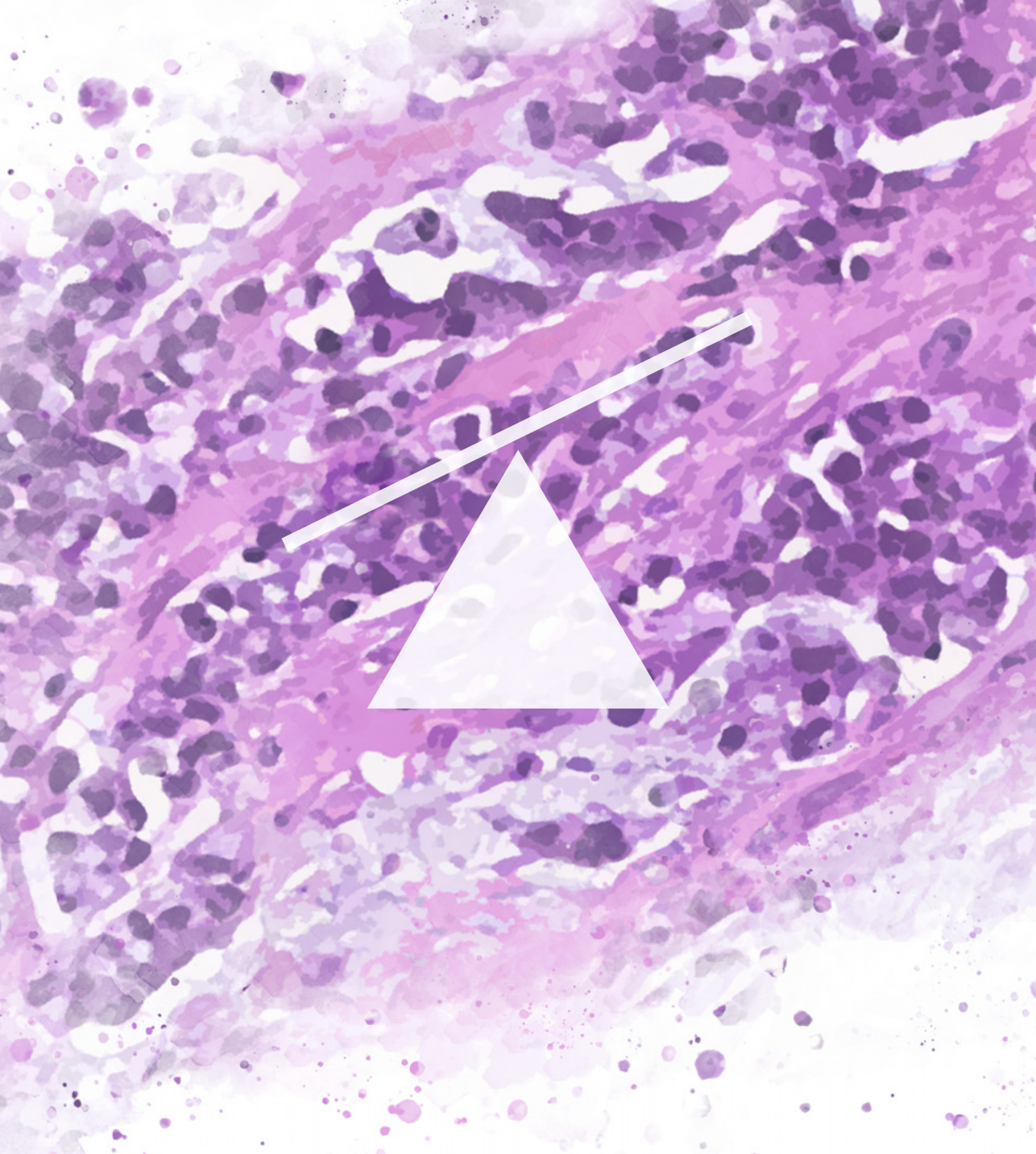
Energy balance and colorectal cancer risk



ENERGY BALANCE AND COLORECTAL CANCER RISK

A ROLE FOR CANCER CELL METABOLISM?

JOSIEN JENNISKENS

ENERGY BALANCE AND COLORECTAL CANCER RISK

A ROLE FOR CANCER CELL METABOLISM?

Energy balance and colorectal cancer risk: a role for cancer cell metabolism?

ISBN: 978-94-6458-564-3

Cover design and layout: Josien Jenniskens
Printing: Ridderprint | www.ridderprint.nl

The work presented in this thesis was performed within GROW, School for Oncology and Reproduction at the Department of Epidemiology, Maastricht University, the Netherlands.

This work was financially supported by the Dutch Cancer Society (KWF 11044).

ENERGY BALANCE AND COLORECTAL CANCER RISK

A ROLE FOR CANCER CELL METABOLISM?

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. dr. P. Habibović
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op vrijdag 4 november 2022 om 13.00 uur

door

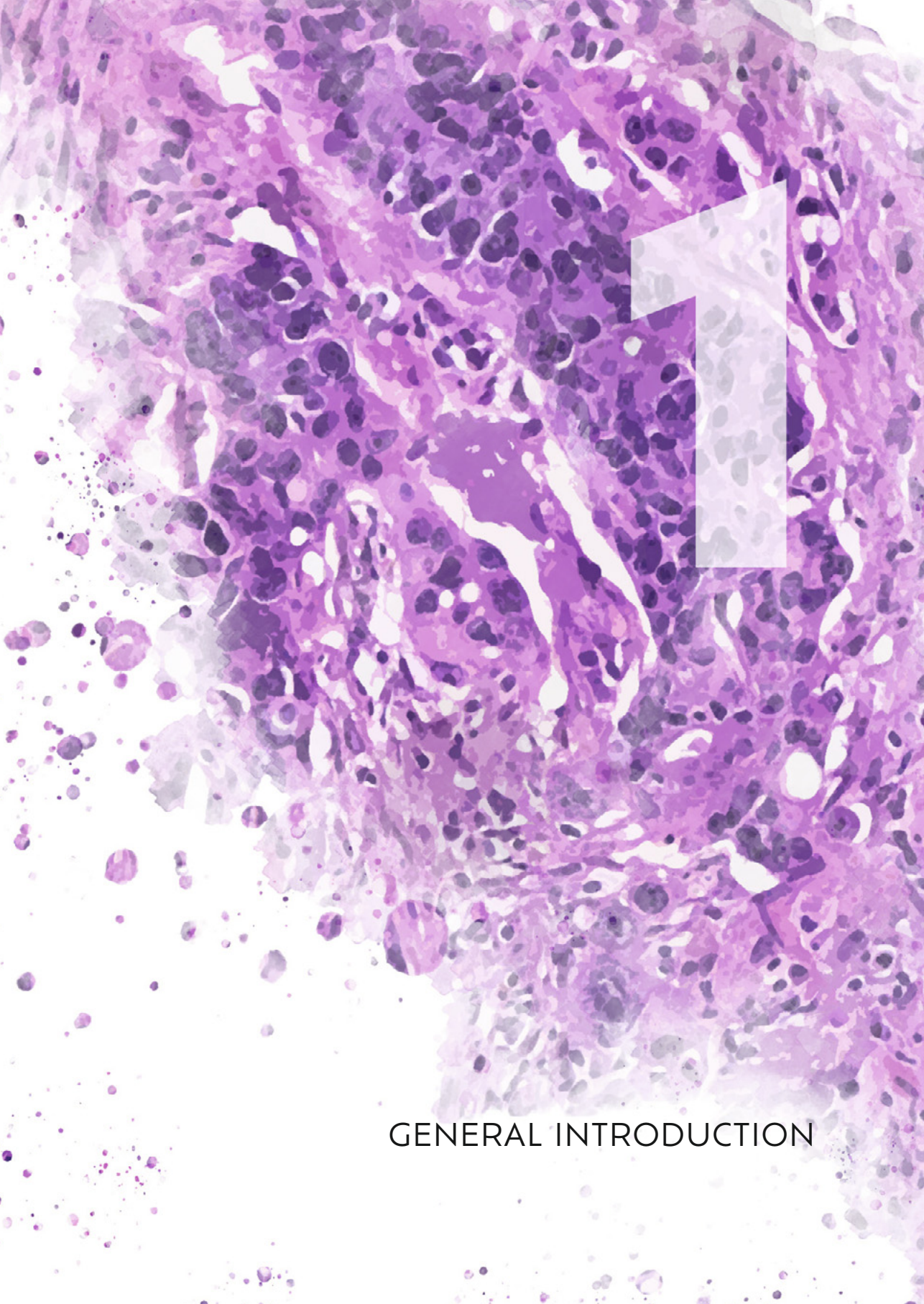
Josien Christina Anna Jenniskens

Promotores
Prof. dr. ir. P.A. van den Brandt
Prof. dr. H.I. Grabsch

Copromotor
Dr. C.C.J.M. Simons

Beoordelingscommissie
Prof. dr. I.C.W. Arts (voorzitter)
Dr. F.J.B. van Duijnhoven (Wageningen University & Research, the Netherlands)
Dr. M. Loughrey (Queen's University Belfast, Northern Ireland, United Kingdom)
Prof. dr. A.A.M. Masclee
Prof. dr. F.J. van Schooten

CONTENTS



1

GENERAL INTRODUCTION

COLORECTAL CANCER

Colorectal cancer has been amongst the five most common cancers worldwide for decades (Figure 1)[1-4]. The total number of new colorectal cancer cases in 2020 was estimated to be 1.9 million worldwide, of which around 1.15 million were colon and 0.73 million were rectal cancers[4]. Cases were more often male than female, with 0.6 million new colon cancer cases and 0.4 million rectal cancer cases in men compared to 0.5 million colon cancer cases and 0.3 million rectal cancer cases in women.

Colorectal cancers originate from the colon or rectum, but it currently is a matter of discussion whether these should be considered a single disease, or rather distinct entities. Even though they are similar in morphology, differences between colon and rectal cancer exist regarding embryological origin, risk factors, genetic mechanisms, and prognosis[5-8]. Figure 2 provides an overview of the various anatomical sites in colorectal cancer with the distribution of incidence per site, separately for men and women.

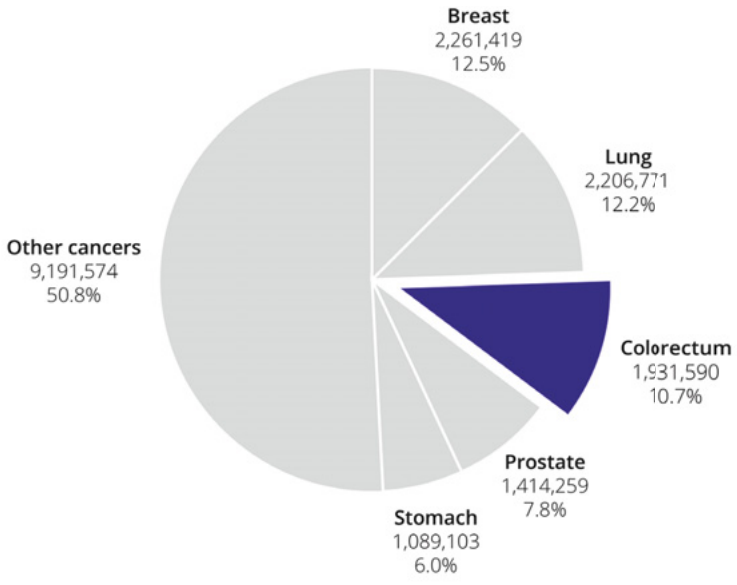


Figure 1 | Distribution of the estimated number of new cancer cases in 2020 (worldwide, both sexes, all ages, non-melanoma skin cancers excluded). *Source: GLOBOCAN 2020, accessed April 2022.*

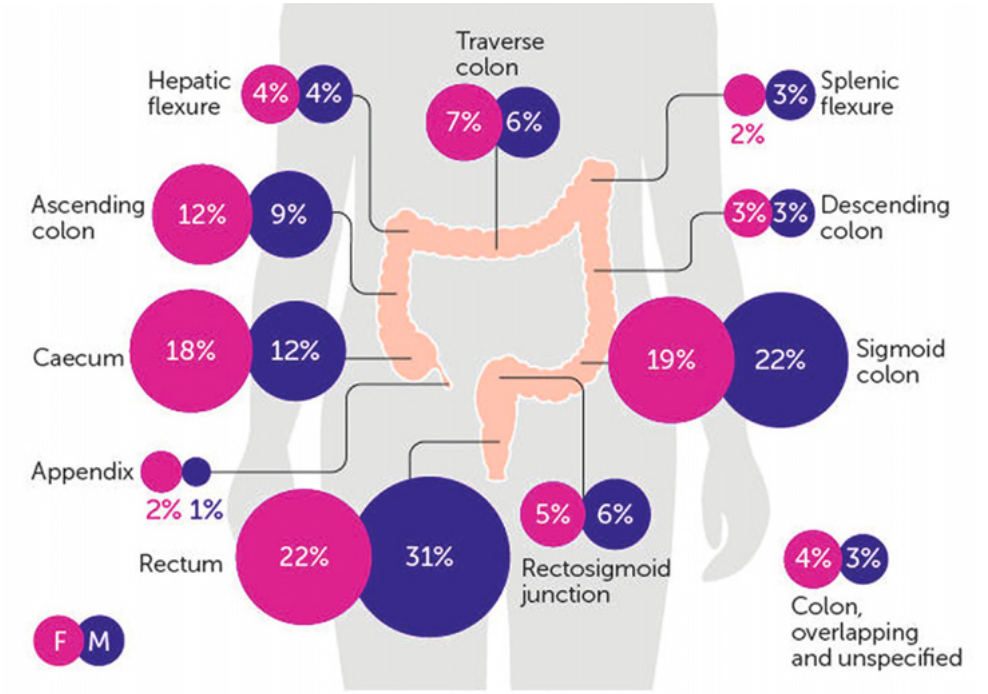


Figure 2 | Percentage distribution of colorectal cancer cases by anatomical site, separately for men (purple) and women (pink). Numbers are based on the United Kingdom (2016-2018). *Source: Cancer Research UK, accessed April 2022.*

In addition to heterogeneity in anatomical location of colorectal cancers, the molecular profile has been shown to vary across tumors[9, 10]. Vogelstein et al[11, 12] developed a model for the transformation from normal colon mucosa to colorectal cancer, referred to as the adenoma-carcinoma sequence (Figure 3). This sequence starts with the transformation of normal colorectal epithelium to a benign epithelial proliferation (i.e. adenoma), which progresses through the accumulation of multiple genetic and epigenetic alterations, ultimately leading to invasive and metastasizing tumors[11-13]. One of the major pathways associated with colorectal cancer is microsatellite instability (MSI), which could happen early in the process due to loss of DNA mismatch repair activity[14]. *APC* mutations seem to occur early in the transition from normal mucosa to early adenoma. Additional mutations seem to be required for the progression to intermediate adenoma, and especially mutations in the *KRAS* or *BRAF* oncogenes are often observed at this stage. Lastly, progression from adenoma to cancer seems to involve mutations in the *TP53* tumor suppressor gene[11-13].

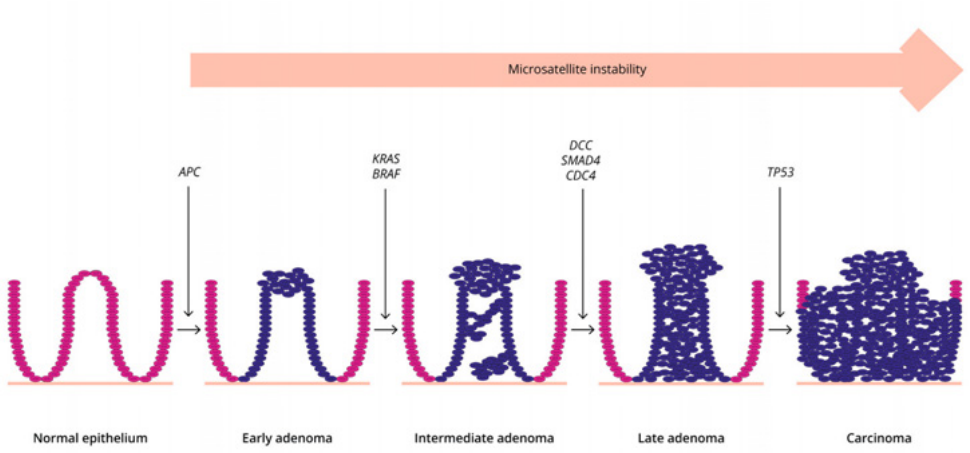


Figure 3 | Simplified illustration of the colorectal adenoma-carcinoma sequence. *Based on figures from Davies[15] and Nguyen[13].*

CANCER CELL METABOLISM

In 2000, Hanahan and Weinberg described six biological capabilities that are acquired during the multistep transformation of normal cells into malignant cells, referred to as the hallmarks of cancer[16]. These were updated in 2011[17] and 2022[18], resulting in a total of 14 hallmarks. The hallmark on metabolic reprogramming, which was added in 2011, has gained interest during the last decades, with a steep increase in the number of publications in the new millennium[19].

THE WARBURG-EFFECT

The increased metabolic demand of cancer cells caused by, amongst others, the enhanced proliferation rate requires adjustments in energy metabolism[17]. Under aerobic conditions (i.e. sufficient oxygen available), normal cells convert glucose into pyruvate via glycolysis. The pyruvate is then shuttled into the mitochondria where it is converted into carbon dioxide via the tricarboxylic acid (TCA) cycle, representing complete oxidation. Under anaerobic conditions (i.e. insufficient oxygen available), the influx of pyruvate into the oxygen-consuming TCA cycle is decreased. In the 1920s, Otto Warburg and colleagues observed that cancer cells, even under aerobic conditions, can show a similar metabolic phenotype as normal cells under anaerobic conditions[20]. This metabolic phenotype in cancer cells is often referred to as the "Warburg-effect" or "aerobic glycolysis" (Figure 4a).

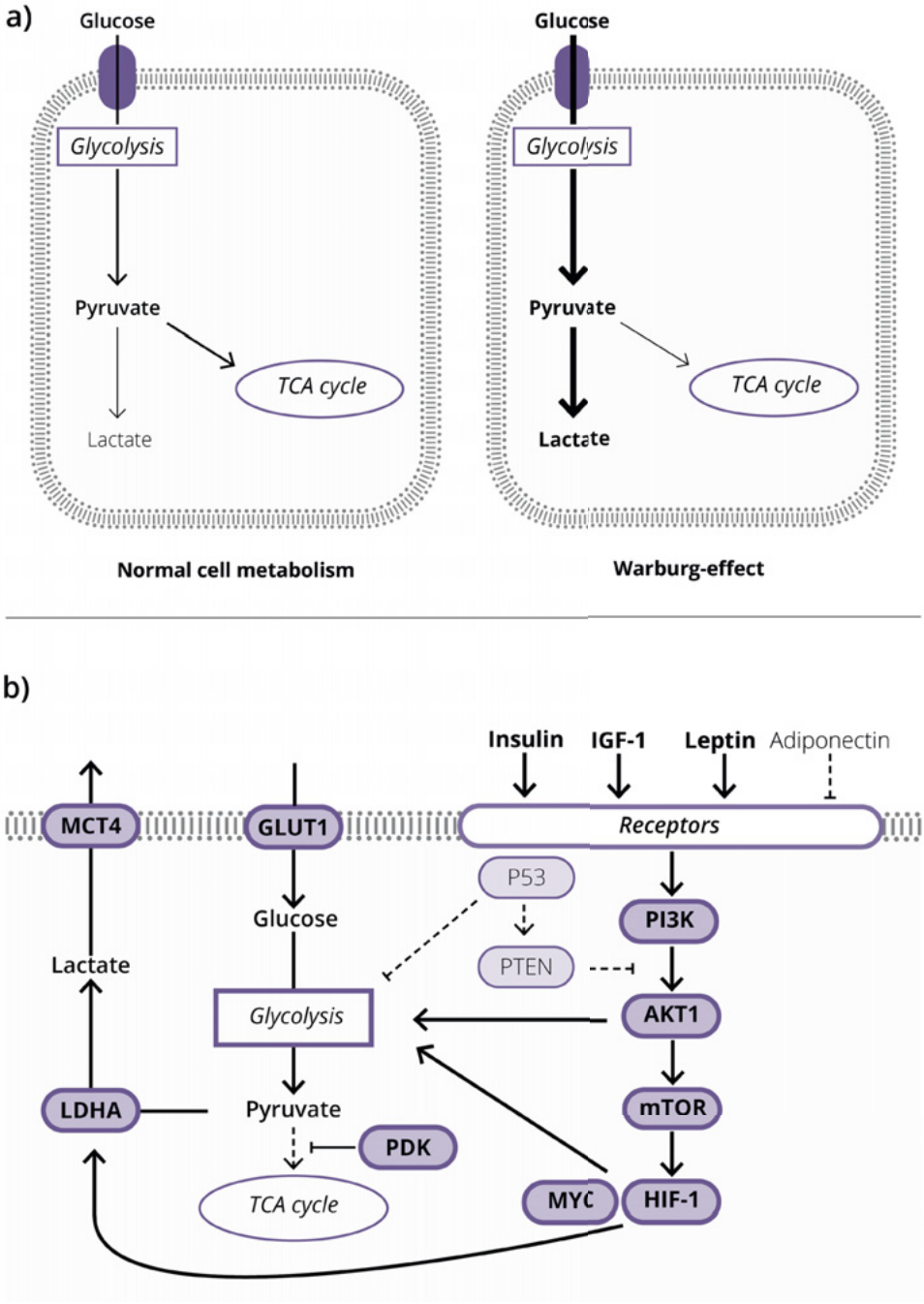


Figure 4 | Simplified schematic overview of the Warburg-effect. **a)** The left cell represents normal cell metabolism (complete oxidation); the right cell represents the Warburg-effect (aerobic glycolysis). *Based on van der Heiden*[26]; **b)** A more detailed overview of the Warburg-effect with important proteins and drivers. An arrow (↑) indicates a stimulatory effect; an arrow with a blunt head (T) indicates an inhibitory effect; a dashed line indicates loss of function. Note: a line does not always indicate a direct link; steps may have been skipped here for simplicity. *Based on Cairns*[27] *and others*[25, 28-32].

