16,331 (14)                     |                       |                |
| 1 (sedentary or standing work)        | 26,709 (23)                     | 236 (19)              |                |
| 2 (light but partly physically active)| 18,390 (16)                     | 208 (17)              |                |
| 3 (light and physically active)       | 22,965 (20)                     | 215 (17)              |                |
| 4 (sometimes physically strenuous)    | 26,904 (23)                     | 262 (21)              |                |
| 5 (physically strenuous most of the time) | 6,388 (5)                     | 63 (5)                 |                |
| **Recreational PA score, n (%)**      | 2,592 (2)                       |                       |                |
| 1 (never)                             | 41,298 (35)                     | 514 (41)              |                |
| 2 (every now and then – not regularly)| 27,916 (24)                     | 311 (25)              |                |
| 3 (1–2 times/week)                    | 21,393 (18)                     | 228 (18)              |                |
| 4 (2–3 times/week)                    | 16,256 (14)                     | 121 (10)              |                |
| 5 (3 times/week)                      | 8,232 (7)                       | 56 (4)                |                |
| **Alcohol intake, grams/day**         | 2.8 (0.0–294.4)                 | 2.3 (0.0–31.2)        | 10,372 (9)     |
| **Alcohol intake, n (%)**             | 10,372 (9)                      |                       |                |
| Zero intake                           | 8,506 (8)                       | 91 (9)                |                |
| ≤median intake                        | 49,673 (46)                     | 510 (49)              |                |
| >median intake                        | 49,136 (46)                     | 441 (42)              |                |
| **Follow-up, years**                  | 15.6 (0.0–30.5)                 | 12.4 (1.0–29.5)       | 0 (0)          |
| **Tumor site, n (%)**                 |                                 |                       |                |
| Right colon                           | 388 (32)                        |                       |                |
| Left colon                            | 361 (29)                        |                       |                |
| Rectum                                | 477 (39)                        |                       |                |
| **Tumor stage, n (%)**                | 102 (8)                         |                       |                |
| Stage I&II                             | 582 (51)                        |                       |                |
| Stage III&IV                          | 566 (49)                        |                       |                |
| **KRAS/BRAF mutation status, n (%)**  | 484 (39)                        |                       |                |
| BRAF-mutated                          | 169 (22)                        |                       |                |
| KRAS-mutated                          | 184 (24)                        |                       |                |
| Both wild type                        | 413 (54)                        |                       |                |

(Continues)
(with knots at the 5th, 50th and 95th percentiles). Nonlinearity was tested with a likelihood ratio test comparing the spline model to a linear model. Models were adjusted for cohort, sex, smoking status, recreational physical activity, occupational physical activity, and alcohol intake.

To evaluate whether the associations between exposures and CRC risk differed by anatomical subsite or KRAS/BRAF mutation subtypes, we estimated subtype-specific HRs per 1 SD increase in metabolic factors using Cox regression models with a competing risks approach using the duplication method. In secondary analyses, we also evaluated differences in association by subtypes defined by MSI status. Heterogeneity in the association between metabolic factors and CRC risk by molecular subtypes was tested using a likelihood ratio test with a $k-1$ degrees of freedom, where $k$ equals the number of subtypes, comparing a model in which the risk association could vary across subtypes to a model in which all associations were held constant.

All computations were conducted in R v3.5.0 (R Foundation for Statistical Computing, Vienna, Austria). All tests were 2-sided when applicable. We used a significance threshold of 0.005 for analyses which were not strictly confirmatory as suggested by Benjamin et al., chosen to account for both the number of tests as well as the correlated exposures.

Results

Baseline characteristics

Approximately 40% of the study cohort were between 40 and 49 years old at baseline (median 42 years, Table 1). A total of 22% were current smokers, and 26% and 36% reported no occupational or recreational physical activity, respectively. Compared to the average cohort participant, the CRC cases were slightly higher with respect to BMI and all other metabolic factors.