The Warburg-effect is an energetically inefficient metabolic phenotype, with an ~18 fold lower adenosine triphosphate (ATP; cellular energy carrier) yield per glucose molecule as compared to complete oxidation[17]. In order to compensate for the lower efficiency of ATP production, cancer cells adapt the expression of several glycolysis related proteins[21-25] (Figure 4b). Glucose import is increased by upregulating glucose transporter 1 (GLUT1) expression. Subsequently, glycolysis is enhanced by increased expression of enzymes like hexokinase 2 (HK2) and pyruvate dehydrogenase kinase 1 (PDK1). Pyruvate kinase M2 (PKM2) catalyzes the final step of the glycolytic pathway. For conversion of the accumulating pyruvate into lactate, the expression of the enzyme lactate dehydrogenase A (LDHA) is upregulated. Finally, to prevent cytoplasmic acidification, expression of monocarboxylate transporter 4 (MCT4) is upregulated for lactate export.

DRIVERS OF THE WARBURG-EFFECT

The phosphoinositide 3-kinase (PI3K)/Akt signaling pathway has been shown to be a major regulator of metabolic reprogramming towards the Warburg-effect[23, 27, 28, 33, 34] (Figure 4b). This pathway is one of the most commonly altered pathways in human cancers and is activated by mutations in oncogenes or tumor suppressor genes or by aberrant signaling from receptors[27, 35]. PI3K can be activated by mutations in the oncogenes phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*) and/or Kirsten rat sarcoma viral oncogene homolog (*KRAS*)[31, 36, 37]. An effector downstream of PI3K is AKT serine/threonine-protein kinase 1 (AKT1), which has been shown to stimulate glycolysis by increasing the expression of glucose transporters and phosphorylating key glycolytic enzymes[38, 39]. Subsequently, AKT1 increases mammalian target of rapamycin (mTOR) signaling[40], which causes metabolic changes by activation of the transcription factor hypoxia inducible factor 1 (HIF-1)[27]. HIF-1 is an important driver of the Warburg-effect by upregulating the transcription of glycolytic enzymes and downregulating mitochondrial oxidative metabolism[41-43]. Furthermore, HIF-1 has been shown to act together with the oncogenic transcription factor MYC[44-46]. It has been shown that both MYC and HIF-1 activity can be upregulated by mutations in well-known oncogenes *KRAS* and/or v-raf murine sarcoma viral oncogene homolog B1 (*BRAF*)[29, 30]. Furthermore, the abovementioned enzyme PKM2 seems to have a regulatory function in the Warburg-effect by functioning as a transcriptional coactivator of HIF-1, too[25, 32]. Lastly, normal functioning of the tumor suppressor P53 inhibits PI3K/Akt signaling through activation of the tumor suppressor phosphatase and tensin homolog (PTEN)[47]. In addition, normal functioning of P53 indirectly inhibits glycolysis via TP53 induced glycolysis regulatory phosphatase (TIGAR)[48]. However, expression of (one of) these tumor suppressors and thereby their inhibitory function is often lost in colorectal cancer (Figure 4b)[27].

ENERGY BALANCE AND COLORECTAL CANCER

It has been estimated that more than 50% of colorectal cancer cases are attributable to modifiable risk factors, and thus potentially preventable[49]. Frequently studied risk factors for colorectal cancer are overweight or obesity, physical inactivity, adult-attained height, alcohol intake, smoking, and several dietary factors[50]. The majority of these risk factors are indicators of energy balance. Energy balance is the state at which the number of calories consumed equals the number of calories used. Energy intake depends on the caloric intake via food and drinks, whereas energy expenditure depends on physical activity and the resting metabolic rate (i.e. the amount of energy needed to fuel the body at rest)[51]. A positive energy balance occurs when energy intake is larger than energy expenditure, which causes an increase in body mass which can ultimately result in overweight or even obesity. Factors related to energy balance (hereafter referred to as energy balance-related factors) both early in life and later in life, have been associated with risk of developing colorectal cancer[52-63].

EPIDEMIOLOGICAL EVIDENCE

The World Cancer Research Fund (WCRF) and the International Agency for Research on Cancer (IARC) regularly publish updated reports on scientific evidence for cancer prevention[64, 65]. In these reports, the evidence for associations is ranked according to probability of a causal relationship. Both reports indicate convincing evidence for associations of adiposity, height, and physical activity with colorectal cancer risk[64, 65]. Limited evidence exists on energy intake in relation to colorectal cancer risk[64].

However, while many studies reported associations between these energy balance-related factors and colorectal cancer, associations often appear to be weak or are inconsistent between studies. For example, a recent meta-analysis based on prospective studies reported statistically significant summary relative risks (RR) of 1.04 per 5 cm increase in height, of 1.06 per 5 kg/m^2 increase in body mass index (BMI), and of 1.02 per 10 cm increase in waist circumference[52]. Furthermore, associations of energy balance-related factors with risk of colorectal cancer have repeatedly been observed to vary across studies and/or populations as well as according to tumor location and/or sex[7, 52, 59-61, 63].

ETIOLOGICAL PATHWAY

Up till now, the mechanism(s) behind the colorectal cancer risk enhancement by disrupted energy balance remain(s) to be elucidated. However, three main factors have been proposed[66-68]: adipokines (e.g. leptin and adiponectin), insulin and insulin-like growth factor 1 (IGF-1) signaling, and sex hormones. Circulating levels of leptin, adiponectin, insulin, and IGF-1 have been shown to be influenced by energy balance-related factors[69-74]. In turn, aberrant signaling of the abovementioned factors has been associated with colorectal cancer risk[75-78]. A common downstream effect of leptin, adiponectin, insulin, and IGF-1 signaling is regulating intracellular PI3K/Akt signaling[28]. As described above, this pathway has been shown to be associated with metabolic reprogramming towards the Warburg-effect.

MOLECULAR PATHOLOGICAL EPIDEMIOLOGY

Molecular pathological epidemiology (MPE) research can provide further insights into carcinogenic mechanisms underlying etiological associations of dietary, lifestyle, or environmental factors with disease risk[79]. MPE is an emerging transdisciplinary field that was introduced by Ogino and Stampfer in 2010[79], incorporating molecular pathology into epidemiological research. In traditional epidemiology, applied to the topic of this thesis, energy balance would be investigated in relation to colorectal cancer risk (Figure 5a), whereas (molecular) pathological research is used to estimate presence of the Warburg-effect in the tumor and predict prognosis or response to treatment (Figure 5b). When these two disciplines are combined, it can be investigated whether energy balance is differentially associated with risk of colorectal cancer subgroups based on estimated presence of the Warburg-effect (Figure 5c). Furthermore, prognosis or response to treatment can be investigated in relation to lifestyle or environmental factors, but this was not considered in the current thesis.

Classification of tumors based on molecular characteristics in MPE research leads to more homogenous tumor subgroups, thereby addressing disease heterogeneity[79-81]. It can then be explored whether certain lifestyle or environmental factors, such as energy balance-related factors, are differentially associated with subgroups of the disease. Hereby, MPE research can provide further insights into etiology and pathogenesis of diseases, strengthening evidence for causal relationships. In addition, by investigating subgroups of a heterogeneous disease instead of investigating the disease as a single entity, weak or masked associations can be revealed.

A challenge of MPE research is misclassification in the categorization of a disease into subtypes[80]. A common way of subtyping is by using immunohistochemical (IHC) staining to assess protein expression. Since the number of cases in MPE studies needs to be very large, it is very time-consuming for a pathologist to score all the available material. Therefore, scoring of the IHC stained tissue is often performed by non-pathologists[82, 83]. However, studies on the validity of scoring results from non-pathologists are currently limited.

AIMS AND HYPOTHESES

In the current thesis we aimed to investigate the potential involvement of the Warburg-effect in the etiological pathway between energy balance and colorectal cancer risk using an MPE approach in a population based cohort study. To achieve this, we investigated associations of adult energy balance-related factors (i.e., BMI; lower body clothing size, as a proxy for waist circumference; physical activity) as well as early-life energy balance-related factors (i.e. height; energy restriction proxies of exposure to the Dutch Hunger Winter, World War II, and the Dutch Economic

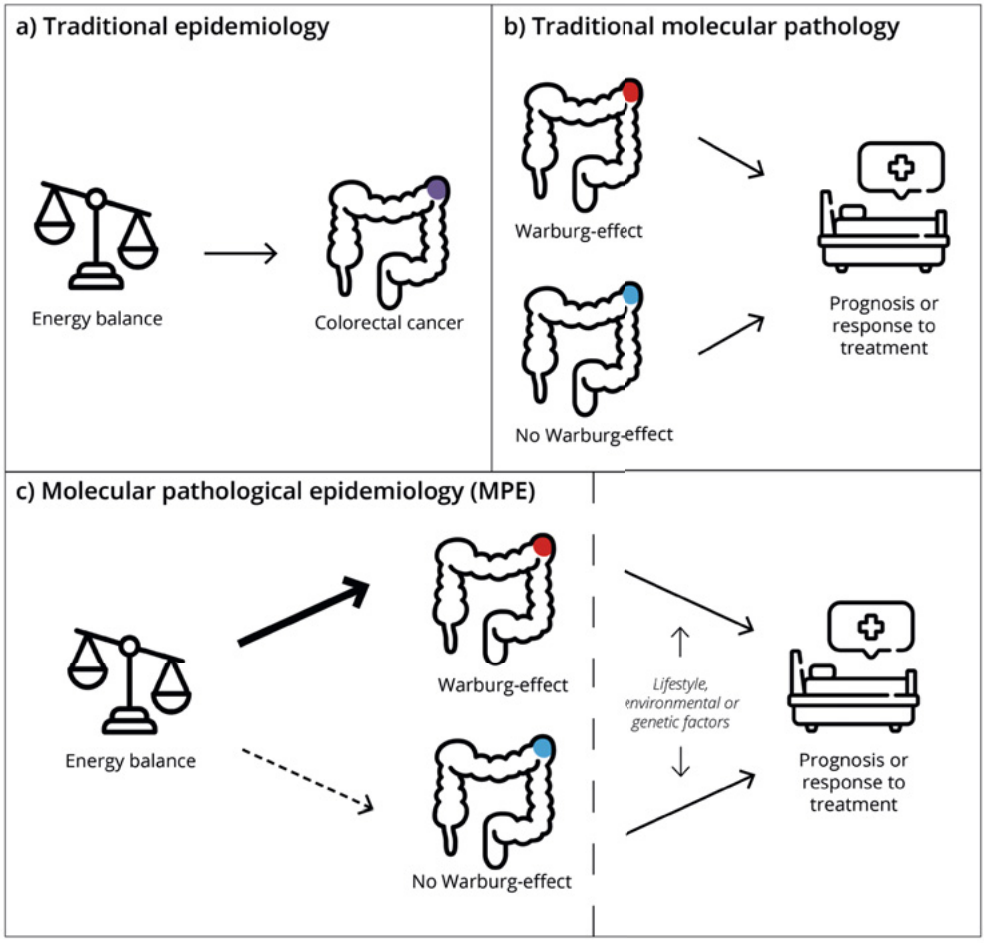


Figure 5 | Schematic overview of **a)** traditional epidemiology; **b)** traditional molecular pathology; **c)** molecular pathological epidemiology. Note: this scheme was based on the current thesis but can be applied to any exposure-disease (subtype) association. The right part of panel c was not considered in the current thesis. *Adapted from Ogino*[80].

Depression; adolescent BMI) with colorectal cancer risk in relation to the estimated presence of the Warburg-effect in the tumor. The presence of the Warburg-effect was estimated by establishing Warburg-subtypes based on IHC expression of six proteins involved in different levels of the Warburg-effect (upstream regulation of the Warburg-effect: PTEN, P53; glucose import: GLUT1; glycolysis: PKM2; conversion of pyruvate into lactate: LDHA; lactate secretion: MCT4). Furthermore, we investigated subgroups of colorectal cancer based on mutations in oncogenes that have been associated with the upstream regulation of the Warburg-effect (*KRAS*, *PIK3CA*, and *BRAF* mutations) as well as on mismatch repair (MMR) status, as a surrogate marker for microsatellite instability (MSI). In addition, we aimed to investigate whether non-pathologists can generate valid and reproducible IHC scoring results.

We hypothesized that:

- Associations between energy balance-related factors and colorectal cancer risk differ across subtypes based on IHC expression of six proteins (LDHA, GLUT1, MCT4, PKM2, P53, PTEN) involved in the Warburg-effect.

- Associations between energy balance-related factors and colorectal cancer risk differ across subgroups based on molecular features (*KRAS*, *PIK3CA*, *BRAF* mutation status and MMR status) that have previously been associated with the upstream regulation of the Warburg-effect.

In addition, we hypothesized that non-pathologists can produce valid and reproducible IHC scoring results, similar to those of an experienced pathologist.

STUDY FRAMEWORK: THE NLCS

All studies included in this thesis were performed using data from the Netherlands Cohort Study (NLCS) on diet and cancer[84]. This is a large prospective cohort study initiated in 1986, which included 120,852 subjects aged 55-69 years at baseline. All participants completed a mailed, self-administered questionnaire on cancer risk factors. This questionnaire provided information on the following energy balance-related factors:

- Adult BMI, from self-reported weight at baseline in kg divided by height in meters squared

- Adult clothing size, as a proxy for waist circumference[85]

- Non-occupational and occupational physical activity

- Adolescent BMI (age 20 years), from self-reported weight at 20 years in kg divided by height in meters squared

- Adult-attained height

- Early-life energy restriction using three proxy measures: place of residence during the Dutch Hunger Winter (1944-45), place of residence during World War II (1940-44), and employment status of the father during the Dutch Economic Depression (1932-40)

A case-cohort approach was applied to improve efficiency of data collection and processing. Cancer cases were derived from the entire cohort, whereas person-years at risk for the total cohort were estimated from a subcohort. The subcohort, including 5000 participants, was randomly sampled from the entire cohort at baseline, and biennially followed-up for vital status information. Incident cancer cases were identified from the total cohort through annual record linkage with the Netherlands Cancer Registry and PALGA, the nationwide Dutch Pathology Registry[86]. In this thesis, we used data from 20.3 years of follow-up (September 17, 1986 until January 1, 2007), including 4,597 incident colorectal cancer cases.

In 2012, the Rainbow-TMA project was initiated, aiming to build a comprehensive biobank of tissue microarrays (TMAs)[87]. For this, formalin-fixed paraffin-embedded (FFPE) tissue blocks from colorectal cancer resections and matched normal tissue from incident cases were requested from participating laboratories. Tissue blocks were successfully collected from 43 pathology laboratories throughout the Netherlands. For TMA-construction, pathologists reviewed scanned Haematoxylin & Eosin (H&E)-stained sections and identified areas with the highest tumor density, from which three 0.6mm diameter cores were sampled per case along with three normal tissue cores (TMA-Grandmaster, 3D-Histech, Hungary).

THESIS OUTLINE

After introducing the topic of this thesis in the current chapter, the validity and reproducibility of IHC scoring by trained non-pathologists on TMAs is described in Chapter 2. Associations of adult and early-life energy balance-related factors with risk of Warburg-subtypes in colorectal cancer are described in Chapters 3 and 4, respectively. In addition, associations of adult and early-life energy balance-related factors with risk of colorectal cancer subgroups based on *KRAS*, *PIK3CA* and *BRAF* mutations and MMR deficiency are described in Chapters 5 and 6, respectively. Lastly, Chapter 7 provides a discussion of the main findings described in this thesis.

REFERENCES

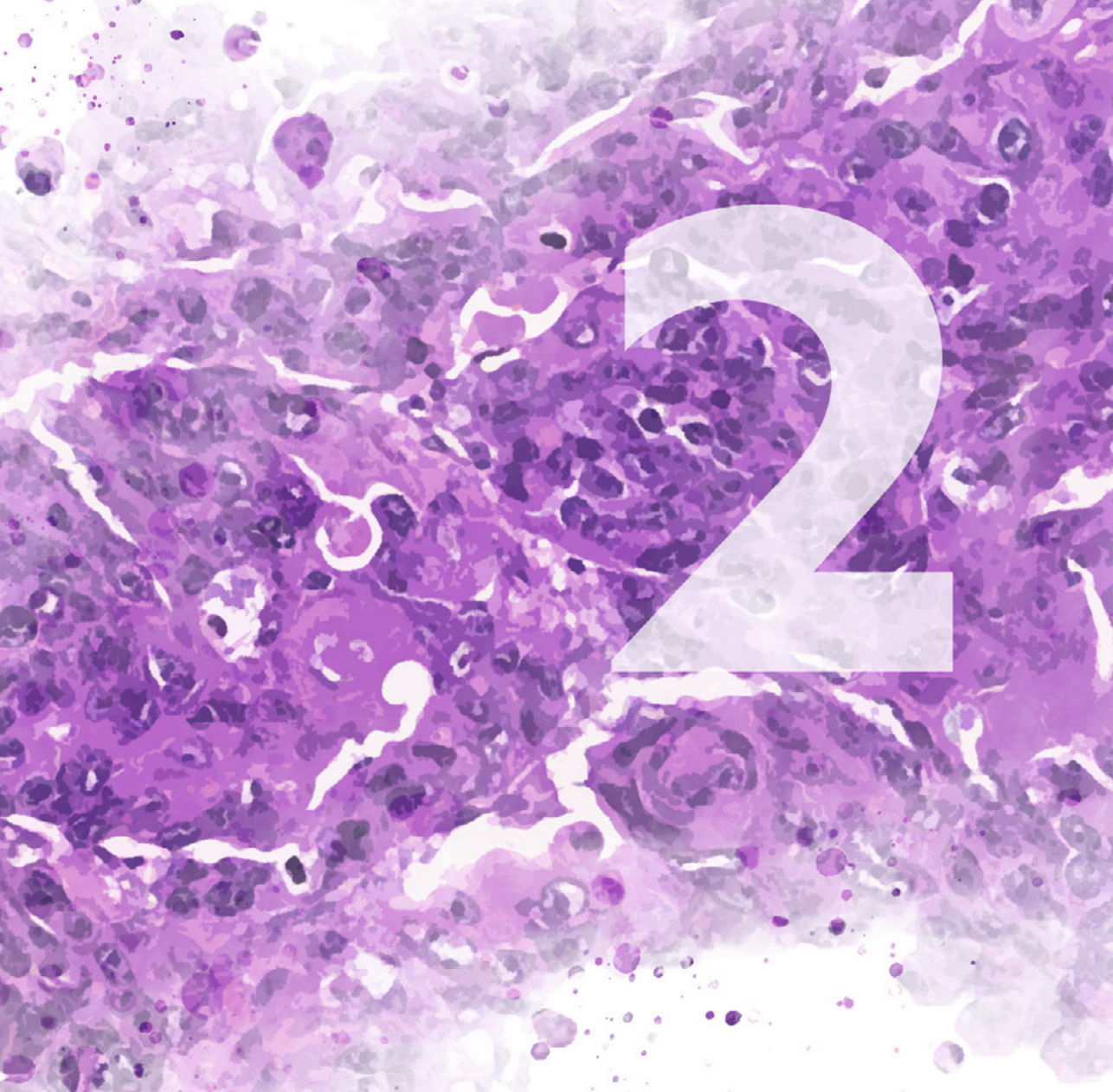
1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer. 2010;127(12):2893-917.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015;65(2):87-108.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
4. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2021;71(3):209-49.
5. Li M, Li J, Zhao A, Gu J. Colorectal cancer or colon and rectal cancer? Oncology. 2007;73(1-2):52-7.
6. Kapiteijn E, Liefers G, Los L, Klein Kranenbarg E, Hermans J, Tollenaar R, Moriya Y, Van De Velde C, Van Krieken J. Mechanisms of oncogenesis in colon versus rectal cancer. The Journal of pathology. 2001;195(2):171-8.
7. Wei EK, Giovannucci E, Wu K, Rosner B, Fuchs CS, Willett WC, Colditz GA. Comparison of risk factors for colon and rectal cancer. International journal of cancer. 2004;108(3):433-42.
8. Li F-y, Lai M-d. Colorectal cancer, one entity or three. Journal of Zhejiang University Science B. 2009;10(3):219-29.
9. Budinska E, Popovici V, Tejpar S, D'Ario G, Lapique N, Sikora KO, Di Narzo AF, Yan P, Hodgson JG, Weinrich S. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. The Journal of pathology. 2013;231(1):63-76.
10. Ogino S, Goel A. Molecular classification and correlates in colorectal cancer. The Journal of Molecular Diagnostics. 2008;10(1):13-27.
11. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.
12. Cho KR, Vogelstein B. Genetic alterations in the adenoma–carcinoma sequence. Cancer. 1992;70(S4):1727-31.
13. Nguyen HT, Duong HQ. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. Oncology Letters. 2018;16(1):9-18.
14. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology. 2010;138(6):2073-87. e3.
15. Davies RJ, Miller R, Coleman N. Colorectal cancer screening: prospects for molecular stool analysis. Nat Rev Cancer. 2005;5(3):199-209.
16. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
17. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. cell. 2011;144(5):646-74.
18. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022;12(1):31-46.
19. Clarivate Web of Science 2022 [Available from: www.webofscience.com.
20. Warburg O. The metabolism of carcinoma cells. The Journal of Cancer Research. 1925;9(1):148-63.
21. Koukourakis MI, Giatromanolaki A, Harris AL, Sivridis E. Comparison of metabolic pathways between cancer cells and stromal cells in colorectal carcinomas: a metabolic survival role for tumor-associated stroma. Cancer research. 2006;66(2):632-7.
22. Kato Y, Maeda T, Suzuki A, Baba Y. Cancer metabolism: new insights into classic characteristics. Jpn Dent Sci Rev. 2018;54(1):8-21.
23. Cairns RA. Drivers of the Warburg phenotype

The Cancer Journal. 2015;21(2):56-61.
24. Vaupel P, Multhoff G. Revisiting the Warburg effect: historical dogma versus current understanding. The Journal of Physiology. 2021;599(6):1745-57.
25. Bensinger SJ, Christofk HR, editors. New aspects of the Warburg effect in cancer cell biology. Semin Cell Dev Biol; 2012: Elsevier.
26. van der Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. science. 2009;324(5930):1029-33.
27. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nature Reviews Cancer. 2011;11(2):85.
28. Huang XF, Chen JZ. Obesity, the PI3K/Akt signal pathway and colon cancer. Obesity reviews. 2009;10(6):610-6.
29. Kimmelman AC. Metabolic dependencies in RAS-driven cancers. Clinical Cancer Research. 2015;21(8):1828-34.
30. Papa S, Choy PM, Bubici C. The ERK and JNK pathways in the regulation of metabolic reprogramming. Oncogene. 2019;38(13):2223-40.
31. Rajalingam K, Schreck R, Rapp UR, Albert Š. Ras oncogenes and their downstream targets. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 2007;1773(8):1177-95.
32. Luo W, Hu H, Chang R, Zhong J, Knabel M, O'Meally R, Cole RN, Pandey A, Semenza GL. Pyruvate kinase M2 is a PHD3-stimulated coactivator for hypoxia-inducible factor 1. Cell. 2011;145(5):732-44.
33. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330(6009):1340-4.
34. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiotherapy and oncology. 2009;92(3):329-33.
35. Wong K-K, Engelman JA, Cantley LC. Targeting the PI3K signaling pathway in cancer. Curr Opin Genet Dev. 2010;20(1):87-90.
36. Jiang W, He T, Liu S, Zheng Y, Xiang L, Pei X, Wang Z, Yang H. The PIK3CA E542K and E545K mutations promote glycolysis and proliferation via induction of the β-catenin/SIRT3 signaling pathway in cervical cancer. Journal of hematology & oncology. 2018;11(1):1-15.
37. Huw L, O'brien C, Pandita A, Mohan S, Spoerke J, Lu S, Wang Y, Hampton G, Wilson T, Lackner M. Acquired PIK3CA amplification causes resistance to selective phosphoinositide 3-kinase inhibitors in breast cancer. Oncogenesis. 2013;2(12):e83-e.
38. Elstrom RL, Bauer DE, Buzzai M, Karnauskas R, Harris MH, Plas DR, Zhuang H, Cinalli RM, Alavi A, Rudin CM. Akt stimulates aerobic glycolysis in cancer cells. Cancer research. 2004;64(11):3892-9.
39. Fan Y, Dickman KG, Zong W-X. Akt and c-Myc differentially activate cellular metabolic programs and prime cells to bioenergetic inhibition. J Biol Chem. 2010;285(10):7324-33.
40. Robey RB, Hay N, editors. Is Akt the "Warburg kinase"?—Akt-energy metabolism interactions and oncogenesis. Semin Cancer Biol; 2009: Elsevier.
41. Semenza GL. Targeting HIF-1 for cancer therapy. Nature reviews cancer. 2003;3(10):721-32.
42. Semenza GL. HIF-1: upstream and downstream of cancer metabolism. Current opinion in genetics & development. 2010;20(1):51-6.
43. Stubbs M, Griffiths JR. The altered metabolism of tumors: HIF-1 and its role in the Warburg effect. Advances in enzyme regulation. 2010;50(1):44-55.
44. Dang CV, Kim J-w, Gao P, Yustein J. The interplay between MYC and HIF in cancer. Nature Reviews Cancer. 2008;8(1):51-6.
45. Dang CV, Le A, Gao P. MYC-induced cancer cell energy metabolism and therapeutic opportunities. Clinical cancer research. 2009;15(21):6479-83.
46. Kim J-w, Gao P, Liu Y-C, Semenza GL,

Dang CV. Hypoxia-inducible factor 1 and dysregulated c-Myc cooperatively induce vascular endothelial growth factor and metabolic switches hexokinase 2 and pyruvate dehydrogenase kinase 1. Mol Cell Biol. 2007;27(21):7381-93.
47. Stambolic V, MacPherson D, Sas D, Lin Y, Snow B, Jang Y, Benchimol S, Mak T. Regulation of PTEN transcription by p53. Mol Cell. 2001;8(2):317-25.
48. Bensaad K, Tsuruta A, Selak MA, Vidal MNC, Nakano K, Bartrons R, Gottlieb E, Vousden KH. TIGAR, a p53-inducible regulator of glycolysis and apoptosis. Cell. 2006;126(1):107-20.
49. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, McCullough ML, Patel AV, Ma J, Soerjomataram I. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. CA: a cancer journal for clinicians. 2018;68(1):31-54.
50. World Cancer Research Fund. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective: Continuous Update Project Expert Report. 2018.
51. Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. Circulation. 2012;126(1):126-32.
52. Abar L, Vieira AR, Aune D, Sobiecki JG, Vingeliene S, Polemiti E, Stevens C, Greenwood DC, Chan DS, Schlesinger S. Height and body fatness and colorectal cancer risk: an update of the WCRF–AICR systematic review of published prospective studies. European journal of nutrition. 2018;57(5):1701-20.
53. Gunnell D, Okasha M, Davey Smith G, Oliver S, Sandhu J, Holly J. Height, leg length, and cancer risk: a systematic review. Epidemiologic reviews. 2001;23(2):313-42.
54. Hughes LA, van den Brandt PA, Goldbohm RA, De Goeij AF, De Bruine AP, Van Engeland M, Weijenberg MP. Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from the Netherlands Cohort Study. International journal of epidemiology. 2010;39(5):1333-44.
55. Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age–period–cohort analysis. European journal of cancer prevention. 2002;11(5):489-95.
56. Svensson E, Møller B, Tretli S, Barlow L, Engholm G, Pukkala E, Rahu M, Tryggvadottir L. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. Cancer Causes & Control. 2005;16(3):215-23.
57. Keinan-Boker L, Vin-Raviv N, Liphshitz I, Linn S, Barchana M. Cancer incidence in Israeli jewish survivors of World War II. Journal of the National Cancer Institute. 2009;101(21):1489-500.
58. Xie SH, Lagergren J. A possible link between famine exposure in early life and future risk of gastrointestinal cancers: Implications from age-period-cohort analysis. International journal of cancer. 2017;140(3):636-45.
59. Garcia H, Song M. Early-life obesity and adulthood colorectal cancer risk: a meta-analysis. Revista Panamericana de Salud Pública. 2019;43:e3.
60. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiology and Prevention Biomarkers. 2007;16(12):2533-47.
61. Robsahm TE, Aagnes B, Hjartåker A, Langseth H, Bray FI, Larsen IK. Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. European Journal of Cancer Prevention. 2013;22(6):492-505.
62. Wolin KY, Yan Y, Colditz GA, Lee I. Physical activity and colon cancer prevention: a meta-analysis. British journal of cancer. 2009;100(4):611-6.
63. Samad A, Taylor R, Marshall T, Chapman

MA. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Colorectal disease. 2005;7(3):204-13.
64. World Cancer Research Fund/American Institute for Cancer Research. Diet, nutrition, physical activity and cancer: a global perspective. Continuous Update Project Expert Report 2018.
65. Wild C, Weiderpass E, Stewart B. World cancer report 2020. Lyon: International Agency for Research on Cancer. 2020;199:512.
66. Park J, Morley TS, Kim M, Clegg DJ, Scherer PEJNRE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. 2014;10(8):455.
67. Hopkins BD, Goncalves MD, Cantley LCJJoco. Obesity and cancer mechanisms: cancer metabolism. 2016;34(35):4277.
68. Friedenreich CM, Ryder-Burbidge C, McNeil JJMO. Physical activity, obesity and sedentary behavior in cancer etiology: Epidemiologic evidence and biologic mechanisms. 2020.
69. Fasshauer M, Blüher MJTips. Adipokines in health and disease. 2015;36(7):461-70.
70. Renehan AG, Frystyk J, Flyvbjerg A. Obesity and cancer risk: the role of the insulin–IGF axis. Trends in Endocrinology & Metabolism. 2006;17(8):328-36.
71. Thomas RJ, Kenfield SA, Jimenez A. Exercise-induced biochemical changes and their potential influence on cancer: a scientific review. British journal of sports medicine. 2017;51(8):640-4.
72. Yu N, Ruan Y, Gao X, Sun J. Systematic review and meta-analysis of randomized, controlled trials on the effect of exercise on serum leptin and adiponectin in overweight and obese individuals. Hormone and Metabolic Research. 2017;49(03):164-73.
73. Stefan N, Häring H-U, Hu FB, Schulze MB. Divergent associations of height with cardiometabolic disease and cancer: epidemiology, pathophysiology, and global implications. The lancet Diabetes & endocrinology. 2016;4(5):457-67.
74. Castejón M, Plaza A, Martinez-Romero J, Fernandez-Marcos PJ, de Cabo R, Diaz-Ruiz A. Energy restriction and colorectal cancer: a call for additional research. Nutrients. 2020;12(1):114.
75. Yin S, Bai H, Jing D. Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients: a systemic review and meta-analysis. Diagnostic pathology. 2014;9(1):1-6.
76. Renehan AG, Zwahlen M, Minder C, T O'Dwyer S, Shalet SM, Egger M. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. The Lancet. 2004;363(9418):1346-53.
77. Joshi RK, Kim WJ, Lee S-A. Association between obesity-related adipokines and colorectal cancer: a case-control study and meta-analysis. World Journal of Gastroenterology: WJG. 2014;20(24):7941.
78. Yoon YS, Kwon AR, Lee YK, Oh SW. Circulating adipokines and risk of obesity related cancers: a systematic review and meta-analysis. Obesity research & clinical practice. 2019;13(4):329-39.
79. Ogino S, Stampfer M. Lifestyle factors and microsatellite instability in colorectal cancer: the evolving field of molecular pathological epidemiology. Oxford University Press; 2010. p. 365-7.
80. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut. 2010:gut. 2010.217182.
81. Hamada T, Keum N, Nishihara R, Ogino S. Molecular pathological epidemiology: new developing frontiers of big data science to study etiologies and pathogenesis. Journal of gastroenterology. 2017;52(3):265-75.
82. Nielsen JS, Jakobsen E, Holund B, Bertelsen K, Jakobsen A. Prognostic significance of

p53, Her-2, and EGFR overexpression in borderline and epithelial ovarian cancer. Int J Gynecol Cancer. 2004;14(6):1086-96.
83. Dijkema IM, Struikmans H, Dullens HF, Kal HB, van der Tweel I, Battermann JJ. Influence of p53 and bcl-2 on proliferative activity and treatment outcome in head and neck cancer patients. Oral Oncol. 2000;36(1):54-60.
84. van den Brandt PA, Goldbohm RA, Veer PVT, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. Journal of clinical epidemiology. 1990;43(3):285-95.
85. Hughes LA, Schouten LJ, Goldbohm RA, van den Brandt PA, Weijenberg MP. Self-reported clothing size as a proxy measure for body size. Epidemiology. 2009;20(5):673-6.
86. van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19(3):553-8.
87. van den Brandt PA. Molecular pathological epidemiology of lifestyle factors and colorectal and renal cell cancer risk. Maastricht Pathology 2018. 11th Joint Meeting of the British Division of the International Academy of Pathology and the Pathological Society of Great Britain & Ireland, 19-22 June 2018. . J Pathol 2018;246 Suppl 1: S9.



2

VALIDITY AND REPRODUCIBILITY OF IMMUNOHISTOCHEMICAL SCORING BY TRAINED NON-PATHOLOGISTS ON TISSUE MICROARRAYS

Josien C. A. Jenniskens, **Kelly Offermans**, Iryna Samarska, Gregorio E. Fazzi, Colinda C. J. M. Simons, Kim M. Smits, Leo J. Schouten, Matty P. Weijenberg, Piet A. van den Brandt, Heike I. Grabsch

Cancer Epidemiology, Biomarkers & Prevention. 2021;30:1867-74

ABSTRACT

BACKGROUND
Scoring of immunohistochemistry (IHC) staining is often done by non-pathologists, especially in large-scale tissue microarray (TMA)-based studies. Studies on the validity and reproducibility of scoring results from non-pathologists are limited. Therefore, our main aim was to assess interobserver agreement between trained non-pathologists and an experienced histopathologist for three IHC markers with different subcellular localization (nucleus/membrane/cytoplasm).

METHODS
Three non-pathologists were trained in recognizing adenocarcinoma and IHC scoring by a senior histopathologist. Kappa statistics were used to analyze interobserver and intraobserver agreement for 6,249 TMA cores from a colorectal cancer series.

RESULTS
Interobserver agreement between non-pathologists (independently scored) and the histopathologist was "substantial" for nuclear and membranous IHC markers (K_{range} = 0.67–0.75 and K_{range} = 0.61–0.69, respectively), and "moderate" for the cytoplasmic IHC marker (K_{range} = 0.43–0.57). Scores of the three non-pathologists were also combined into a "combination score" (if at least two non-pathologists independently assigned the same score to a core, this was the combination score). This increased agreement with the pathologist ($K_{nuclear}$ = 0.74; $K_{membranous}$ = 0.73; $K_{cytoplasmic}$ = 0.57). Interobserver agreement between non-pathologists was "substantial" ($K_{nuclear}$= 0.78; $K_{membranous}$ = 0.72; $K_{cytoplasmic}$ = 0.61). Intraobserver agreement of non-pathologists was "substantial" to "almost perfect" ($K_{nuclear,range}$ = 0.83–0.87; $K_{membranous,range}$ = 0.75–0.82; $K_{cytoplasmic}$ = 0.69). Overall, agreement was lowest for the cytoplasmic IHC marker.

CONCLUSIONS
This study shows that adequately trained non-pathologists are able to generate reproducible IHC scoring results, that are similar to those of an experienced histopathologist. A combination score of at least two non-pathologists yielded optimal results.

IMPACT
Non-pathologists can generate reproducible IHC results after appropriate training, making analyses of large-scale molecular pathological epidemiology studies feasible within an acceptable time frame.

INTRODUCTION

The introduction of the tissue microarray (TMA) technology by Kononen and colleagues[1] in 1998 has enabled large-scale studies using archival formalin-fixed paraffin-embedded (FFPE) tissue blocks[2,3]. The TMA technology has the advantage that sampling of cores leaves the donor block relatively intact, allowing it to be sampled multiple times[3,4]. Furthermore, immunohistochemistry (IHC) on TMAs is cost effective and less time consuming than performing IHC on full tissue sections[2-6]. In addition, a higher level of assay standardization can be achieved, improving reproducibility of results[3,4,6-8].

Several studies have shown a high degree of concordance between IHC results obtained from TMA sections and full sections when three 0.6 mm cores per case were used[9-13]. Interestingly, a study by Gavrielides and colleagues[14] found slightly higher interobserver agreement for HER2 scoring on TMAs compared with full sections, suggesting a potential benefit of the restricted field of view.

Manual scoring of TMA sections can take a considerable amount of time if individual scores need to be provided for hundreds or thousands of cores [7,15]. Although scoring by automated image analysis has been proposed as a potential alternative to manual scoring, IHC markers present in tumor cells and other cell populations at the same time are challenging to assess automatically[16].

Scoring of IHC stained sections is often done by non-pathologists[17,18]. However, studies on the validity of results from non-pathologists are limited. Jaraj and colleagues[19] suggested that after adequate training, non-pathologists are able to produce valid and reproducible IHC results for a cytoplasmic marker. However, it has been suggested that apart from the expert histopathologist knowledge, the agreement of IHC results between observers might also be affected by the subcellular localization of the marker of interest (nucleus/membrane/cytoplasm)[20]. There is a limited number of studies investigating scoring agreement of markers with different subcellular localizations. One of these studies reported similar overall kappa values for scoring of staining in different subcellular compartments[21,22], whereas another study reported considerably lower agreement for scoring of cytoplasmic immunostaining[23].

We hypothesized that there is good interobserver agreement between trained non-pathologists and pathologists for IHC scoring on TMAs, and that the interobserver agreement does not depend on the subcellular localization of the staining. Therefore, the aims of the current study were to (i) assess interobserver agreement between trained non-pathologists and an experienced pathologist, and (ii) assess agreement of three IHC markers with different subcellular localization (nucleus/membrane/cytoplasm).

MATERIALS AND METHODS

STUDY POPULATION, TISSUE COLLECTION, AND TMA CONSTRUCTION

For TMA construction, tissue blocks from colorectal cancer resections of cases from the Netherlands Cohort Study (NLCS) were collected retrospectively from Dutch hospitals[24-26]. Hematoxylin & eosin (H&E)-stained sections were reviewed and the area with the highest tumor density was identified. From this area, three 0.6-mm-diameter cores with tumor and three cores with normal epithelium were sampled per case for TMA construction (TMA-Grandmaster, 3D-Histech). In total, 78 TMA blocks were constructed containing 7,963 tumor cores. Ethical approval was obtained from Medical Ethical Committee MUMC, number METC 2019-1085.

IMMUNOHISTOCHEMISTRY

Five micrometers thick serial sections were cut from all 78 TMA blocks and subjected to IHC using an automated immunostainer (DAKO Autostainer Link 48, Glostrup). TP53, GLUT1, and PTEN were chosen as markers to assess interobserver and intraobserver agreement in scoring nuclear, membranous, and cytoplasmic immunoreactivity, respectively, as these are established IHC markers routinely used in clinical setting. Details of primary antibodies and staining protocols are shown in Table 1. Staining protocols for all markers were optimized to eliminate background and nonspecific staining. Sections were counterstained with Mayer's Hematoxylin (VWR International B.V.), dehydrated, and mounted with a glass coverslip and xylene-based mounting medium (DPX, Sigma-Aldrich). All TMA sections were scanned using the Aperio scanner (Leica Microsystems) at 40 magnification at the University of Leeds (Leeds, UK) Scanning Facility.

Table 1 | Overview staining protocols, all performed using the DAKO Autostainer Link 48.

Antibody	Clone	Supplier (cat. no.)	Antigen retrieval	Dilution	Incubation time
Pan-CK	AE1/AE3	DAKO (GA05361-2)	PT high[a]	RTU[b]	10 min.
TP53	DO-7	DAKO (M700101-2)	PT high[a]	RTU[b]	20 min.
GLUT1	-	Thermo Fisher Scient. (RB-9052-P)	PT low[c]	1:200	20 min.
PTEN	6H2.1	DAKO (M362729-2)	PT high[a]	1:100	20 min.

Note: visualisation system, Envision FLEX Visualization Kit (K8008, DAKO); chromogen, 3,30-diaminobenzidine (DAB).
[a]High pH retrieval (K8004) for 20 minutes on the Dako PT link (Agilent Technologies).
[b]RTU: ready-to-use.
[c]Low pH retrieval (K8005) for 20 minutes on the Dako PT link (Agilent Technologies).

QUALITY CONTROL

Presence of adenocarcinoma was confirmed for every individual core by reviewing the H&E-stained TMA sections. In case of tumor identification difficulties because of poor tumor differentiation or a large number of inflammatory cells, pan-cytokeratin staining was used to identify tumor cells.

Table 2 | Percentage of slides evaluated per assessor for all IHC markers.