Associations between background and lifestyle variables and metabolic factors are presented in Supporting Information Table S3. Age, BMI, and sex were the largest contributor to total variance for all metabolic factors ($R^2$ in full model 5–29%). Cohort, smoking status, recreational physical activity, and alcohol intake were associated with all factors, though with very small contributions to total variance ≤1%.

CANCER CASE CHARACTERISTICS

After up to 30.5 years of follow-up time (median: 15.6 years, 1.7 million person-years total follow-up, Table 1), we identified 1,250 verified CRC cases (52% men, 1,122 in VIP and 128 in MONICA). The estimated cumulative incidence at age 80 was 4.8% for men (95% CI: 4.3, 5.2%), and 3.9% for women (95% CI: 3.5, 4.3%), which is almost identical to the cumulative incidence in the underlying population in northern Sweden (4.8 and 3.8%, for men and women, respectively, data from NORDCAN© 2017 Association of the Nordic Cancer Registries, IARC. http://www-dep.iarc.fr/NORDCAN/, assessed July 26, 2018). Male sex, low recreational physical activity, and smoking were associated with an increased risk of CRC (Supporting Information Table S3).

The median age at diagnosis was 64.4 years, the relatively low age reflecting the recruitment ages of 40, 50 and 60 years in the VIP cohort. We observed 388 right-colon (32%), 361 left-colon (29%), and 477 (39%) rectal tumors (Table 1). The cases were approximately equally distributed between early tumor stages (49% stage I&II) and more advanced stages (51% stage III&IV). A total of 169 cases (22%) were BRAF mutated, 184 (24%) were KRAS mutated, and 104 (14%) were MSI. Subtypes by KRAS and BRAF mutation status reflected distinct groups of clinical characteristics. BRAF mutations were more common in women, in patients who were older at diagnosis, in right-sided colon tumors, and in MSI tumors (Table 2). KRAS-mutated tumors were approximately equally distributed across tumor sites, whereas double-wild type tumors were more often situated in the rectum.

Metabolic factors and CRC risk

No associations between metabolic factors and CRC risk showed any strong evidence of non-linearity (all $p_{\text{nonlinearity}} > 0.01$, Fig. 3), and the results here are thus for linear associations. Higher BMI was associated with an increased risk of CRC [HR per 1 SD increase, approximately per 4 kg/m$^2$: 1.07 (95% CI: 1.00, 1.13), Fig. 3]. Higher total cholesterol, triglyceride levels, and diastolic blood pressure were also associated with an increased risk of CRC, independently of BMI and other covariates (HRs: 1.07 to 1.11 per 1 SD increase in the factors). Associations were similar in men and women (all $p_{\text{heterogeneity}} > 0.3$, Supporting Information Table S4). Associations for the lipid and blood pressure variables were weaker in the smaller MONICA cohort, but tests of
heterogeneity between the cohorts was only borderline significant for systolic blood pressure ($p_{\text{heterogeneity}} = 0.006$, Supporting Information Table S4). Differences were likely due to low power in the MONICA subsample. Therefore, in order to maximize power, main analyses between metabolic factors and the risk of CRC subtypes were conducted for men and women and both cohorts combined, and assuming linear associations. Diabetes, defined as self-reported diabetes, glucose ≥7.8 mmol/L, or glucose tolerance ≥5.8 mmol/L, or glucose tolerance ≥7.8 mmol/L, was not associated with CRC risk (HR: 1.07 (95% CI: 0.95, 1.20)).

**Metabolic factors and CRC risk by clinical and molecular tumor traits**

Associations between metabolic factors and CRC risk by anatomical site, KRAS and BRAF mutation status, and MSI status are presented in Figure 4. Associations were slightly stronger for colon cancer compared to rectal cancer for most factors. However, there was no significant heterogeneity for any factor (all $p_{\text{heterogeneity}} > 0.1$). Regarding subtypes by KRAS and BRAF mutation status, no metabolic factor displayed any clear heterogeneous association across subtypes (all $p_{\text{heterogeneity}} > 0.05$). The same was true for subtypes defined by MSI status (all $p_{\text{heterogeneity}} > 0.05$). The results were similar in men and women, and in separate analyses of right-sided colon, left-sided colon, or rectal cancer (data not shown). Results were similar in complete-case analysis (where participants with missing exposure and potential confounder data were excluded) compared to results from the multiple imputation analyses, for both overall CRC risk and CRC subtype risk analyses (Supporting Information Table S5).