Assessor	Experience	Nuclear (TP53)	Membranous (GLUT1)	Cytoplasmic (PTEN)	Intra-observer[b]
1	NP	100%	25%[a]	100%	X
2	NP	100%	100%	100%	10%
3	NP	100%	100%	100%	10%
4	P	10%	10%	10%	X

Abbreviations: NP, non-pathologist; P, pathologist.
[a]Assessor 1 left the project early because of an unforeseen work relocation.
[b]Percentage of slides rescored per protein.

IMMUNOHISTOCHEMICAL SCORING

Three non-pathologists (G.E. Fazzi: histology technician; K. Offermans: PhD student; J.C.A. Jenniskens: PhD student) were trained by a senior histopathologist (H.I. Grabsch) in (i) recognizing adenocarcinoma on H&E-stained TMA sections; (ii) recognizing immunoreactivity and distinguishing between immunoreactivity in the nucleus, membrane, and cytoplasm; and (iii) scoring of two TMA sections (~200 cores) for every immunostaining to ensure that the same criteria were used by all assessors.

After training, the three non-pathologists scored all tumor cores for TP53, GLUT1, and PTEN immunostainings. The scores from the three non-pathologists were combined into a "combination score." If at least two non-pathologists independently assigned the same score to a core, this score became the combination score. If all non-pathologists assigned different scores, the core was categorized as "no agreement". Because not all cores were scored by three non-pathologists for GLUT1 (Table 2), the remaining scores of the combination score were based on two non-pathologists. When comparing scores from pairs of trained non-pathologists to the score of the pathologist, non-pathologists' scores were combined as described for the combination score of three non-pathologists.

For evaluation of intraobserver agreement, two non-pathologists (assessor 2 and 3) evaluated 10% randomly selected TMA sections (range: 538–681 cores) per marker for a second time after a period of at least 5 months. These scores were only used to assess intraobserver agreement. To assess interobserver agreement between pathologist and

non-pathologists, an experienced pathologist (I. Samarska) evaluated the same 10% randomly selected TMA sections for every marker. The contribution of each assessor to the IHC scoring of the different markers is shown in Table 2.

TP53 positivity was defined as unequivocal strong nuclear staining and scored semiquantitatively as published previously[13,27], with minor adaptations, as: (i) no positive tumor nuclei; (ii) ≤10% positive tumor nuclei; (iii) 11% to 50% positive tumor nuclei; (iv) 51% to 90% positive tumor nuclei; and (v) 91% to 100% positive tumor nuclei (Figure 1A).

GLUT1 positivity was defined as any membranous (complete or incomplete) immunostaining of tumor cells, and scored as published previously[28,29]: (i) no tumor cells with membranous immunostaining; (ii) ≤10% tumor cells with membranous immunostaining; (iii) 11% to 50% tumor cells with membranous immunostaining; (iv) >50% tumor cells with membranous immunostaining (Figure 1B).

PTEN scoring was performed as described previously[30], comparing cytoplasmic immunostaining intensity of the tumor cells with that of adjacent stromal cells. PTEN immunostaining was classified as: (i) negative (no PTEN staining in the tumor cells); (ii) weak (staining intensity in the tumor cells weaker than in the stromal cells); (iii) moderate (similar staining intensity in tumor and stromal cells); or (iv) strong (staining intensity in the tumor cells stronger than in the stromal cells), see Figure 1C. In case of heterogeneous immunostaining, the region with the highest staining intensity prevailed.

Uninterpretable (e.g., folded cores) or missing cores were categorized as "uninterpretable" and excluded from analyses for all markers.

STATISTICAL ANALYSES

Interobserver and intraobserver agreement was assessed using all cores that passed quality control. Cohen's kappa was used for assessing interobserver agreement between assessor pairs and for assessing intraobserver agreement within one assessor[31]. Fleiss' kappa was used for assessing interobserver agreement between more than two assessors[32]. All kappa values were weighted (33), taking into account the magnitude of the disagreement (e.g., ≤10% vs. >50% is worse than ≤10% vs. 11%–50%). A weight of 0.5 was chosen for scoring an adjacent category and a weight of zero for non-adjacent categories. Non-weighted Fleiss' kappa was used for assessing the variation in interobserver agreement between scoring categories. To calculate kappa confidence intervals, the bootstrap method was used with 1,000 repetitions[34-36]. The interpretation of kappa values is shown in Supplementary Table S1. Agreement between each pair of assessors was determined, as well as agreement between the combination score of two or three non-pathologists and the pathologist's score (for the latter, cores for which no agreement was reached were excluded from analyses). Data were analyzed using Stata (version 15.1, Statacorp).

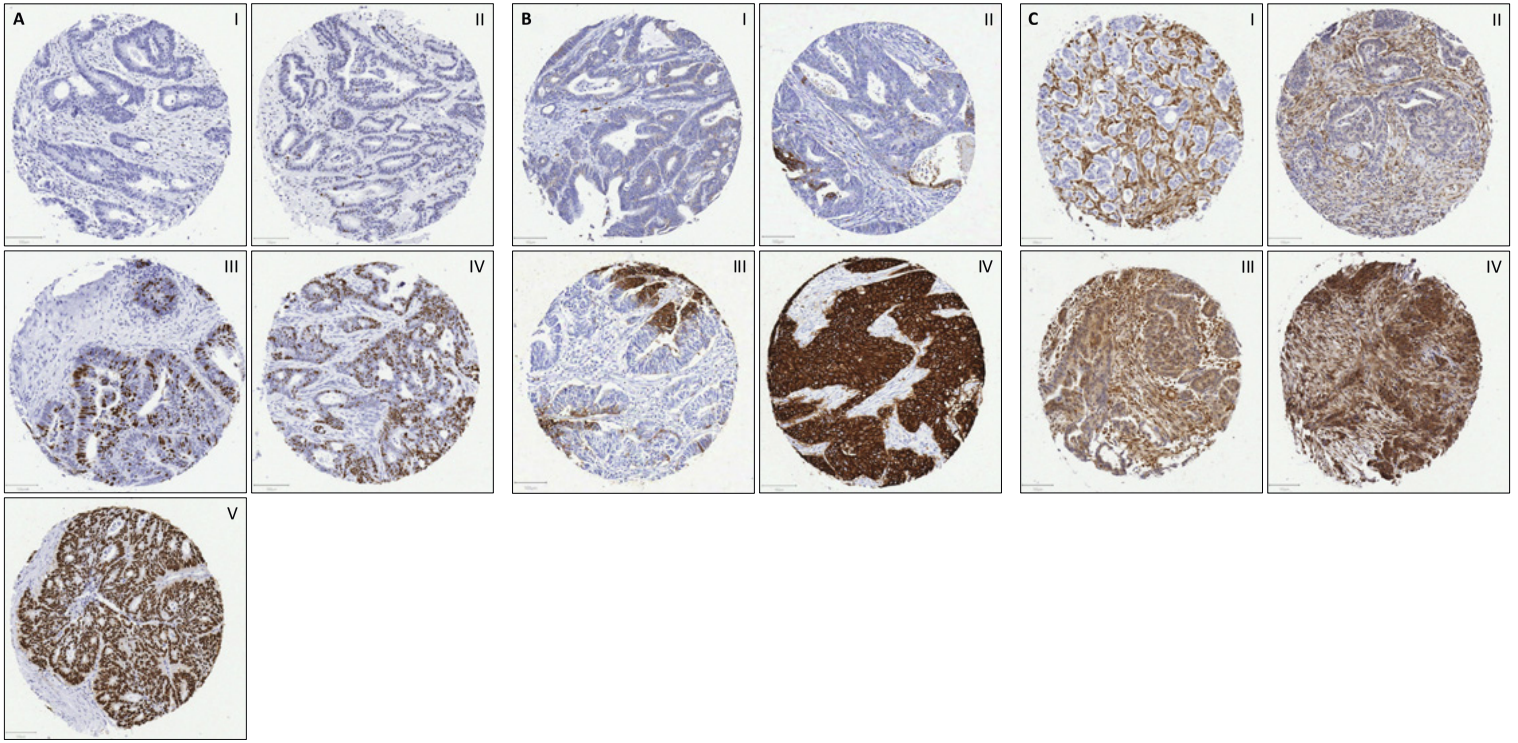


Figure 1 | IHC scoring of nuclear (TP53), membranous (GLUT1), and cytoplasmic (PTEN) protein expression on TMAs. The localization of the antigen–antibody complex is visualized in brown (3,30-diaminobenzidine) and counterstained in blue (hematoxylin). **A)** Nuclear positivity was defined as unequivocal strong nuclear staining and scored as: **(I)** negative nuclear immunoreactivity; **(II)** ≤10% nuclear immunoreactivity; **(III)** 11%–50% nuclear immunoreactivity; **(IV)** 51%–90% nuclear immunoreactivity; **(V)** 90%–100% nuclear immunoreactivity. **B)** Membranous positivity was defined as any membranous (complete or incomplete) immunostaining of tumor cells and scored as: **(I)** negative membranous immunoreactivity; **(II)** ≤10% membranous immunoreactivity; **(III)** 11%–50% membranous immunoreactivity; **(IV)** >50% membranous immunoreactivity. **C)** Cytoplasmic protein expression was evaluated by comparing the cytoplasmic immunostaining intensity of the tumor cells to that of adjacent stromal cells, and scored as: **(I)** negative cytoplasmic immunoreactivity; **(II)** weak positive cytoplasmic immunoreactivity; **(III)** moderate positive cytoplasmic immunoreactivity; **(IV)** strong positive cytoplasmic immunoreactivity. In case of heterogeneous immunostaining, the region with the highest staining intensity prevailed.

RESULTS

In total, 78 TMA blocks containing 7,963 tumor cores were available. After quality control, 1,714 (21.5%) cores were excluded (464 missing cores; 1,135 cores lacking tumor tissue; 115 uninterpretable tissue cores), leaving 6,249 tumor cores for analyses. All cores were evaluated by at least two assessors (Table 2). Frequency distributions of scores assigned by all assessors for nuclear (TP53), membranous (GLUT1), and cytoplasmic (PTEN) immunoreactivity are shown in
Supplementary Tables S2–S4.

INTEROBSERVER AGREEMENT

Non-pathologist versus pathologist
Weighted kappa values of interobserver agreement between non-pathologists and pathologist are shown in Table 3 (non-weighted kappa values in Supplementary Table S5). Kappa values of each individual non-pathologist with the pathologist showed "substantial" agreement for nuclear (K_{range} = 0.67–0.75) and membranous immunostainings (K_{range} = 0.61–0.69), and "moderate" for cytoplasmic immunostaining (K_{range} = 0.43–0.57). The combination score of the three non-pathologists showed "substantial" agreement with the pathologist's score for nuclear (K = 0.74) and membranous immunoreactivity (K = 0.73), and "moderate" agreement for cytoplasmic immunoreactivity (K = 0.57). The combination score of two non-pathologists showed similar agreement with the pathologist's score as the combination score of three non-pathologists ($K_{nuclear,range}$ = 0.75–0.81; $K_{membranous,range}$ = 0.75–0.79; $K_{cytoplasmic,range}$ = 0.54–0.65). For the majority of scores (range, 90.3%–98.6%), equal or adjacent scoring categories were assigned (Table 4) by pathologist and non-pathologists.

In Supplementary Table S6, the agreement per scoring category is shown by non-weighted kappa values. The lowest and highest scoring categories show higher agreement among non-pathologist assessors ($K_{nuclear}$ 0.83 and 0.79; $K_{membranous}$ 0.68 and 0.82; $K_{cytoplasmic}$ 0.61 and 0.51, respectively), than the scoring categories in between ($K_{nuclear,range}$ = 0.35–0.56; $K_{membranous,range}$ = 0.45–0.53; $K_{cytoplasmic,range}$ = 0.49–0.53). Adding the pathologist assessor, this again led to highest agreement in the most extreme categories for nuclear and membranous stainings ($K_{nuclear}$ 0.86 and 0.67; $K_{membranous}$ 0.74 and 0.76, respectively). For cytoplasmic stainings the agreement was highest for the lowest scoring category, and decreased with increasing scoring categories ($K_{category0}$ = 0.60; $K_{category1}$ = 0.53; $K_{category2}$ = 0.37; $K_{category3}$ = 0.32).

Non-pathologist versus non-pathologist
Interobserver agreement among non-pathologists is shown in Table 3 (non-weighted kappa values in Supplementary Table S5). Overall kappa values between all three non-pathologists were similar to those comparing the combination score and the pathologist's score ($K_{nuclear}$ 0.78 vs. 0.74; $K_{membranous}$ 0.72 vs. 0.73; $K_{cytoplasmic}$ 0.61 vs. 0.56, respectively). Scores for nuclear and membranous immunoreactivity showed the highest kappa values among non-pathologists, with an overall weighted kappa of 0.78

Table 3 | Interobserver agreement (weighted) between non-pathologists and pathologist.

	Nuclear K (95% CI)	Membranous K (95% CI)	Cytoplasmic K (95% CI)
NP vs P[a]			
1 vs 4	0.75 (0.72-0.79)	0.61 (0.55–0.67)[f]	0.57 (0.53–0.61)
2 vs 4	0.67 (0.63–0.71)	0.69 (0.65–0.73)	0.43 (0.38-0.48)
3 vs 4	0.70 (0.67–0.74)	0.69 (0.66–0.73)	0.56 (0.52–0.60)
1+2 vs 4[b,c]	0.80 (0.77–0.84)	0.77 (0.70–0.83)[f]	0.57 (0.51–0.62)
1+3 vs 4[b,c]	0.81 (0.77–0.84)	0.79 (0.73–0.85)[f]	0.65 (0.60–0.70)
2+3 vs 4[b,c]	0.75 (0.72–0.79)	0.75 (0.72–0.79)	0.54 (0.50–0.60)
Combination score[c,d] vs 4	0.74 (0.71–0.78)	0.73 (0.69–0.77)	0.57 (0.52–0.61)
NP vs NP[e]			
1 vs 2	0.74 (0.73–0.75)	0.69 (0.67–0.72)[f]	0.55 (0.54–0.57)
1 vs 3	0.79 (0.79–0.80)	0.66 (0.64–0.69)[f]	0.64 (0.62–0.65)
2 vs 3	0.80 (0.79–0.81)	0.81 (0.80–0.82)	0.65 (0.64–0.67)
1 vs 2 vs 3[g]	0.78	0.72[f]	0.61

Abbreviations: NP, non-pathologist; P, pathologist.
Nuclear: TP53; membranous: GLUT1; cytoplasmic: PTEN.
[a]Based on a random 10% of TMA sections (range, 538–681 cores).
[b]Comparison of a combination of two non-pathologists with the pathologist: if the two non-pathologists independently assigned the same score to a core, this was the combined score. If the non-pathologists assigned a different score, the core was categorized as no agreement.
[c]Cores where no agreement was reached between non-pathologists (combination score = no agreement) were excluded for analyses.
[d]The combination score is based on all three non-pathologist's scores: if at least two assessors independently assigned the same score to a core, this was the combination score. If none of the assessors assigned the same score, the core was categorized as no agreement.
[e]Based on all cores (N = 6,249).
[f]Assessor 1 left the project early because of an unforeseen work relocation, 1,457 cores were evaluated.
[g]Confidence interval for weighted kappa of multiple assessors (>2) could not be calculated using Stata.

Table 4 | Percentage of discrepancies between assessors.

	Nuclear Difference in categories[b]				Membranous[a] Difference in categories[b]				Cytoplasmic Difference in categories[b]			
	0	**1**	**2**	**3/4**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
Interobserver												
NP vs P[c]												
1 vs 4	57.9	32.4	8.0	1.8	62.5	29.4	6.0	2.2	54.1	42.1	3.9	0.0
2 vs 4	51.4	41.0	6.7	0.9	70.5	24.5	4.0	1.0	63.0	33.9	2.7	0.5
3 vs 4	61.7	32.0	5.1	1.2	70.6	25.2	3.1	1.0	62.1	36.5	1.4	0.0
Combination vs 4	69.8	27.1	2.9	0.2	73.7	22.7	3.0	0.7	64.6	34.8	0.7	0.0
NP vs NP[d]												
1 vs 2	72.0	24.2	3.5	0.3	68.3	29.2	2.1	0.4	65.4	34.0	0.6	0.0
1 vs 3	76.3	22.2	1.4	0.1	65.5	30.9	3.2	0.4	73.2	26.5	0.3	0.0
2 vs 3	76.3	22.4	1.2	0.1	81.4	17.2	1.3	0.1	76.0	23.8	0.2	0.0
Intraobserver												
NP vs NP[c]												
2 vs 2	82.6	16.3	1.0	0.0	82.4	16.6	1.0	0.0	78.8	21.2	0.0	0.0
3 vs 3	84.1	15.3	0.6	0.0	74.2	24.8	0.8	0.3	80.0	20.0	0.0	0.0

Note: Uninterpretable cores were excluded.
Abbreviations: NP, non-pathologist; P, pathologist. Nuclear: TP53; membranous: GLUT1; cytoplasmic: PTEN.
[a]Assessor 1 left the project early because of an unforeseen work relocation.
[b]Difference in categories assigned by the two assessors: 0=same category assigned (no discrepancy); 1=adjacent categories were assigned (e.g.,<10% positive and 11%–50% positive); 2=difference between assigned categories was 2 (e.g., <10% positive and >50% positive); 3/4=difference between assigned categories was 3 or 4 (e.g., negative and >50%).
[c]Based on a random 10% of TMA sections.
[d]Based on all TMA sections.

(K_{range} = 0.74–0.80) and 0.72 (K_{range} = 0.66–0.81), respectively. Agreement was lowest for cytoplasmic immunoreactivity, with an overall kappa of 0.61 (K_{range} = 0.55–0.65). In the majority of non-pathologists' scores (range, 96.2%–99.8%), equal or adjacent scoring categories were assigned (Supplementary Table S6).

INTRAOBSERVER AGREEMENT OF NON-PATHOLOGISTS

Weighted intraobserver kappa values of two non-pathologists are shown in Table 5 (non-weighted kappa values in Supplementary Table S7). The intraobserver agreement was highest for scoring nuclear and membranous immunoreactivity, showing "almost perfect" agreement ($K_{observer2}$ = 0.83; $K_{observer3}$ = 0.87), and "substantial" to "almost perfect" agreement ($K_{observer2}$ = 0.82; $K_{observer3}$ = 0.75), respectively. Scoring of cytoplasmic immunoreactivity showed "substantial" agreement ($K_{observer2}$ = 0.69; $K_{observer3}$ = 0.69). In the majority of scores (range, 98.9%–100%), equal or adjacent categories were assigned at the first and second timepoint (Supplementary Table S6).

Table 5 | Intraobserver agreement (weighted) of two non-pathologists, based on 10% randomly selected TMA sections.

	Assessor 2 K (95% CI)	Assessor 3 K (95% CI)
Nuclear	0.83 (0.80-0.86)	0.87 (0.84-0.90)
Membranous	0.82 (0.79-0.85)	0.75 (0.72-0.78)
Cytoplasmic	0.69 (0.64-0.74)	0.69 (0.64-0.74)

Abbreviations: CI, confidence interval; TMA, tissue microarray.
Nuclear: TP53; membranous: GLUT1; cytoplasmic: PTEN.

DISCUSSION

TMAs are increasingly used to analyze protein expression by IHC in large-scale studies[2,3,5,37]. Scoring is often done by non-pathologists[17,18]; however, only few studies reported validity and reproducibility of scoring results[38,39]. To the best of our knowledge, our study is one of the first to investigate agreement of TMA-based scoring of immunoreactivity in different subcellular localizations by non-pathologists. Our study showed that interobserver agreement between an experienced histopathologist and trained non-pathologists was "moderate" to "substantial." Agreement with the pathologist's score did not further increase when a combination score from three instead of two trained non-pathologists was used.

INTEROBSERVER AGREEMENT NON-PATHOLOGISTS VERSUS PATHOLOGIST

Our study demonstrates that non-pathologists can generate reproducible results. These results are in line with a previous study by Jaraj and colleagues [19], reporting comparable kappa values for interobserver agreement between pathologists and non-pathologists. Even though it was not their main objective, two other studies reported comparable interobserver agreement between pathologists and non-pathologists[22,40]. However, some of the studies reported weighted kappa values[19,22], but did not state what weights were assigned to adjacent scoring categories, making a direct comparison of kappa values with our study impossible.

Considering the subjectivity of immunoreactivity scoring, several studies recommended that scoring should be done by multiple assessors to improve interobserver agreement[39,41,42]. Our study confirmed that combining scores from multiple non-pathologists into a combination score increased interobserver agreement with the pathologist's score. Combining scores of three non-pathologists instead of two did not change interobserver agreement with the pathologist, indicating that IHC scoring by two non-pathologists seems to be sufficient to yield reliable IHC results.

IMMUNOREACTIVITY SCORING IN DIFFERENT SUBCELLULAR LOCALIZATIONS

A limited number of studies investigated scoring agreement of immunoreactivity in different subcellular localizations, showing inconsistent results[21-23]. We showed that scoring of nuclear and membranous immunoreactivity generally leads to higher interobserver agreement compared with cytoplasmic immunoreactivity, consistent with results of Bolton and colleagues[23]. However, this is in contrast to two other studies which did not find a difference in the intraobserver and interobserver agreement when scoring nuclear, membranous and cytoplasmic immunoreactivity[21,22]. These discrepant results might be explained by the use of different IHC scoring methods between studies.

The IHC markers selected for the current study were chosen to provide a range of subcellular localizations (nucleus/membrane/cytoplasm) for scoring purposes. These markers are generalizable to other IHC stainings considering the subcellular localization.

INTEROBSERVER AGREEMENT AMONG NON-PATHOLOGISTS

Hitherto, few studies reported interobserver agreement of IHC results among non-pathologists. In the current study, we found "substantial" to "almost perfect" agreement among trained non-pathologists, which is in line with previously published results on TMAs and whole tissue sections[17-19].

INTRAOBSERVER AGREEMENT OF NON-PATHOLOGISTS

IHC studies often report intraobserver kappa values as a measure of reproducibility. Our study shows that non-pathologists are able to generate reproducible IHC scores after appropriate training, which is in line with previous studies[17-19,40]. Interestingly, intraobserver kappa values of non-pathologists in the current study were similar to those previously reported for pathologists[23,43]. In general, across all three markers, disagreements were limited to one-category discordances (e.g., <10% vs. 11%–50%)

for all comparisons.

LIMITATIONS

Our study has some limitations. We have no information on intraobserver and interobserver agreement of pathologists, as this was beyond the scope of this article. Furthermore, the current study used TMA cores to assess interobserver and intraobserver agreement. It has been described in the literature that interobserver agreement increases when using TMA cores compared with whole tissue sections[14]. Thus, it remains to be clarified whether the agreement among non-pathologists and between non-pathologists and pathologists is similar in full tissue sections. However, the aim of this study was specifically to investigate IHC scoring on TMAs, because non-pathologists will mainly be involved in IHC scoring in large-scale studies using TMAs. Also, we did not directly compare the scoring performance between trained non-pathologists and untrained non-pathologists; thus, we are not able to draw direct conclusions on the necessity of training, and in particular whether similar results would have been obtained without training.

RECOMMENDATIONS

We propose some recommendations which could improve comparability of IHC studies. First, it is important to report what weights were used for analyses of weighted kappa values. In addition, we think it would be of value to report both weighted and non-weighted kappa values. Second, it should be mentioned clearly in the methods what the IHC scoring experience of assessors was. If done by non-pathologists, it is important to report their training. Third, our results showed that disagreements were mostly limited to one-category discordances, suggesting that less refined scoring protocols may potentially improve agreement. This is in line with previous studies[44,45], in which the authors showed that agreement improved when using scoring protocols with less categories. However, we acknowledge that the number of categories of the scoring protocol depends on the novelty and clinical relevance of the biomarker being studied. Scoring protocols for potential new biomarkers might comprise more categories compared with well-known biomarkers. Finally, we suggest that IHC scoring should be performed by at least two non-pathologists to be able to assess interobserver agreement among assessors. Ideally, these non-pathologists are trained by an expert pathologist and a certain percentage of samples (e.g., 10%) are double-scored by the pathologist to ensure quality of scoring.

CONCLUSION

In this large study investigating interobserver and intraobserver agreement of TMA-based immunoreactivity scores between pathologists and non-pathologists, we have shown that non-pathologists can generate reproducible IHC scoring results that are similar to those of an experienced pathologist. A combination score of at least two non-pathologists yielded optimal results. Future studies are required to validate our findings and to examine the practical implications and impact of potential misclassification, by comparing effect estimates for established stain-outcome associations when using the pathologist's score versus the non-pathologists' combination score.

REFERENCES

1. Kononen J, Bubendorf L, Kallioniemi A, Bärlund M, Schraml P, Leighton S, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–7.
2. Tawfik El-Mansi M, Williams AR. Validation of tissue microarray technology using cervical adenocarcinoma and its precursors as a model system. Int J Gynecol Cancer 2006;16:1225–33.
3. Ilyas M, Grabsch H, Ellis IO, Womack C, Brown R, Berney D, et al. Guidelines and considerations for conducting experiments using tissue microarrays. Histopathology 2013;62:827–39.
4. Camp RL, Neumeister V, Rimm DL. A decade of tissue microarrays: progress in the discovery and validation of cancer biomarkers. J Clin Oncol 2008;26: 5630–7.
5. Domeny-Duarte P, Niero L, Domingues MAC. Tissue microarrays of bone marrow aspirate clot allow assessment of multiple samples. Pathology, research and practice. 2020;216:152721.
6. Shergill IS, Shergill NK, Arya M, Patel HRH. Tissue microarrays: a current medical research tool. Curr Med Res Opin 2004;20:707–12.
7. Bubendorf L, Nocito A, Moch H, Sauter G. Tissue microarray (TMA) technology: miniaturized pathology archives for high-throughput in situ studies. J Pathol 2001;195:72–9.
8. Al Kuraya K, Simon R, Sauter G. Tissue microarrays for high-throughput molecular pathology. Ann Saudi Med 2004;24:169–74.
9. Boone J, van Hillegersberg R, van Diest PJ, Offerhaus GJA, Borel Rinkes IHM, Ten Kate FJW. Validation of tissue microarray technology in squamous cell carcinoma of the esophagus. Virchows Arch 2008;452:507–14.
10. Zhang D, Salto-Tellez M, Putti TC, Do E, Koay ESC. Reliability of tissue microarrays in detecting protein expression and gene amplification in breast cancer. Mod Pathol 2003;16:79–84.
11. Hassan S, Ferrario C, Mamo A, Basik M. Tissue microarrays: emerging standard for biomarker validation. Curr Opin Biotechnol 2008;19:19–25.
12. Hoos A, Urist MJ, Stojadinovic A, Mastorides S, Dudas ME, Leung DH, et al. Validation of tissue microarrays for immunohistochemical profiling of cancer specimens using the example of human fibroblastic tumors. Am J Pathol 2001; 158:1245–51.
13. Jourdan F, Sebbagh N, Comperat E, Mourra N, Flahault A, Olschwang S, et al. Tissue microarray technology: validation in colorectal carcinoma and analysis of p53, hMLH1, and hMSH2 immunohistochemical expression. Virchows Arch 2003;443:115–21.
14. Gavrielides MA, Conway C, O'Flaherty N, Gallas BD, Hewitt SM. Observer performance in the use of digital and optical microscopy for the interpretation of tissue-based biomarkers. Anal Cell Pathol 2014;2014:157308.
15. Jawhar NM. Tissue microarray: a rapidly evolving diagnostic and research tool. Ann Saudi Med 2009;29:123–7.
16. O'Hurley G, Sjostedt E, Rahman A, Li B, Kampf C, Ponten F, et al. Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers. Mol Oncol 2014;8:783–98.
17. Nielsen JS, Jakobsen E, Holund B, Bertelsen K, Jakobsen A. Prognostic significance of p53, Her-2, and EGFR overexpression in borderline and epithelial ovarian cancer. Int J Gynecol Cancer 2004;14:1086–96.
18. Dijkema IM, Struikmans H, Dullens HF, Kal HB, van der Tweel I, Battermann JJ. Influence of p53 and bcl-2 on proliferative

activity and treatment outcome in head and neck cancer patients. Oral Oncol 2000;36:54–60.
19. Jaraj SJ, Camparo P, Boyle H, Germain F, Nilsson B, Petersson F, et al. Intra- and interobserver reproducibility of interpretation of immunohistochemical stains of prostate cancer. Virchows Arch 2009;455:375–81.
20. Lejeune M, Jaen J, Pons L, Lopez C, Salvado MT, Bosch R, et al. Quantification of diverse subcellular immunohistochemical markers with clinicobiological relevancies: validation of a new computer-assisted image analysis procedure. J Anat 2008;212:868–78.
21. Zlobec I, Terracciano L, Jass JR, Lugli A. Value of staining intensity in the interpretation of immunohistochemistry for tumor markers in colorectal cancer. Virchows Arch 2007;451:763–9.
22. Kirkegaard T, Edwards J, Tovey S, McGlynn LM, Krishna SN, Mukherjee R, et al. Observer variation in immunohistochemical analysis of protein expression, time for a change? Histopathology 2006;48:787–94.
23. Bolton KL, Garcia-Closas M, Pfeiffer RM, Duggan MA, Howat WJ, Hewitt SM, et al. Assessment of automated image analysis of breast cancer tissue microarrays for epidemiologic studies. Cancer Epidemiol Biomarkers Prev 2010;19:992–9.
24. van den Brandt PA, Goldbohm RA, van't Veer P, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990;43:285–95.
25. van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 1990;19:553–8.
26. van den Brandt PA. Molecular pathological epidemiology of lifestyle factors and colorectal and renal cell cancer risk. Maastricht Pathology 2018. 11th Joint Meeting of the British Division of the International Academy of Pathology and the Pathological Society of Great Britain & Ireland, 19–22 June 2018. J Pathol 2018;246:S1–S46.
27. Resnick MB, Routhier J, Konkin T, Sabo E, Pricolo VE. Epidermal growth factor receptor, c-MET, b-catenin, and p53 expression as prognostic indicators in stage II colon cancer: a tissue microarray study. Clin Cancer Res 2004;10:3069–75.
28. Cooper R, Sarioglu S, Sökmen S, Füzün M, Küpelioglu A, Valentine H, et al. Glucose transporter-1 (GLUT-1): a potential marker of prognosis in rectal carcinoma? Br J Cancer 2003;89:870–6.
29. Sakashita M, Aoyama N, Minami R, Maekawa S, Kuroda K, Shirasaka D, et al. Glut1 expression in T1 and T2 stage colorectal carcinomas: its relationship to clinicopathological features. Eur J Cancer 2001;37:204–9.
30. Richman SD, Adams R, Quirke P, Butler R, Hemmings G, Chambers P, et al. Pretrial inter-laboratory analytical validation of the FOCUS4 personalised therapy trial. J Clin Pathol 2016;69:35–41.
31. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960; 20:37–46.
32. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull 1971;76:378–82.
33. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol Bull 1968;70:213–20.
34. Efron B, Tibshirani RJ. An introduction to the bootstrap Boca Raton, FL: Chapman & Hall; 1993.
35. Lee J, Fung KP. Confidence interval of the kappa coefficient by bootstrap resampling. Psychiatry Res 1993;49:97–8.
36. Reichenheim ME. Confidence intervals for the kappa statistic. Stata J 2004;4: 421–8.

37. Shiraishi T, Shinto E, Nearchou IP, Tsuda H, Kajiwara Y, Einama T, et al. Prognostic significance of mesothelin expression in colorectal cancer disclosed by area-specific four-point tissue microarrays. Virchows Arch 2020;477:409–20.
38. Cross SS. Observer accuracy in estimating proportions in images: implications for the semiquantitative assessment of staining reactions and a proposal for a new system. J Clin Pathol 2001;54:385–90
39. Adams EJ, Green JA, Clark AH, Youngson JH. Comparison of different scoring systems for immunohistochemical staining. J Clin Pathol 1999;52:75–7.
40. Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. Arch Pathol Lab Med 2011;135:233–42.
41. Masmoudi H, Hewitt SM, Petrick N, Myers KJ, Gavrielides MA. Automated quantitative assessment of HER-2/neu immunohistochemical expression in breast cancer. IEEE Trans Med Imaging 2009;28:916–25.
42. Obuchowski NA. How many observers are needed in clinical studies of medical imaging? AJR Am J Roentgenol 2004;182:867–9.
43. Ali A, Bell S, Bilsland A, Slavin J, Lynch V, Elgoweini M, et al. Investigating various thresholds as immunohistochemistry cutoffs for observer agreement. Appl Immunohistochem Mol Morphol 2017;25:599–608.
44. de Jong D, Rosenwald A, Chhanabhai M, Gaulard P, Klapper W, Lee A, et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications–a study from the Lunenburg Lymphoma Biomarker Consortium. J Clin Oncol 2007;25:805–12.
45. Rüschoff J, Dietel M, Baretton G, Arbogast S, Walch A, Monges G, et al. HER2 diagnostics in gastric cancer—guideline validation and development of standardized immunohistochemical testing. Virchows Arch 2010;457:299–307.

Supplementary Table S1 | Interpretation of Kappa values.

Kappa coefficient	Agreement
0 - 0.20	Slight
0.21 - 0.40	Fair
0.41 - 0.60	Moderate
0.61 - 0.80	Substantial
0.81 - 1.00	Almost perfect

Supplementary Table S2 | Scoring percentages of nuclear (TP53) immunoreactivity in tumor cells between non-pathologists, the combination score and the pathologist score. In total, 6249 tumor cores that passed quality control were evaluated.

	All cores N_{cores} = 6249				Random 10% of cores N_{cores} = 538	
	NP1	**NP2**	**NP3**	**Combi NP[c]**	**Combi NP[c]**	**P4**
Negative	39.5	31.6	30.7	32.7	29.6	31.0
1-10%	12.1	14.0	18.7	15.4	15.2	11.5
11-50%	10.8	10.6	10.8	9.7	11.9	19.1
51-90%	9.9	7.9	14.4	10.1	10.4	22.1
91-100%	26.0	34.3	23.8	27.2	27.9	15.4
Uninterpretable[a]	1.7	1.6	1.6	1.3	1.3	0.7
No agreement[b]	-	-	-	3.5	3.7	-

Abbreviations: NP, non-pathologist; P, pathologist
[a]Excluded from IHC analyses because of missing core, no tumor tissue in core, no immunostaining, or folded core.
[b]If none of the assessors assigned the same score, the core was categorized as no agreement for the combination score.
[c]The combination score is based on the non-pathologist's scores: if at least two assessors independently assigned the same score to a core, this was the combination score.

Supplementary Table S3 | Scoring percentages of membranous (GLUT1) immunoreactivity in tumor cells between non-pathologists, the combination score, and the pathologist score. In total, 6249 tumor cores that passed quality control were evaluated.

	All cores N_{cores} = 6249				Random 10% of cores N_{cores} = 681	
	NP1[c]	**NP2**	**NP3**	**Combi NP[d]**	**Combi NP[d]**	**P4**
Negative	36.2	24.3	23.8	21.6	24.1	25.4
1-10%	23.5	17.6	15.9	12.4	12.8	18.2
11-50%	17.8	31.2	33.8	27.1	26.0	25.8
51-100%	20.3	25.0	24.0	22.0	26.6	30.1
Uninterpretable[a]	2.1	2.0	2.5	1.7	0.9	0.4
No agreement[b]	-	-	-	15.2	9.7	-

Abbreviations: NP, non-pathologist; P, pathologist
[a]Excluded from IHC analyses because of missing core, no tumor tissue in core, no immunostaining, or folded core.
[b]If none of the assessors assigned the same score, the core was categorized as no agreement for the combination score.
[c]Assessor 1 left the project early because of an unforeseen work relocation, 1457 cores were evaluated.
[d]The combination score is based on the non-pathologist's scores: if at least two assessors independently assigned the same score to a core, this was the combination score.

Supplementary Table S4 | Scoring percentages of cytoplasmic (PTEN) immunoreactivity in tumor cells between non-pathologists, the combination score, and the pathologist score. In total, 6249 tumor cores that passed quality control were evaluated.

	All cores N_{cores} = 6249				Random 10% of cores N_{cores} = 645	
	NP1	**NP2**	**NP3**	**Combi NP[c]**	**Combi NP[c]**	**P4**
Negative	14.5	7.0	7.5	7.8	8.5	12.9
Weak positive	49.7	51.8	52.3	53.0	51.6	45.7
Moderate positive	22.5	32.6	31.2	28.7	30.2	20.2
Strong positive	8.6	5.6	5.4	5.8	4.7	17.8
Uninterpretable[a]	4.8	3.1	3.5	3.3	3.3	3.4
No agreement[b]	-	-	-	1.4	1.7	-

Abbreviations: NP, non-pathologist; P, pathologist
[a]Excluded from IHC analyses because of missing core, no tumor tissue in core, no immunostaining, or folded core.
[b]If none of the assessors assigned the same score, the core was categorized as no agreement for the combination score.
[c]The combination score is based on the non-pathologist's scores: if at least two assessors independently assigned the same score to a core, this was the combination score.

Supplementary Table S5 | Interobserver agreement (weighted) between non-pathologists and pathologist.

	Nuclear K (95% CI)	Membranous K (95% CI)	Cytoplasmic K (95% CI)
NP vs P[a]			
1 vs 4	0.63 (0.59-0.68)	0.50 (0.43-0.57)[f]	0.47 (0.42-0.52)
2 vs 4	0.51 (0.47-0.56)	0.60 (0.56-0.64)	0.31 (0.26-0.37)
3 vs 4	0.57 (0.52-0.62)	0.60 (0.56-0.64)	0.45 (0.39-0.50)
1+2 vs 4[b,c]	0.69 (0.64-0.74)	0.69 (0.61-0.77)[f]	0.47 (0.41-0.53)
1+3 vs 4[b,c]	0.70 (0.65-0.75)	0.72 (0.64-0.80)[f]	0.55 (0.49-0.61)
2+3 vs 4[b,c]	0.62 (0.57-0.67)	0.67 (0.62-0.72)	0.43 (0.37-0.49)
Combination score[c,d] vs 4	0.61 (0.57-0.66)	0.64 (0.59-0.69)	0.45 (0.40-0.50)
NP vs NP[e]			
1 vs 2	0.62 (0.60–0.63)	0.58 (0.55-0.61)[f]	0.45 (0.44-0.47)
1 vs 3	0.68 (0.67-0.70)	0.54 (0.51-0.58)[f]	0.56 (0.55-0.58)
2 vs 3	0.69 (0.67-0.70)	0.74 (0.73-0.76)	0.59 (0.58-0.61)
1 vs 2 vs 3[f]	0.66 (0.65-0.67)	0.62 (0.59-0.64)[f]	0.54 (0.52-0.55)

Abbreviations: NP, non-pathologist; P, pathologist.
Nuclear: TP53; membranous: GLUT1; cytoplasmic: PTEN.
[a]Based on a random 10% of TMA sections (range 538-681 cores).
[b]Comparison of a combination of two non-pathologists with the pathologist: if the two non-pathologists independently assigned the same score to a core, this was the combined score. If the non-pathologists assigned a different score, the core was categorized as no agreement.
[c]Cores where no agreement was reached between non-pathologists (combination score = no agreement) were excluded for analyses.
[d]The combination score is based on all three non-pathologist's scores: if at least two assessors independently assigned the same score to a core, this was the combination score. If none of the assessors assigned the same score, the core was categorized as no agreement.
[e]Based on all cores (N=6249).
[f]Assessor 1 left the project early because of an unforeseen work relocation, 1457 cores were evaluated.

Supplementary Table S6 | Interobserver agreement (non-weighted) per scoring category of non-pathologist assessors (NP1-3; n=) and all (NP1-3 & P4) assessors.