**Discussion**

In this population-based, prospective cohort study of 117,687 participants in the Northern Sweden Health and Disease Study (NSHDS), we investigated factors involved in energy metabolism and the metabolic syndrome in relation to the risk of colorectal cancer (CRC) by molecular subtypes defined by mutations in the KRAS and BRAF oncogenes. Consistent with previous studies, higher BMI, total cholesterol, triglyceride levels, and diastolic blood pressure were independently associated with an increased risk of CRC. We observed no clear differences in the associations between metabolic factors and CRC risk by KRAS and BRAF mutation status in the tumor. Secondary analyses of MSI status also showed similar associations across subtypes.

Mutations in the KRAS and BRAF oncogenes are important early events in CRC development. The mutations are mutually exclusive within cell clones, and combined KRAS/BRAF mutation status is largely representative of separate developmental pathways with substantial differences in clinical and molecular characteristics. Furthermore, several CRC risk factors have been reported to demonstrate subtype-specific risk relationships based on KRAS/BRAF mutation status, namely low plasma adiponectin, aspirin use, and alcohol intake. This wide spectrum of risk factors supports distinct etiological pathways.

The associations between BMI, total cholesterol, triglycerides, and blood pressure and overall CRC risk in our study were in line with previous studies. The associations appeared mainly present for colon cancer, with no clear difference between the right or left colon, which is also consistent with previous studies. We saw no association with CRC risk for blood glucose variables or diabetes. Meta-analyses of
these exposures suggest a positive association, however, there is large heterogeneity between studies and conclusive evidence for a relationship has not been established.

Our study was, to our knowledge, the first molecular pathological epidemiology study on several metabolic factors related to the metabolic syndrome and CRC risk by molecular subtypes. The homogenous risk relationships for BMI across molecular subtypes by KRAS and BRAF mutations and MSI status are consistent with previous larger studies. We observed similar homogenous risk relationships across molecular subtypes for other metabolic factors. Results to date therefore suggest that metabolic factors related to the metabolic syndrome probably do not influence colorectal carcinogenesis through a single colorectal developmental pathway.

Some factors related to the metabolic syndrome, namely low plasma adiponectin or high BMI in women, have been reported to be associated with a higher risk of KRAS-mutated CRC. Our research, using several metabolic factors related to the metabolic syndrome in a larger study cohort, did not support these previous findings. One explanation is that adiponectin is a potent anti-inflammatory agent, and inflammation is a driver of KRAS-mutated carcinogenesis. Moreover, adiponectin levels may be more strongly inversely associated with abdominal fat mass than with BMI. We mitigated the weaknesses of BMI as a marker of metabolic heath by using several metabolic factors and found similar results for each factor in relation to CRC molecular subtypes.

There are some limitations to our study. Missing data for KRAS or BRAF mutation status were due to lack of tumor tissue available (e.g. cases with more advanced tumors, who are less likely to undergo surgical resection) and unsuccessful analysis (due mostly to insufficient amounts of DNA in the accessible tissue sample). Although the proportions of missing data for molecular tumor characteristics were similar to or lower than in other molecular pathological epidemiology studies or population-based CRC studies, the risk of selection bias must be addressed.
We imputed plausible values using multivariate multiple imputation, which can effectively yield unbiased results in studies with missing at random data such as ours.32 The missing at random assumption, i.e., that all predictors of missing status were included in the imputation model, may not be entirely true. However, we were able to account for several important determinants of missing tumor data in our study. Furthermore, the results in complete-case and multiple imputation analyses were similar, indicating that large discrepancies in the results of analyses including other potential predictors of missing status independent of the predictors included in our study are unlikely. In the KRAS mutation analysis, only the codons routinely analyzed in clinical practice when the tumor analyses were started, i.e. KRAS codons 12 and 13 and no NRAS mutations, were assessed. Although these codons cover most KRAS mutations, we observed a lower frequency of KRAS mutations than could be expected using a broader analysis. Finally, using a few molecular markers to divide CRC into subtypes probably does not capture all the existing intertumoral heterogeneity.1 This can result in bias if the subtypes studied do not represent pathogenic mechanisms.49 Still, distinct pathogenic mechanisms specific to each KRAS and BRAF-mutation subtype are supported by the fact that the mutations are early oncogenic events,3 that the subtypes are associated with distinct clinical and molecular features,5,40 and that the subtypes have been shown to differ in their associations to several unrelated risk factors.