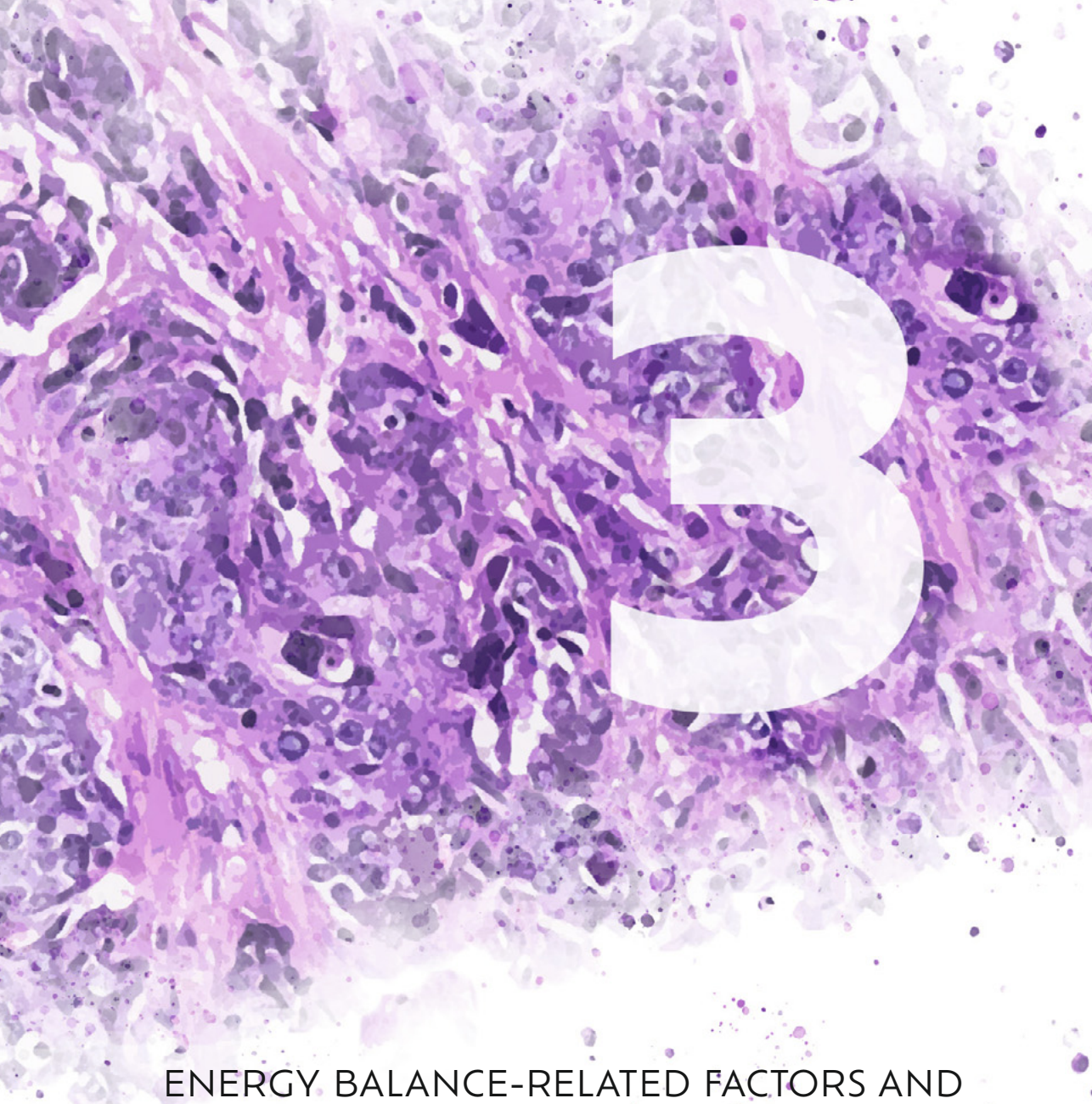
	Agreement non-pathologist assessors (NP1-3) N_{cores} = 6249			Agreement all assessors (NP1-3 & P4) N_{cores} = 538-681		
	Nuclear[a] κ	Membranous[b,c] κ	Cytoplasmic[d] κ	Nuclear[a] κ	Membranous[b,c] κ	Cytoplasmic[d] κ
Scoring categories						
0	0.83	0.68	0.61	0.86	0.74	0.60
1	0.56	0.45	0.53	0.57	0.44	0.53
2	0.45	0.53	0.49	0.46	0.56	0.37
3	0.35	0.82	0.51	0.32	0.76	0.32
4	0.79	-	-	0.67	-	-
Not interpretable	0.65	0.71	0.62	0.64	0.59	0.54
Combined	0.66	0.62	0.53	0.62	0.64	0.47

Confidence intervals for kappa values per category could not be calculated using Stata.
[a]Nuclear (TP53): 0 = negative; 1 = <10% positive nuclei; 2 = 11-50% positive nuclei; 3 = 51-90% positive nuclei; 4 = 91-100% positive nuclei.
[b]Membranous (GLUT1): 0 = negative; 1 = <10% positive membranes; 2 = 11-50% positive membranes; 3 = 51-100% positive membranes.
[c]Assessor 1 left the project early because of an unforeseen work relocation.
[d]Cytoplasmic (PTEN): 0 = negative; 1 = weak positive cytoplasm; 2 = moderate positive cytoplasm; 3 = strong positive.

Supplementary Table S7 | Intraobserver agreement (non-weighted) of two non-pathologists, based on 10% randomly selected TMA sections (N=538-681).

	Assessor 2 **K (95% CI)**	**Assessor 3** **K (95% CI)**
Nuclear	0.73 (0.69-0.77)	0.80 (0.76-0.84)
Membranous	0.75 (0.72-0.79)	0.65 (0.60-0.69)
Cytoplasmic	0.63 (0.58-0.69)	0.65 (0.60-0.70)

Nuclear: TP53; membranous: GLUT1; cytoplasmic: PTEN.



3

ENERGY BALANCE-RELATED FACTORS AND RISK OF COLORECTAL CANCER EXPRESSING DIFFERENT LEVELS OF PROTEINS INVOLVED IN THE WARBURG-EFFECT

Josien C. A. Jenniskens, Kelly Offermans, Colinda C. J. M. Simons, Iryna Samarska, Gregorio E. Fazzi, Kim M. Smits, Leo J. Schouten, Matty P. Weijenberg, Heike I. Grabsch, Piet A. van den Brandt

Cancer Epidemiology, Biomarkers & Prevention. 2022;31:633-46

ABSTRACT

BACKGROUND
Energy balance–related factors [body mass index (BMI), waist circumference, physical activity] have been associated with colorectal cancer risk. Warburg-effect activation via PI3K/Akt signaling is one of the proposed mechanisms. We investigated whether energy balance–related factors were associated with risk of Warburg-subtypes in colorectal cancer.

METHODS
We investigated this using immunohistochemistry for six proteins involved in the Warburg-effect (LDHA, GLUT1, MCT4, PKM2, P53, PTEN) on tissue microarrays of 2,399 incident colorectal cancer cases from the prospective Netherlands Cohort Study (n_{total} = 120,852; $n_{subcohort}$ = 5,000; aged 55–69 in 1986; 20.3 years follow-up). Data analyses included 3,911 subcohort members and 1,972 colorectal cancer cases with complete covariate data. Expression levels of all proteins were combined into a pathway-based sum score and categorized into three "Warburg-subtypes" (Warburg-low/moderate/high). Multivariable Cox regression analyses were used to estimate associations of BMI, clothing size (waist circumference proxy), and physical activity with Warburg-subtypes in colorectal cancer.

RESULTS
BMI and clothing size were positively associated with Warburg-moderate and Warburg-high colon cancer risk in men ($P_{heterogeneity}$ = 0.192). In women, clothing size was positively associated with Warburg-low and Warburg-high colon cancer ($P_{heterogeneity}$ = 0.005). Non-occupational physical activity was inversely associated with Warburg-low and Warburg-moderate colon cancer in women ($P_{heterogeneity}$ = 0.045), but positively associated with Warburg-high rectal cancer in men ($P_{heterogeneity}$ = 0.089).

CONCLUSIONS
The Warburg-effect might be involved in associations between adiposity and colon cancer risk, though additional mechanisms could be at play in women as well. The inverse association between physical activity and colon cancer might be explained by mechanisms other than the Warburg-effect.

IMPACT
Further research is needed to reproduce these results and investigate possible additional mechanisms.

INTRODUCTION

Energy balance–related factors are known to influence risk of colorectal cancer. Measures of adiposity, such as body mass index (BMI) and waist circumference, have been associated with increased risk of colorectal cancer[1,2], whereas physical activity is inversely associated with colorectal cancer risk[2-4]. Up till now, the underlying biological mechanisms are not fully understood. Several reviews published on this subject[5-7] implied three main factors: adipocyte-derived cytokines (adipokines), insulin and insulin-like growth factor 1 (IGF-1) signaling, and sex hormones.

Circulating levels of adipokines are influenced by the quantity of adipose tissue, with a larger number of adipocytes leading to higher circulating leptin and lower adiponectin levels[8]. Conversely, physical activity has been linked to lower circulating leptin and higher adiponectin levels, even independent of weight loss[9,10]. Similarly, increased serum levels of insulin and free IGF-1 have been reported for overweight and obese individuals[11], whereas reduced levels were observed in more physically active individuals[9]. High levels of leptin, insulin, and IGF-1 have all been associated with increased colorectal cancer risk[12-15], whereas high adiponectin levels have been associated with a decreased risk[14,15].

Adipokine, insulin, and IGF-1 signaling share a common downstream effect, namely activation of the PI3K/Akt signaling pathway[16]. Apart from its well-known properties like cell survival and growth, the PI3K/Akt signaling pathway has been associated with the so-called metabolic switch[17,18]. Upon activation, the expression of glucose transporters and enzymes involved in glycolysis increases[19,20]. Upregulation of aerobic glycolysis in cancer cells was first observed in the 1920s by Warburg and colleagues, hence the term "Warburg-effect"[21]. While it was initially thought that the Warburg-effect was an effect rather than a cause of cancer, it is increasingly being considered a carcinogenic step[22]. This is further supported by the addition of "Reprogramming Energy Metabolism" as an Emerging Hallmark of Cancer in 2011[23].

Previous studies investigated the suggested link between energy balance–related factors and colorectal cancer mainly using circulating biomarkers (e.g., leptin or insulin; refs. 12–15). However, differentiating between cause and effect is difficult with circulating biomarkers, especially when they are measured at the time of cancer diagnosis, as the biomarker status may be influenced by the tumor. In the current study, we aimed to investigate whether this suggested link could be captured in the primary tumor itself by upregulation of the Warburg-effect.

We aimed to capture the Warburg-effect by ensuring that the different steps of the pathway were represented by at least one protein (Supplementary Table S1). These steps include: upstream regulation of the Warburg-effect (PTEN, P53), glucose import (GLUT1), glycolysis (PKM2), conversion of pyruvate into lactate (LDHA), and lactate secretion (MCT4). The expression levels of these six proteins (PTEN, P53, GLUT1, PKM2, LDHA, MCT4) were combined into a sum score, which was divided

into three subgroups, representing tumors with a low, moderate, or high likelihood of presence of the Warburg-effect, hereafter referred to as the Warburg-subtypes (Warburg-low, Warburg-moderate, Warburg-high, respectively).

We hypothesized that associations between energy balance–related factors (BMI; lower body clothing size, as a proxy for waist circumference; physical activity) and risk of colorectal cancer differ across Warburg-subtypes.

MATERIALS AND METHODS

STUDY DESIGN AND POPULATION

The Netherlands Cohort Study (NLCS) was initiated in 1986 and included 120,852 subjects ages 55–69 years at baseline. All participants completed a mailed, self-administered questionnaire on diet, smoking habits, anthropometry, history of selected diseases, physical activity, and other cancer risk factors[24]. The NLCS was approved by Institutional Review Boards from Maastricht University (Maastricht, the Netherlands) and the Netherlands Organization for Applied Scientific Research. Ethical approval was obtained from the Medical Ethical Committee of Maastricht University Medical Center+ (Maastricht, the Netherlands). The NLCS was conducted in accordance with the Declaration of Helsinki. All cohort members consented to participate in the NLCS by completing the questionnaire. For data processing and analysis, the case–cohort method was used[25]. Accumulated person-years in the cohort were estimated from a subcohort (n = 5,000), randomly sampled from the whole cohort immediately after baseline. These subcohort members were actively followed up biennially for vital status information and by linkage to municipal population registries. Only one male subcohort member was lost to follow-up.

Follow-up for cancer incidence was established by annual record linkage with the Netherlands Cancer Registry and PALGA, the nationwide Dutch Pathology Registry[26], covering 20.3 years of follow-up (September 17, 1986 until January 1, 2007). Completeness of cancer incidence follow-up by the Netherlands Cancer Registry and PALGA was estimated to be over 96%[27]. After excluding cases and subcohort members who reported a history of cancer (except skin cancer) at baseline, a total of 4,597 incident colorectal cancer cases and 4,774 subcohort members were available (Figure 1).

Formalin-fixed paraffin-embedded (FFPE) tissue blocks from primary tumor and matched normal colon tissue from 3,872 colorectal cancer cases were requested from participating laboratories as part of the Rainbow-TMA (tissue microarray) project during 2012–2017[28]. Colorectal cancer cases were selected on the basis of available linkage to a PALGA-record (which provides access to pathology labs) and surgical specimen with pathology report, or coloscopic resection. Cases treated with neoadjuvant therapy

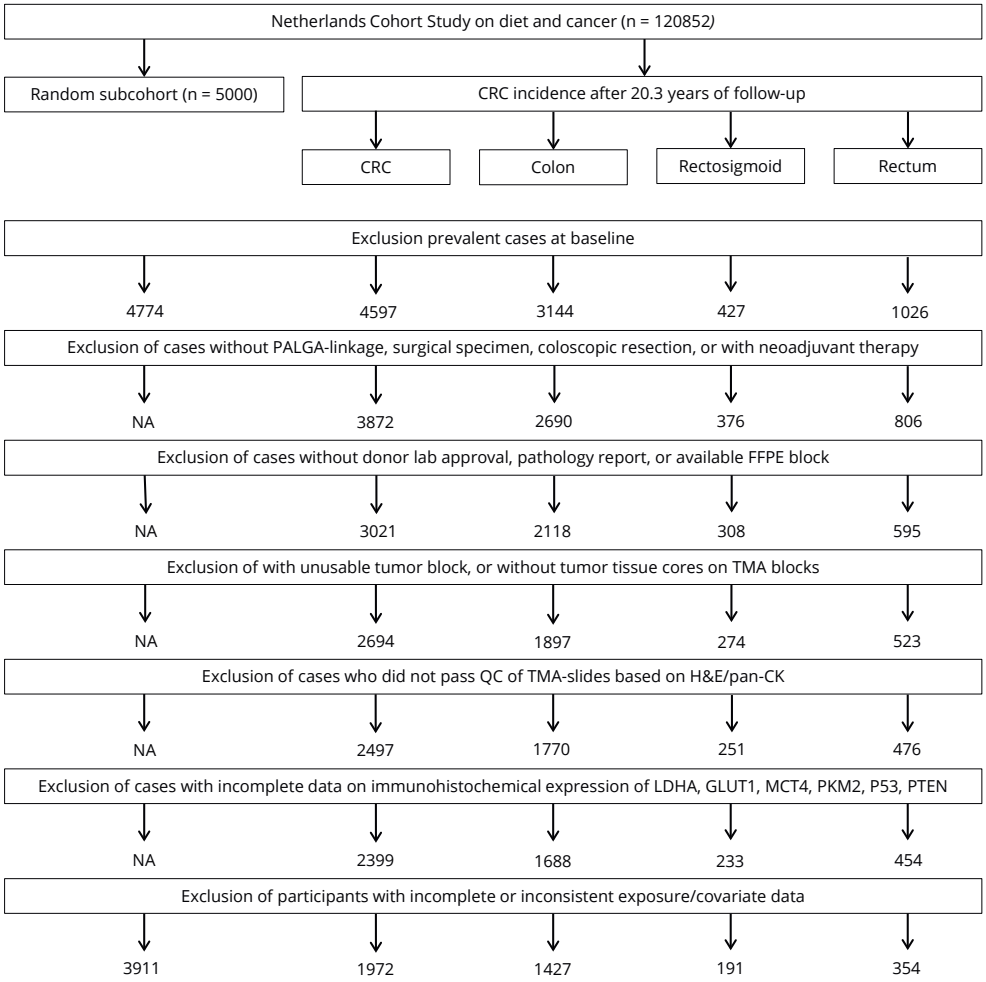


Figure 1 | Flow diagram of the number of colorectal cancer cases and subcohort members, NLCS 1986–2006. *Abbreviations: CRC, colorectal cancer; NA, not applicable; PALGA, Dutch Pathology Registry; FFPE, formalin-fixed paraffin-embedded; TMA, tissue microarray; QC, quality control; H&E, hematoxylin & eosin; pan-CK, pan-cytokerin.*

were excluded. Tissue blocks from 3,021 colorectal cancer cases were successfully collected from 43 pathology laboratories throughout the Netherlands (78% retrieval rate).

For TMA construction, pathologists reviewed scanned hematoxylin & eosin (H&E)-stained sections and identified areas with the highest tumor density, from which three 0.6-mm-diameter cores were sampled per case along with three normal tissue cores (TMA-Grandmaster, 3D-Histech). In total, tumor tissue of 2,694 colorectal cancer cases was successfully assembled in 78 TMA blocks (Figure 1).

IMMUNOHISTOCHEMISTRY

Five micrometers thick sections were cut from all 78 TMA blocks, H&E stained according to standard protocol, and subjected to immunohistochemistry (IHC). IHC was performed using an automated immunostainer (DAKO Autostainer Link 48) for GLUT1, P53, and PTEN, and manually for LDHA, MCT4, and PKM2. Details of the primary antibodies and staining protocols are shown in Supplementary Table S2. All TMA sections were scanned using an Aperio scanner (Leica Microsystems) at 40x magnification at the University of Leeds (Leeds, United Kingdom) Scanning Facility or at the Department of Pathology, Aachen University Hospital (Aachen, Germany).

Three non-pathologists (G.E. Fazzi: histology technician; K. Offermans: PhD student; J.C.A. Jenniskens: PhD student) were trained by a senior histopathologist (H.I. Grabsch) in recognizing adenocarcinoma and IHC scoring[29]. Presence of adenocarcinoma was confirmed for every individual core by reviewing H&E-stained TMA sections, in combination with pan-cytokeratin stained sections if necessary. Requiring at least one core per case, 2,497 cases passed quality control (Figure 1).

After quality control, all cores were scored by at least two assessors (Supplementary Table S3 shows contribution of each assessor), independently and blinded for case characteristics. IHC scoring protocols for all markers are described in the Supplementary Materials and Methods and shown in Supplementary Figure S1. Kappa values on interobserver and intraobserver scoring agreement are shown in Supplementary Table S4.

Figure 2 illustrates the stepwise process of combining multiple core-level scores into case-level Warburg-subtypes. If at least two assessors assigned the same score to a core, this score became the "combination score." Remaining discrepancies were resolved by consensus agreement of two non-pathologist assessors or by an experienced pathologist, resulting in a final score for each core. Case-level protein expression was determined by taking the average of the final scores of available cores (range: 1–3 cores per case) and rounding it to the nearest scoring category. The average score per case was subdivided into three subgroups, representing low, moderate, or high expression. Cutoffs for PTEN and P53 were based on previous literature[30,31], cut-offs for other proteins were determined on the basis of distribution of cases (Supplementary Table S4 shows cut-offs per protein).

CREATING WARBURG-SUBTYPES

To create Warburg-subtypes, we used a pathway-based sum score of case-level protein expression levels of LDHA, GLUT1, MCT4, PKM2, P53, and PTEN (Figure 2). Cases with incomplete protein expression data were excluded (Figure 1). Expression of LDHA, GLUT1, MCT4, PKM2, and P53 are positively associated with the Warburg-effect[18,32], whereas PTEN expression is inversely associated with the Warburg-effect[32]. Therefore, for all proteins, except PTEN, high protein expression was given a score of 2, moderate expression a score of 1, and low expression a score of 0. For PTEN, this score was reversed; high PTEN expression was given a score of 0, moderate expression a score

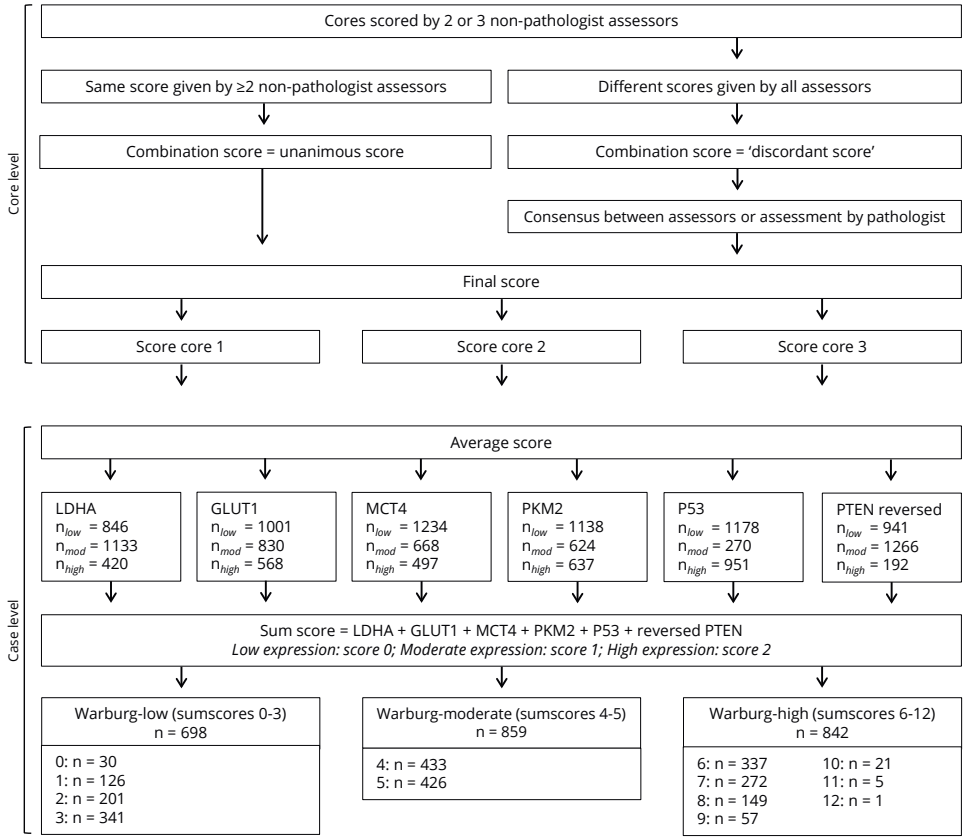


Figure 2 | Flow diagram of getting from multiple core-level scores to case-level Warburg-subtypes. n_{low} = number of people with low expression; n_{mod} = number of people with moderate protein expression; n_{high} = number of people with high protein expression.

of 1, and low expression a score of 2. The sum score is the sum of scores of all proteins (range: 0–12), whereby a higher score indicates a higher likeliness of presence of the Warburg-effect. For statistical efficiency, cases were then divided into tertiles based on the sum score to establish Warburg-subtypes. Distribution of the sum score did not differ according to sex or tumor location, leading to the following cut-offs for all cases: cases with sum scores 0–3 were classed as "Warburg-low" (n = 698, 29.1%), sum scores 4–5 as "Warburg-moderate" (n = 859, 35.8%), and sum scores 6–12 as "Warburg-high" (n = 842, 35.1%; Figure 2). Clinical characteristics of the cases stratified on Warburg-subtypes are shown in Supplementary Table S5.