The main strength of our study is the use of extensive and high quality prediagnostic health data from a large cohort, in combination with follow-up clinical and molecular tumor data on the cases. This molecular pathological epidemiology approach makes it possible not only to decipher whether certain risk factors are associated with specific subtypes of CRC, but can also provide valuable insights into potential etiological mechanisms by connecting lifestyle factors directly to pathological pathways.5 This was, to our knowledge, the first molecular pathological epidemiology study of CRC and several metabolic factors other than BMI. By studying several metabolic factors, we were able to assess several aspects of metabolic health, which do not always co-exist with obesity,50,51 and are independently associated with CRC risk.21–23 The prospective cohort design with long follow-up (up to 31 years) reduces the risk of recall bias and reverse causation. The use of data collected by health professionals, rather than self-reported BMI and diabetes for example, also helps prevent reporting bias. Another major strength of our study is the truly population-based nature of the NSHDS cohorts, which reduces the risk of selection bias. Also, follow-up in our study used the essentially complete Swedish cancer registry.52 The unselected patient population may explain the high BRAF-mutation frequency in our study (22%, typically 4–18%53), similar to that of another unselected Scandinavian CRC study population.48

In conclusion, in this population-based cohort study, including 117,687 participants, higher levels of BMI, total cholesterol, triglycerides, and blood pressure were associated with an increased risk of CRC, regardless of molecular subtype defined by KRAS and BRAF mutation status or MSI status. Our results suggest that poor metabolic health may be a universal risk mechanism for colorectal cancer, acting across multiple developmental pathways.
Acknowledgements
We thank all the participants in the VIP and MONICA cohorts of the NSHDs for their time and willingness to participate. We thank Robert Johansson, Veronika Hellström, Åsa Ågren, and their colleagues at the Department of Biobank Research and Department of Public Health and Clinical Medicine, Nutritional Research, Umeå University for excellent assistance. We also would like to thank the staff at Biobanken Norr, Västerbotten County Council, as well as Åsa Stenberg and Roger Stening at the Department of Medical Biosciences, Pathology, Umeå University, for invaluable assistance with the tumor tissue retrieval and analyses.