ENERGY BALANCE-RELATED FACTORS

All NLCS participants returned a mailed, self-administered questionnaire on anthropometry, physical activity, diet, and other risk factors at baseline in 1986[24]. BMI at baseline (kg/m^2) was calculated using baseline weight (kg) divided by height squared (m^2). Participants were asked to report their lower body clothing size (trouser/skirt) from their clothing label (Dutch sizes). This has previously been shown to be an adequate proxy for waist circumference when predicting cancer risk in the NLCS[33]. To estimate levels of non-occupational physical activity, participants were asked to report the average daily time spent on activities like walking, cycling, or doing sports, as described in more detail previously[34]. For occupational physical activity, energy expenditure and sitting time were estimated for the longest held job, which was self-reported at baseline. Jobs were classified as low, moderate, or high activity, as described previously[34]. Energy expenditure was classified as <8, 8–12, and >12 kJ/minute, and sitting time as sitting for >6, 2–6, and <2 working hours/day. Data on occupational physical activity were only available for the subcohort and for cases until 17.3 years of follow-up, because funding for later data entry and classification of occupations was unavailable. Furthermore, we did not analyze occupational physical activity measures in women because many did not have paid jobs[34].

COX REGRESSION MODELS

After excluding participants with incomplete or inconsistent data on exposure variables or confounders, 3,911 subcohort members and 1,972 colorectal cancer cases were available for analyses (Figure 1). Associations between energy balance–related factors and colorectal cancer risk were investigated stratified on sex, tumor location, and Warburg-subtypes. Cox proportional hazards models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI) for the association between colorectal cancer and BMI (according to sex-specific quartiles, and per 5 kg/m^2 increase), clothing size (according to sex-specific quartiles, and per two sizes increase), non-occupational physical activity (in categories of <30, 30–60, 60–90, >90 minutes per day, and per 30 minutes/day increase), and, for men, occupational physical activity (energy expenditure in categories of <8, 8–12, >12 kJ/minute; sitting time in categories of >6, 2–6, and <2 working hours/day). Standard errors of the HRs were estimated using the Huber–White sandwich estimator to account for additional variance introduced by sampling the subcohort from the total cohort[35]. The proportional hazards assumption was tested using the scaled Schoenfeld residuals[36] and by introducing time–covariate interactions into the models.

All multivariable models were adjusted for age, total energy intake (kcal/day), family history of colorectal cancer (yes/no), and alcohol intake (0; 0.1–4; 5–14; >15 g/day). BMI and clothing size models were additionally adjusted for non-occupational physical activity (minutes/day), and BMI models for height (cm). All physical activity models were additionally adjusted for BMI. Moreover, clothing size and BMI models were mutually adjusted as an indication of fat distribution, where clothing size adjusted for BMI represents a proxy for abdominal fatness, and BMI adjusted for clothing size as a proxy for subcutaneous fatness[33,37]. Potential additional confounders were

smoking status (never/former/current), level of education (primary or lower vocational education; secondary or medium vocational education; higher vocational education or university), red meat consumption (g/day), and processed meat consumption (g/day). These potential confounders were included in multivariable models if they introduced a ≥10% change in HRs.

Heterogeneity in associations between risk factors and Warburg-subtypes was tested to evaluate differences across tumors expressing different levels of proteins involved in the Warburg-effect. This was done using an adapted version of the competing risks procedure in Stata developed for the case-cohort design, as described previously[38,39]. Sensitivity analyses were performed by excluding the first 2 years of follow-up. Furthermore, analyses were performed for two instead of three Warburg-subtypes (Warburg-low: sum score 0–4; Warburg-high: sum score 5–12) to increase power. All analyses were conducted in Stata Statistical Software: Release 16 (StataCorp.).

RESULTS

Baseline lifestyle characteristics of subcohort members and colon and rectal cancer cases, overall and according to Warburg-subtypes, are shown in Table 1. Overweight and obesity were more often observed in colorectal cancer cases compared with subcohort members, especially for Warburg-moderate and Warburg-high colorectal cancer cases in men, and for Warburg-low colorectal cancer cases in women. Clothing size, as a proxy for waist circumference, showed similar trends. Male colon cancer cases showed equal levels of high non-occupational physical activity as subcohort members, but slightly lower levels of occupational energy expenditure and higher levels of occupational sitting time. In contrast, male rectal cancer cases more often showed high levels of non-occupational physical activity and low occupational sitting time, especially for the Warburg-high subtype. Female cases less often showed high levels of non-occupational physical activity compared with female subcohort members, especially in the Warburg-low and Warburgh-moderate groups for colon and in the Warburg-high group for rectal cancer. Warburg-high cases generally had the lowest level of education compared with Warburg-low and Warburg-moderate cases, except for men with colon cancer. Furthermore, a family history of colorectal cancer occurred less frequently in the Warburg-high subgroup compared with Warburg-low or Warburg-moderate subgroups for rectal cancer.

Tables 2–5 show multivariable-adjusted Cox regression models for energy balance–related factors in Warburg-subtypes, stratified on sex and tumor location. Age-adjusted Cox regression models are shown in Supplementary Tables S6–S9; results were similar to those of multivariable-adjusted models. Age was included as a time-varying covariate in all models, because of violation of the proportional hazards assumption.

ADIPOSITY

Both BMI and clothing size showed a positive association with total colon cancer risk in men (Tables 2 and 3), with HRs (95% CI) of 1.24 (1.07–1.44) per 5 kg/m^2 increment and of 1.31 (1.14–1.49) per two sizes increment. HRs for the same increments were enhanced for Warburg-moderate and Warburg-high subtypes [Warburg-moderate: HR_{BMI} (95% CI): 1.26 (1.01–1.57); $HR_{clothing}$: 1.45 (1.18–1.78); Warburg-high: HR_{BMI}: 1.39 (1.11–1.75); $HR_{clothing}$: 1.28 (1.06–1.56)], whereas the Warburg-low subtype showed weaker associations. After additional adjustment for clothing size, as a proxy for subcutaneous fatness, a similar association was found for Warburg-high [$HR_{5kg/m2}$ (95% CI): 1.42 (1.07–1.88)], whereas associations for total and Warburg-moderate colon cancer diminished (Supplementary Table S10). In contrast, adjustment for BMI in clothing size models, as a proxy for abdominal fatness, led to similar associations for Warburg-moderate [$HR_{two\ sizes}$ (95% CI): 1.41 (1.11–1.78)], but weaker associations for Warburg-high colon cancer [$HR_{two\ sizes}$ (95% CI): 1.15 (0.93–1.42)] (Supplementary Table S11). Neither BMI nor clothing size models showed statistically significant heterogeneity between Warburg-subtypes.

For rectal cancer, no associations with BMI or clothing size were observed in men (Table 2 and 3). After mutual adjustment, neither BMI nor clothing size showed statistically significant associations (Supplementary Table S10 and S11).

In women, BMI was not associated with colon cancer risk (Table 2), whereas clothing size showed a weak positive association with colon cancer risk [$HR_{two\ sizes}$ (95% CI): 1.09 (0.95–1.24)] (Table 3). This association was stronger for Warburg-low and Warburg-high subtypes [per two sizes: $HR_{Warburg\text{-}low}$ (95% CI): 1.27 (0.96–1.69); $HR_{Warburg\text{-}high}$: 1.20 (1.01–1.42)], whereas the association for Warburg-moderate seemed to be inverse [$HR_{two\ sizes}$ (95% CI): 0.84 (0.69–1.02)]. Statistically significant heterogeneity between Warburg-subtypes was observed for clothing size models per two sizes (continuous $P_{heterogeneity}$ = 0.007), as well as for models on quartiles of clothing size (categorical $P_{heterogeneity}$ = 0.018). Mutual adjustment, as an indication of fat distribution, resulted in inverse associations for BMI (Supplementary Table S10), and stronger associations for clothing size (Supplementary Table S11). However, BMI and clothing size showed high correlation in women (Spearman rank correlation: 0.76).

For rectal cancer, no associations were found for either BMI or clothing size in women (Table 2 and 3). Neither BMI nor clothing size showed statistically significant associations after mutual adjustment (Supplementary Tables S10 and S11).

PHYSICAL ACTIVITY

Non-occupational physical activity was not associated with total colon cancer risk in men (Table 4). Stratification on Warburg-subtypes did not lead to different associations. Energy expenditure at work was associated with a non-significant decreased risk of colon cancer (Table 5), and similar associations were shown for all Warburg-subtypes. Lower occupational sitting time was associated with a statistically significant decreased risk of colon cancer (Table 5), with HR (95% CI) for sitting <2 hours/day versus >6 hours/

day of 0.69 (0.53–0.91), and statistically significant trend over categories ($P_{trend,categories}$ = 0.007). After stratification on Warburg-subtypes, associations were in the same direction but reached statistical significance only for the Warburg-moderate group [HR (95% CI): 0.64 (0.43–0.94); $P_{trend,categories}$ = 0.025). Tests for heterogeneity did not reach statistical significance for any of these exposures.

For rectal cancer, non-occupational physical activity showed a positive association in men (Table 4), with HR (95% CI) of 1.05 (0.99–1.10) per 30 minutes/day and a statistically significant trend over categories ($P_{trend,categories}$ = 0.004). The association was stronger for the Warburg-high subtype [$HR_{30min/day}$ (95% CI): 1.10 (1.02–1.19)], with a statistically significant trend over categories ($P_{trend,categories}$ < 0.001). Heterogeneity between Warburg-subtypes was not statistically significant. For occupational physical activity, no clear associations were found with rectal cancer risk (Table 5).

In women, non-occupational physical activity was associated with decreased colon cancer risk (Table 4), with HR (95% CI) of 0.96 (0.90–1.02) per 30 minutes/day, and a statistically significant trend over categories ($P_{trend,categories}$ = 0.008). After stratification on Warburg-subtypes, a similar association was found for Warburg-low [$HR_{30min/day}$ (95% CI): 0.96 (0.86–1.08); $P_{trend,categories}$ = 0.042], a stronger effect for Warburg-moderate [$HR_{30min/day}$ (95% CI): 0.87 (0.78–0.96); $P_{trend,categories}$ = 0.006), and no effect for Warburg-high. A statistically significant difference in associations per 30 minutes/day increase was found between Warburg-subtypes ($P_{heterogeneity}$ = 0.050), but not for categorical models.

For rectal cancer, no statistically significant associations for non-occupational physical activity were found in women (Table 4).

SENSITIVITY ANALYSES

Sensitivity analyses excluding the first 2 years of follow-up did not lead to essential changes. Furthermore, analyses with two instead of three Warburg- subtypes generally led to similar conclusions. Associations that were found for the Warburg-moderate subtype (e.g., occupational sitting time with colon cancer in men) when using three Warburg-subtypes, resulted in similar associations for Warburg-low and Warburg-high subtypes when two Warburg-subtypes were used.

Table 1 | Baseline characteristics [mean (SD) or %] of subcohort members and colorectal cancer cases in Warburg-subtypes, by sex and tumor location; NLCS, 1986–2006.

	Subcohort	Colon				Rectum			
		Total	Warburg low	Warburg moderate	Warburg high	Total	Warburg low	Warburg moderate	Warburg high
Men									
N	1971	772	215	280	277	227	76	76	75
Overweight/obesity[a] (%)	46.6	51.0	47.4	50.0	54.9	49.8	48.7	51.3	49.3
Clothing size[b]	51.7 (2.7)	52.1 (2.6)	51.8 (2.5)	52.4 (2.7)	52.2 (2.6)	51.8 (2.5)	51.7 (2.7)	51.7 (2.1)	51.9 (2.7)
Non-occ. PA >60 min/day (%)	51.1	51.0	54.9	47.1	52.0	60.8	55.3	55.3	72.0
Occ. energy exp. >12 kJ/min (%)[c]	13.0	12.5	10.8	13.2	13.2	12.0	12.1	11.7	12.1
Occ. sitting time <2 hrs/day (%)[c]	25.9	24.6	27.4	24.5	22.4	30.4	27.3	26.7	37.9
Age (years)	61.3 (4.2)	61.6 (4.2)	61.8 (4.2)	61.3 (4.1)	61.8 (4.2)	60.7 (3.9)	60.8 (3.6)	61.1 (4.3)	60.2 (3.6)
Total energy intake (kcal/day)	2164 (500)	2121 (461)	2089 (455)	2164 (494)	2102 (427)	2240 (478)	2263 (448)	2230 (518)	2228 (471)
Family history of CRC (%)	5.4	10.8	10.2	11.1	10.8	9.3	10.5	10.5	6.7
Alcohol consumption (g/day)	15.1 (17.1)	15.1 (15.9)	15.3 (16.1)	15.6 (15.9)	14.5 (15.7)	17.4 (17.5)	16.4 (17.4)	16.9 (16.3)	18.8 (19.0)
Processed meat intake (g/day)	15.9 (16.9)	15.3 (14.7)	15.4 (15.9)	15.7 (14.4)	14.9 (14.0)	17.9 (17.6)	15.6 (11.7)	20.0 (22.6)	18.1 (16.6)
Red meat intake (g/day)	93.8 (41.2)	91.9 (40.0)	88.4 (40.1)	95.7 (40.2)	90.7 (39.7)	94.4 (39.6)	95.9 (38.2)	86.5 (39.1)	101 (40.7)
Never cigarette smokers (%)	12.7	12.7	12.6	12.9	12.6	8.8	7.9	11.8	6.7
High education[d] (%)	19.8	23.6	21.4	24.6	24.4	17.3	21.3	15.8	14.9

(continued)

	Subcohort	Colon				Rectum			
		Total	Warburg low	Warburg moderate	Warburg high	Total	Warburg low	Warburg moderate	Warburg high
Women									
N	1940	655	170	216	269	127	37	51	39
Overweight/obesity[a] (%)	43.6	45.3	48.2	43.5	45.0	50.4	56.8	52.9	41.0
Clothing size[b]	43.4 (2.9)	43.6 (3.3)	44.0 (4.2)	43.1 (2.9)	43.7 (3.0)	43.6 (2.7)	43.6 (2.5)	43.8 (2.8)	43.4 (2.7)
Non-occ. PA >60 min/day (%)	44.9	40.6	38.8	37.5	44.2	43.3	54.1	43.1	33.3
Age (years)	61.4 (4.3)	61.9 (4.1)	62.0 (4.1)	62.0 (4.1)	61.8 (4.1)	61.4 (4.2)	60.5 (4.1)	62.1 (4.2)	61.5 (4.3)
Total energy intake (kcal/day)	1684 (392)	1679 (385)	1661 (353)	1711 (411)	1664 (382)	1684 (346)	1664 (315)	1667 (340)	1724 (386)
Family history of CRC (%)	6.0	9.9	10.6	8.8	10.4	10.2	10.8	11.8	7.7
Alcohol consumption (g/day)	6.0 (9.5)	5.7 (9.5)	5.7 (10.1)	5.6 (8.8)	5.8 (9.6)	5.6 (8.7)	4.4 (5.7)	6.7 (9.9)	5.3 (9.4)
Processed meat intake (g/day)	10.3 (11.6)	10.0 (11.0)	10.7 (11.4)	10.5 (11.6)	9.0 (10.3)	11.4 (10.5)	11.2 (9.8)	10.1 (7.7)	13.2 (13.8)
Red meat intake (g/day)	81.0 (38.1)	77.4 (34.6)	77.0 (34.2)	77.6 (34.1)	77.4 (35.2)	88.3 (42.5)	100 (49.3)	83.9 (34.7)	83.1 (43.8)
Never cigarette smokers (%)	57.3	57.7	59.4	58.8	55.8	57.5	56.8	49.0	69.2
High education[d] (%)	9.5	10.0	10.7	11.1	8.7	5.6	2.7	10.0	2.6

Abbreviations: SD, standard deviation; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; (non-)occ., (non-)occupational; PA, physical activity; exp., expenditure

[a]BMI ≥25

[b]Based on fewer participants due to extra missings

[c]Based on fewer participants due to shorter follow-up (17.3 years), only available for men

[d]High education = university or higher vocational education

Table 2 | Multivariable-adjusted HRs[a] and 95% CIs for associations between BMI and Warburg-subtypes in CRC, by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
BMI quartiles (kg/m²)											
Men – colon											
< 23.4	22.2	7993	174	1.00 (ref.)	57	1.00 (ref.)	57	1.00 (ref.)	60	1.00 (ref.)	
23.4-24.9	24.2	8343	199	1.08 (0.84-1.38)	55	0.91 (0.60-1.36)	81	1.34 (0.92-1.94)	63	1.00 (0.68-1.46)	
25.0-26.6	25.7	7683	203	1.18 (0.91-1.52)	52	0.91 (0.60-1.38)	63	1.13 (0.76-1.68)	88	1.49 (1.03-2.17)	
> 26.6	27.8	7003	196	1.34 (1.03-1.73)	51	1.05 (0.69-1.59)	79	1.64 (1.12-2.41)	66	1.32 (0.89-1.96)	0.073
P-trend				0.021		0.860		0.035		0.038	
per 5 kg/m²		31022	772	1.24 (1.07-1.44)	215	1.05 (0.82-1.35)	280	1.26 (1.01-1.57)	277	1.39 (1.11-1.75)	0.192
Men – rectum											
< 23.4	22.2	7993	56	1.00 (ref.)	20	1.00 (ref.)	20	1.00 (ref.)	16	1.00 (ref.)	
23.4-24.9	24.2	8343	53	0.87 (0.58-1.31)	14	0.67 (0.33-1.37)	17	0.77 (0.39-1.51)	22	1.24 (0.63-2.42)	
25.0-26.6	25.7	7683	69	1.26 (0.85-1.86)	25	1.35 (0.71-2.56)	26	1.34 (0.72-2.47)	18	1.09 (0.54-2.18)	
> 26.6	27.8	7003	49	1.01 (0.66-1.54)	17	1.05 (0.52-2.12)	13	0.74 (0.36-1.51)	19	1.27 (0.62-2.60)	0.499
P-trend				0.507		0.437		0.902		0.634	
per 5 kg/m²		31022	227	1.08 (0.86-1.35)	76	1.13 (0.77-1.67)	76	1.04 (0.72-1.49)	75	1.06 (0.74-1.52)	0.933
Women – colon											
<22.8	21.5	9014	186	1.00 (ref.)	54	1.00 (ref.)	64	1.00 (ref.)	68	1.00 (ref.)	
22.8-24.7	23.8	8914	147	0.80 (0.62-1.03)	27	0.50 (0.31-0.82)	51	0.79 (0.53-1.18)	69	1.03 (0.72-1.48)	
24.8-27.0	25.7	8141	160	0.98 (0.76-1.27)	41	0.86 (0.55-1.34)	53	0.93 (0.63-1.37)	66	1.12 (0.78-1.63)	
>27.0	29.2	8158	162	1.02 (0.79-1.33)	48	1.05 (0.68-1.62)	48	0.86 (0.57-1.30)	66	1.15 (0.78-1.68)	0.246
P-trend				0.595		0.535		0.628		0.424	
per 5 kg/m²		34228	655	1.05 (0.93-1.19)	170	1.12 (0.91-1.39)	216	0.94 (0.77-1.14)	269	1.09 (0.92-1.30)	0.372

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
<22.8	21.5	9014	35	1.00 (ref.)	11	1.00 (ref.)	12	1.00 (ref.)	12	1.00 (ref.)	
22.8-24.7	23.8	8914	26	0.73 (0.43-1.24)	4	0.34 (0.11-1.10)	11	0.93 (0.40-2.17)	11	0.88 (0.37-2.08)	
24.8-27.0	25.7	8141	31	0.93 (0.55-1.57)	10	0.91 (0.37-2.22)	15	1.44 (0.63-3.32)	6	0.51 (0.19-1.37)	
>27.0	29.2	8158	35	1.06 (0.64-1.77)	12	1.07 (0.46-2.47)	13	1.30 (0.54-3.13)	10	0.85 (0.35-2.07)	0.985
P-trend				0.652		0.561		0.395		0.513	
per 5 kg/m^2		34228	127	1.10 (0.89-1.38)	37	1.12 (0.75-1.66)	51	1.30 (0.94-1.80)	39	0.90 (0.59-1.33)	0.428

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), height (cm; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes; no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Median BMI per quartile based on the subcohort.