References
1. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015;21:1350–6.
2. Yamauchi M, Lochhead P, Morikawa T, et al. Colorectal cancer: a tale of two sides or a continuum? Gut 2012;61:794–7.
3. Merkel M, Riemer P, Blaker H, et al. Similar but different: distinct roles for KRAS and BRAF oncosignatures in colorectal cancer development and therapy resistance. Oncometarget 2015;6:20785–800.
4. Ogino S, Lochhead P, Chan AT, et al. Molecular pathological epidemiology of epigenetics: emerging integrative science to analyze environment, host, and disease. Mod Pathol 2013;26:465–84.
5. Ogino S, Chan AT, Fuchs CS, et al. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut 2011;60:397–411.
6. Hughes LAE, Simons C, van den Brandt PA, et al. Lifestyle, diet, and colorectal cancer risk according to (Epi)genetic instability: current evidence and future directions of molecular pathological epidemiology. Curr Colorectal Cancer Rep 2017;13:455–69.
7. Nishihara R, Lochhead P, Kuchiba A, et al. Aspirin use and risk of colorectal cancer according to BRAF mutation status. JAMA 2013;309:2563–71.
8. Jarvis D, Mitchell JS, Law PJ, et al. Mendelian randomisation analysis strongly implicates adiposity with risk of developing colorectal cancer. Br J Cancer 2016;115:266–72.
9. Gao C, Patel CJ, Michailidou K, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. Int J Epidemiology 2016;45:896–908.
10. Bhaskaran K, Douglas I, Forbes H, et al. Body mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UKadults. Lancet 2014;384:755–65.
11. Iyengar NM, Gucalp A, Dannenberg AJ, et al. Obesity and cancer mechanisms: tumor microenvironment and inflammation. J Clin Oncol 2016;34:4270–6.
12. Buron Pust A, Alisson R, Blanks R, et al. Heterogeneity of colorectal cancer risk by tumour characteristics: large prospective study of UKwomen. Int J Cancer 2017;140:1082–90.
13. Slattery ML, Anderson K, Curtin K, et al. Lifestyle factors and Ki-ras mutations in colon cancer tumors. Mutat Res 2001;483:73–81.
14. Slattery ML, Curtin K, Wolfik RK, et al. Diet, physical activity, and body size associations with rectal tumor mutations and epigenetic changes. Cancer Causes Control 2003;14:327–45.
15. Brandstedt J, Wangeford S, Nodin B, et al. Associations of anthropometric factors with KRAS and BRAF mutation status of primary colorectal cancer in men and women: a cohort study. PLoS One 2014;9:e89064.
16. Hughes LAE, Williamson EJ, van Engeland M, et al. Body size and risk for colorectal cancers showing BRAF mutations or microsatellite instability: a pooled analysis. Int J Epidemiol 2012;41:1066–72.
17. Hanyuda A, Ogino S, Qian ZR, et al. Body mass index and risk of colorectal cancer according to tumor lymphocytic infiltrate. Int J Cancer 2016;139:854–68.
18. Hoffmeister M, Blaker H, Kloor M, et al. Body mass index and microsatellite instability in colorectal cancer: a population-based study. Cancer Epidemiol Biomarkers Prev 2013;22:2303–11.
19. Inamura K, Song M, Jung S, et al. Prediagnosis plasma Adiponectin in relation to colorectal cancer risk according to KRAS mutation status. J Natl Cancer Inst 2016;108.
20. Kadawaki T, Yamauchi T, Koubota N, et al. Adiponectin and adiponectin receptors in insulin resistance, diabetes, and the metabolic syndrome. J Clin Invest 2006;116:1784–92.
21. Jinjuvadia R, Loizah P, Jinjuvadia C, et al. The association between metabolic syndrome and colorectal neoplasm systematic review and meta-analysis. J Clin Gastroenterol 2013;47:33–44.
22. Stocks T, Lukanova A, Bjorge T, et al. Metabolic factors and the risk of colorectal cancer in 580,000 men and women in the metabolic syndrome and cancer project (me-can). Cancer 2011;117:2398–407.
23. Rodríguez-Broadbent H, Law PJ, Sud A, et al. Mendelian randomization implicates hyperlipidaemia as a risk factor for colorectal cancer. Int J Cancer 2017;140(12):2701–2708.
24. Chen C, Lacke E, Stock C, et al. Colonoscopy and sigmoidoscopy use among older adults in different countries: a systematic review. Prev Med 2017;103:33–42.
25. Norberg M, Wall S, Roman K, et al. The Vasterbotten intervention Programme: background, design and implications. Glob Health Action 2010;3:1–15.
26. Benckert M, Lilja M, Soderberg S, et al. Improved metabolic health among the obese in six population surveys 1986 to 2009: the northern Sweden MONICA study. BMC Obes 2015;2:7.
27. Tohin MD, Sheehan NA, Scarrah KJ, et al. Adjusting for treatment effects in studies of quantititative traits: antihypertensive therapy and systolic blood pressure. Stat Med 2005;24:2911–35.
28. Wu J, Province MA, Coon H, et al. An investigation of the effects of lipid-lowering medications: genome-wide linkage analysis of lipids in the HyperGEN study. BMC Genet 2007;8:60.
29. Ekelöf V, Wikberg ML, Edin S, et al. The prognostic role of KRAS, BRAF, Pik3ca and Pten in colorectal cancer. Br J Cancer 2013;108:2153–63.
30. Benlbôch S, Paya A, Alenda C, et al. Detection of BRAF V600E mutation in colorectal cancer: comparison of automatic sequencing and real-time chemistry methodology. J Mol Diagn 2006;8:540–3.
31. Van Guelpen B, Dahlin AM, Hultdin J, et al. One-carbon metabolism and Cpx Island methylator phenotype status in incident colorectal cancer: a nested case-referent study. Cancer Causes Control 2010;21:557–66.
32. van Buuren S, mice KG-O. Multivariate imputation by chained equations in R. J Stat Softw 2011;45:67.
33. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011;30:377–99.
34. Johansson I, Hallmans G, Wikman A, et al. Validation and calibration of food-frequency questionnaire measurements in the northern Swedish health and disease cohort. Public Health Nutr 2002;5:847–96.
35. Marshall A, Altman DG, Helder RL, et al. Combining estimates of interest in prognostic modeling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol 2009;9:57.
36. Liu L, Nevo D, Nishihara R, et al. Utility of inverse probability weighting in molecular pathological epidemiology. Eur J Epidemiol 2017;33(4):381–392.
37. Wang M, Spiegelman D, Kuchiba A, et al. Statistical methods for studying disease subtype heterogeneity. Stat Med 2016;35:782–800.
38. Benjamin DJ, Berger JO, Johannesson M, et al. Redefine statistical significance. Nat Hum Behav 2018;2:6–10.
39. Engholm G, Ferlay J, Christensen N, et al. NORDCAN—a Nordic tool for cancer information, planning, quality control and research. Acta Oncol 2010;49:725–36.
40. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology 2010;138:2088–100.
41. Jayasekara H, Maclonis RJ, Williamson EJ, et al. Lifetime alcohol intake is associated with an increased risk of KRAS plus BRAF/–KRAS but not BRAF plus colorectal cancer. Int J Cancer 2017;140:1485–93.
42. Jiang Y, Ben Q, Shen H, et al. Diabetes mellitus and incidence and mortality of colorectal cancer: a systematic review and meta-analysis of cohort studies. Eur J Epidemiol 2011;26:683–76.
43. Xu JM, Ye Y, Wu H, et al. Association between markers of glucose metabolism and risk of colorectal cancer. BMJ Open 2016;6(6):e011430.
44. Ouchi N, Walsh K. Adiponectin as an anti-inflammatory factor. Clin Chim Acta 2007;380:24–30.
45. Kaitaaina S, Thummalapallni R, Barbie DA. Inflammation as a driver and vulnerability of KRAS mediated oncogenesis. Semin Cell Dev Biol 2016;58:127–35.
46. Kuo SM, Halpert MM. Lack of association between body mass index and plasma adiponectin levels in healthy adults. Int J Obesity 2011;35:1487–94.
47. Turer AT, Khera A, Ayers CR, et al. Adipose tissue mass and location affect circulating adiponectin levels. Diabetologia 2011;54:2515–24.
48. Sorbye H, Dragomir A, Sundström M, et al. High BRAF mutation frequency and marked survival
differences in subgroups according to KRAS/BRAN mutation status and tumor tissue availability in a prospective population-based metastatic colorectal cancer cohort. PLoS One 2015;10:e0131046.

49. Richiardi L, Barone-Adesi F, Pearce N. Cancer subtypes in aetiological research. Eur J Epidemiol 2017;32:353–61.

50. Haring R, Rosvall M, Volker U, et al. A network-based approach to visualize prevalence and progression of metabolic syndrome components. PLoS One 2012;7:e39461.

51. Stefan N, Haring HU, Hu FB, et al. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. Lancet Diabetes Endocrinol 2013;1:152–62.

52. Barlow L, Westergren K, Holmberg L, et al. The completeness of the Swedish cancer register: a sample survey for year 1998. Acta Oncol 2009;48:27–33.

53. Kudryavtseva AV, Lipatova AV, Zaretsky AR, et al. Important molecular genetic markers of colorectal cancer. Oncotarget 2016;7:53959–83.