Table 3 | Multivariable-adjusted HRs[a] and 95% CIs for associations between lower body clothing size and Warburg-subtypes in CRC, by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Clothing size											
Men – colon											
≤50	50	10903	220	1.00 (ref.)	67	1.00 (ref.)	74	1.00 (ref.)	79	1.00 (ref.)	
52	52	9750	257	1.30 (1.04-1.63)	80	1.35 (0.94-1.92)	89	1.34 (0.95-1.88)	88	1.23 (0.88-1.72)	
54	54	5156	135	1.30 (0.99-1.70)	29	0.92 (0.57-1.47)	46	1.32 (0.88-1.98)	60	1.61 (1.10-2.37)	
≥56	56	2618	90	1.74 (1.27-2.38)	24	1.56 (0.94-2.60)	35	1.98 (1.26-3.11)	31	1.66 (1.04-2.65)	0.344
P-trend				0.001		0.299		0.006		0.006	
per 2 sizes		28428	702	1.31 (1.14-1.49)	200	1.18 (0.96-1.47)	244	1.45 (1.18-1.78)	258	1.28 (1.06-1.56)	0.292
Men – rectum											
≤50	50	10903	77	1.00 (ref.)	28	1.00 (ref.)	28	1.00 (ref.)	25	1.00 (ref.)	
52	52	9750	69	1.01 (0.71-1.43)	20	0.81 (0.45-1.46)	20	0.81 (0.45-1.46)	20	0.89 (0.49-1.63)	
54	54	5156	46	1.31 (0.88-1.95)	17	1.30 (0.69-2.43)	17	1.30 (0.69-2.43)	14	1.22 (0.62-2.42)	
≥56	56	2618	17	0.97 (0.55-1.70)	6	0.89 (0.36-2.21)	6	0.89 (0.36-2.21)	8	1.41 (0.62-3.24)	0.872
P-trend				0.488		0.773		0.773		0.384	
per 2 sizes		28428	209	1.02 (0.84-1.25)	71	0.91 (0.65-1.29)	71	0.91 (0.65-1.29)	67	1.13 (0.81-1.59)	0.690
Women – colon											
≤40	40	6574	126	1.00 (ref.)	37	1.00 (ref.)	42	1.00 (ref.)	47	1.00 (ref.)	
42	42	8582	162	0.97 (0.73-1.28)	32	0.64 (0.39-1.06)	67	1.20 (0.79-1.83)	63	1.02 (0.68-1.54)	
44	44	9270	167	0.89 (0.68-1.17)	37	0.65 (0.40-1.06)	54	0.85 (0.55-1.31)	76	1.12 (0.75-1.66)	
≥46	46	9454	188	1.00 (0.76-1.32)	61	1.08 (0.69-1.67)	48	0.74 (0.47-1.15)	79	1.18 (0.79-1.76)	0.018
P-trend				0.907		0.493		0.044		0.350	
per 2 sizes		33880	643	1.09 (0.95-1.24)	167	1.27 (0.96-1.69)	211	0.84 (0.69-1.02)	265	1.20 (1.01-1.42)	0.007

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
≤40	40	6574	21	1.00 (ref.)	3	1.00 (ref.)	9	1.00 (ref.)	9	1.00 (ref.)	
42	42	8582	29	0.99 (0.55-1.80)	11	2.56 (0.69-9.47)	12	0.98 (0.40-2.39)	6	0.48 (0.16-1.44)	
44	44	9270	36	1.14 (0.65-2.00)	14	3.30 (0.98-11.2)	10	0.75 (0.29-1.91)	12	0.86 (0.35-2.08)	
≥46	46	9454	40	1.19 (0.68-2.08)	9	1.88 (0.53-6.70)	19	1.37 (0.59-3.21)	12	0.81 (0.32-2.04)	0.546
P-trend				0.442		0.426		0.507		0.983	
per 2 sizes		33880	126	1.04 (0.84-1.27)	37	1.06 (0.74-1.52)	50	1.04 (0.75-1.45)	39	1.02 (0.71-1.47)	0.998

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.
[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes; no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.
[b]Median clothing size per category based on the subcohort.

Table 4 | Multivariable-adjusted HRs[a] and 95% CIs for associations between non-occupational physical activity and Warburg-subtypes in CRC, by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Non-occupational physical activity (min/day)											
Men – colon											
≤30	21.4	4997	131	1.10 (0.85-1.43)	36	1.28 (0.81-2.00)	48	0.97 (0.67-1.41)	47	1.13 (0.76-1.68)	
31-60	42.9	10100	247	1.00 (ref.)	61	1.00 (ref.)	100	1.00 (ref.)	86	1.00 (ref.)	
61-90	73.6	6001	164	1.16 (0.90-1.48)	53	1.54 (1.02-2.32)	51	0.88 (0.60-1.28)	60	1.21 (0.85-1.75)	
>90	130.0	9925	230	0.96 (0.77-1.20)	65	1.11 (0.76-1.62)	81	0.83 (0.61-1.15)	84	1.01 (0.72-1.40)	0.611
P-trend				0.505		0.917		0.270		0.812	
per 30 min/day		31022	772	0.99 (0.95-1.03)	215	0.99 (0.93-1.06)	280	0.98 (0.92-1.05)	277	0.99 (0.93-1.05)	0.987
Men – rectum											
≤30	21.4	4997	19	0.57 (0.33-0.97)	10	0.90 (0.42-1.93)	9	0.77 (0.35-1.71)	0	-	
31-60	42.9	10100	70	1.00 (ref.)	24	1.00 (ref.)	25	1.00 (ref.)	21	1.00 (ref.)	
61-90	73.6	6001	58	1.41 (0.96-2.05)	21	1.51 (0.83-2.75)	14	0.96 (0.48-1.91)	23	1.85 (1.00-3.43)	
>90	130.0	9925	80	1.20 (0.85-1.69)	21	0.90 (0.50-1.65)	28	1.16 (0.66-2.04)	31	1.60 (0.90-2.86)	0.099
P-trend				0.004		0.882		0.332		<0.001	
per 30 min/day		31022	227	1.05 (0.99-1.10)	76	0.97 (0.88-1.06)	76	1.05 (0.96-1.16)	75	1.10 (1.02-1.19)	0.089
Women – colon											
≤30	19.3	7756	177	1.17 (0.92-1.49)	51	1.31 (0.87-1.96)	62	1.23 (0.85-1.78)	64	1.04 (0.73-1.48)	
31-60	42.9	10923	212	1.00 (ref.)	53	1.00 (ref.)	73	1.00 (ref.)	86	1.00 (ref.)	
61-90	75.0	8000	148	0.94 (0.73-1.20)	37	0.94 (0.61-1.47)	46	0.84 (0.56-1.25)	65	1.02 (0.72-1.44)	
>90	115.7	7550	118	0.81 (0.62-1.05)	29	0.80 (0.50-1.28)	35	0.69 (0.45-1.06)	54	0.91 (0.63-1.32)	0.704
P-trend				0.008		0.042		0.006		0.574	
per 30 min/day		34228	655	0.96 (0.90-1.02)	170	0.96 (0.86-1.08)	216	0.87 (0.78-0.96)	269	1.02 (0.94-1.11)	0.050

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
≤30	19.3	7756	30	0.98 (0.60-1.61)	6	0.77 (0.27-2.17)	13	1.12 (0.53-2.36)	11	1.00 (0.45-2.24)	
31-60	42.9	10923	42	1.00 (ref.)	11	1.00 (ref.)	16	1.00 (ref.)	15	1.00 (ref.)	
61-90	75.0	8000	33	1.05 (0.65-1.68)	13	1.53 (0.69-3.43)	13	1.09 (0.51-2.32)	7	0.64 (0.26-1.58)	
>90	115.7	7550	22	0.73 (0.42-1.25)	7	0.83 (0.30-2.24)	9	0.83 (0.36-1.92)	6	0.56 (0.21-1.48)	0.919
P-trend				0.352		0.649		0.569		0.143	
per 30 min/day		34228	127	1.01 (0.89-1.14)	37	1.06 (0.86-1.30)	51	0.91 (0.77-1.08)	39	1.05 (0.84-1.32)	0.329

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.
[a]Hazard Ratios were adjusted for age (years; continuous), BMI (kg/m^2), total energy intake (kcal/day; continuous), family history of CRC (yes; no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.
[b]Median daily minutes of physical activity per category based on the subcohort.

Table 5 | Multivariable-adjusted HRs[a] and 95% CIs for associations between occupational physical activity and Warburg-subtypes in CRC in men, by tumor location, NLCS 1986-2003.

	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Colon										
Energy expenditure	25073	574		157		212		205		
< 8 kJ/minute	15144	364	1.00 (ref.)	100	1.00 (ref.)	135	1.00 (ref.)	129	1.00 (ref.)	
8-12 kJ/minute	6368	138	0.88 (0.70-1.12)	40	0.97 (0.65-1.46)	49	0.83 (0.58-1.19)	49	0.87 (0.60-1.25)	
>12 kJ/minute	3561	72	0.79 (0.58-1.08)	17	0.75 (0.43-1.29)	28	0.78 (0.50-1.22)	27	0.83 (0.52-1.33)	0.972
P-trend			0.107		0.355		0.202		0.356	
Sitting time	25073	574		157		212		205		
>6 hours/day	6511	184	1.00 (ref.)	53	1.00 (ref.)	71	1.00 (ref.)	60	1.00 (ref.)	
2-6 hours/day	11617	249	0.72 (0.57-0.92)	61	0.62 (0.42-0.92)	89	0.66 (0.47-0.93)	99	0.90 (0.63-1.28)	
<2 hours/day	6944	141	0.69 (0.53-0.91)	43	0.78 (0.50-1.21)	52	0.64 (0.43-0.94)	46	0.69 (0.45-1.06)	0.617
P-trend			0.007		0.264		0.025		0.087	
Rectum										
Energy expenditure	25073	184		66		60		58		
< 8 kJ/minute	15144	106	1.00 (ref.)	42	1.00 (ref.)	34	1.00 (ref.)	30	1.00 (ref.)	
8-12 kJ/minute	6368	56	1.33 (0.94-1.90)	16	1.00 (0.55-1.79)	19	1.40 (0.76-2.57)	21	1.69 (0.95-3.00)	
>12 kJ/minute	3561	22	0.88 (0.54-1.44)	8	0.80 (0.36-1.77)	7	0.90 (0.38-2.12)	7	0.94 (0.41-2.14)	0.734
P-trend			0.841		0.630		0.799		0.523	
Sitting time	25073	184		66		60		58		
>6 hours/day	6511	57	1.00 (ref.)	24	1.00 (ref.)	16	1.00 (ref.)	17	1.00 (ref.)	
2-6 hours/day	11617	71	0.67 (0.46-0.97)	24	0.53 (0.29-0.95)	28	0.92 (0.49-1.75)	19	0.61 (0.32-1.18)	
<2 hours/day	6944	56	0.92 (0.62-1.37)	18	0.72 (0.39-1.35)	16	0.93 (0.45-1.92)	22	1.19 (0.63-2.23)	0.414
P-trend			0.714		0.308		0.841		0.568	

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study.
[a]Hazard Ratios were adjusted for age (years; continuous), BMI (kg/m2), total energy intake (kcal/day; continuous), family history of CRC (yes; no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

DISCUSSION

The role of metabolic reprogramming, in particular the Warburg-effect, in cancer development is becoming increasingly recognized[22,23]. We investigated whether the associations between energy balance-related factors and colorectal cancer risk differ between tumors expressing low versus high levels of proteins involved in the Warburg-effect. In this prospective cohort study, we found positive associations for BMI and clothing size with risk of Warburg-moderate and Warburg-high colon cancer in men. In women, clothing size showed positive associations with Warburg-low and Warburg-high colon cancer. Non-occupational physical activity was inversely associated with Warburg-low and Warburg-moderate colon cancer in women, and occupational sitting time was inversely associated with Warburg-moderate colon cancer in men. In contrast, non-occupational physical activity was positively associated with Warburg-high rectal cancer in men. Statistically significant heterogeneity between Warburg-subtypes was found for clothing size and non-occupational physical activity and risk of colon cancer in women, whereas differences found in men did not show statistically significant heterogeneity.

Previous studies classifying colorectal cancer into so-called metabolic subtypes[40] or specifically glycolysis-related subtypes[41] focused mainly on prognosis. Up to now, there are no studies relating etiologic research to metabolic/glycolysis/Warburg subtypes in colorectal cancer. However, studies have investigated the relationship between proposed precursors (i.e., leptin, adiponectin, insulin, IGF-1) of the Warburg effect in relation to colorectal cancer development[12-14]. It has been suggested that leptin, adiponectin, insulin, and IGF-1 can influence PI3K/Akt signaling[16], leading to the metabolic switch towards the Warburg effect[17,18]. Leptin, insulin, and IGF-1 have been associated with increased risk of colorectal cancer[12-14], whereas adiponectin has been associated with a decreased colorectal cancer risk[14]. However, results of these associations are rather inconsistent, which is likely caused by differences in study design[12,14], because it is difficult to differentiate between cause and effect with circulating biomarkers, especially when biomarker levels in the serum are measured at the time of cancer diagnosis (e.g., case–control design). The current study, however, has a prospective cohort design, where the proposed mechanism is measured in tumor tissue instead of serum, hereby adding further insights in the link between energy balance and colorectal cancer risk.

The observed differences in associations between energy balance–related factors and Warburg subtypes in colorectal cancer, further varying by sex and tumor location, suggest that various carcinogenic mechanisms may explain the link between energy balance and colorectal cancer.

ADIPOSITY

The current results suggest a role of the Warburg-effect in colon cancer risk enhancement resulting from increased adiposity in both men and women (associations Warburg-high), but an additional mechanism might be involved in women (associations Warburg-low). A possible explanation for this additional mechanism in women might

be related to sex hormones. Gunter and colleagues[42] have previously reported a positive association between estradiol (an endogenous estrogen) levels and colorectal cancer risk in postmenopausal women, potentially through promotion of cancer cell proliferation by estradiol, as well as for hyperinsulinemia and IGF-1 with colorectal cancer risk. They proposed that at least two pathways are related to the obesity-colorectal cancer link in postmenopausal women, one involving estradiol and one involving hyperinsulinemia and IGF-1 signaling. These proposed pathways might support the current results, where the pathway involving insulin and IGF-1 reflects the associations with the Warburg-high subtype. The associations we found for the Warburg-low subtype, which are thus not linked to the Warburg effect, might be explained by aberrant signaling of sex hormones.

Mutually adjusted BMI and clothing size models might give some further insight in the associations between adiposity and Warburg-subtypes. Although we do not have data on visceral adipose tissue (VAT) or subcutaneous adipose tissue (SAT), it has been shown that waist circumference could be used as a surrogate of VAT, especially when BMI is added to the model, and that BMI better predicts SAT in men, especially when adjusting for waist circumference[37]. The difference we found between Warburg-moderate and Warburg-high subtypes in men with colon cancer in mutually adjusted models of BMI and clothing size might be explained by differences in fat distribution. VAT seems to be an indicator of adiponectin levels, whereas SAT influences leptin levels[43]. The role of adiponectin in the enhancement of the Warburg effect is less established compared with that of leptin[5]. This suggests that adiponectin might not be associated with the Warburg-effect, which might possibly explain why clothing size associations with Warburg-high colon cancer were attenuated after adjustment for BMI (proxy VAT), whereas BMI associations did not change after adjustment for clothing size (proxy SAT), and vice versa for Warburg-moderate. Though we should be careful with drawing any conclusions because we only have proxy measures for VAT and SAT, we believe these differences are interesting and studies with well-defined data on fat distribution are required to further investigate this. For women, we refrain from interpreting results from mutually adjusted BMI and clothing size models, since the changes in HRs might be caused by multicollinearity (for comparison, the Spearman rank correlation was 0.53 for men and 0.76 for women).

PHYSICAL ACTIVITY

Interestingly, both higher levels of non-occupational physical activity in women and less occupational sitting time in men were inversely associated with Warburg-low and -moderate colon cancer. These observations might indicate that the Warburg effect is not involved, at least not to a great extent, in the relation between measures of physical activity and colon cancer. Potentially, some of the proteins of our Warburg panel are involved in the association, but not all. Very few studies have investigated whether any of these proteins are involved in the etiologic link between physical activity and colon cancer. Shirvani and colleagues[44] found increased levels of the tumor-suppressor P53 in mice after exercise training. However, Slattery and colleagues[45] did not find a difference in associations of physical activity for tumors with and

without a P53 mutation in a case–control study. Further research, preferably in large prospective cohorts, is required to examine which markers might be involved in this link. In women, inverse associations for non-occupational physical activity were found for Warburg-low colon cancer as well, indicating that the Warburg effect is not involved. A possible explanation for this association might be the decreased exposure to fecal carcinogens of the colonical mucosal surface, due to a shorter bowel transit time[4]. This is further underlined by the results of Song and colleagues[46], who found a reduction in transit time for highly physically active women, but not men.

We observed a positive association for physical activity and risk of Warburg-high rectal cancer in men. Positive, though non-significant, associations have previously been reported for rectal cancer within the NLCS[34], as well as other studies[47-49], but results from metaanalyses suggest that physical activity is not related to rectal cancer in men[2,4]. For now, we do not have an explanation for this counterintuitive finding. Further research is necessary to see whether this relation will be observed in other populations.

STRENGTHS AND LIMITATIONS

A major strength of this study is the large prospective population-based cohort design with long follow-up (20.3 years) and availability of tumor material from a large number of incident colorectal cancer cases. This enabled us to combine extensive epidemiologic lifestyle data with molecular pathologic profiling of tumor tissue for a large number of incident colorectal cancer cases. However, despite the large sample size, the number of cases in final statistical analyses was limited for some groups (especially rectal cancer) due to heterogeneity in sex and tumor location. In particular, the absence of clear differences in associations between Warburg-subtypes in rectal cancer might be caused by a lack of statistical power. Other large prospective cohort studies with availability of tumor specimens might help to further investigate the mechanisms for rectal cancer, as well as confirming our findings for colon cancer. Furthermore, due to the heterogeneity based on sex and tumor location, multiple testing might have been a problem in the current study. Even though our analyses were hypothesis-driven, this is a common issue in MPE studies[50]. It is therefore important that our analyses will be replicated in large (MPE) studies. Another common problem in MPE studies is selection bias based on referral hospital[50]. However, for the current study, FFPE blocks of incident cancer cases within the NLCS were collected from 43 hospitals throughout the Netherlands, both academic and peripheral, minimizing the risk of selection bias. A third problem with MPE studies is the usage of TMAs instead of full sections. Even though the TMA-technique enables large-throughput analyses at reduced costs[51], it may not provide a full picture of the tumor. Still, in the construction of TMAs, cores were sampled in different regions to capture potential tumor heterogeneity. In addition, IHC scoring on these TMA sections is often performed by nonpathologists (e.g., PhD students or technicians), because the number of cores are often very large in MPE studies, making it impossible for a pathologist to score all available material. However, we have previously shown that non-pathologists can produce reproducible IHC-scoring results, similar to those of a pathologist, after sufficient training by an experienced pathologist[29].

To enable replication of the current results, we used a transparent way of making subtypes by using a simple sum score of six proteins involved in the Warburg effect. We acknowledge that this also entails some disadvantages, as it probably does not reflect all factors involved in the Warburg effect. For example, a case classified as Warburg-low might still show high expression for one of the proteins, whereas a case classified as Warburg-high might show high expression in only half of the proteins. This might be the reason why several associations were found for either Warburg-low and Warburg-moderate, or Warburg-high and Warburg-moderate. By using a Warburg-moderate group here, we were able to compare the more extreme cases and reduce misclassification in Warburg-low and Warburg-high subtypes. In addition, the six proteins used for the sum score and subsequent Warburg subtypes are a selection of the total pathway, and might not capture the full picture. However, capturing the complete pathway is nearly impossible considering time and budgetary constraints. We aimed to capture the Warburg effect by incorporating proteins from different levels of the pathway in this sum score (i.e., from upstream regulators, glucose import, glycolysis, to lactate secretion), attempting to provide a comprehensive view of the Warburg-effect.

Because we were the first to study the associations between energy balance–related factors and risk of Warburg subtypes in colorectal cancer, our results should be interpreted with caution because validation of the current findings is needed.

CONCLUSION

In this large prospective cohort study, we found that associations of energy balance–related factors and colorectal cancer risk differed between Warburg subtypes, further varying by sex and tumor location. The Warburg effect seems to be involved in associations between adiposity and risk of colon cancer, both in men and women, though additional mechanisms are probably at play in women as well. The link between physical activity and colon cancer is probably explained by mechanisms other than the Warburg-effect. Further research is needed to reproduce these results, and investigate possible additional mechanisms.

REFERENCES

1. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiol Biomarkers Prev 2007;16:2533–47.
2. Robsahm TE, Aagnes B, Hjarta ker A, Langseth H, Bray FI, Larsen IK. Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. Eur J Cancer Prev 2013;22:492–505.
3. Wolin KY, Yan Y, Colditz GA, Lee I. Physical activity and colon cancer prevention: a meta-analysis. Br J Cancer 2009;100:611–6.
4. Samad A, Taylor R, Marshall T, Chapman MA. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Colorectal Dis 2005;7:204–13.
5. Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer - mechanisms underlying tumour progression and recurrence. Nat Rev Endocrinol 2014;10:455–65.
6. Hopkins BD, Goncalves MD, Cantley LC. Obesity and cancer mechanisms: cancer metabolism J Clin Oncol 2016;34:4277–83.
7. Friedenreich CM, Ryder-Burbidge C, McNeil JJMO. Physical activity, obesity and sedentary behavior in cancer etiology: epidemiologic evidence and biologic mechanisms. Mol Oncol 2021;15:790–800.
8. Fasshauer M, Bluher M. Adipokines in health and disease. Trends Pharmacol Sci 2015;36:461–70.
9. Thomas RJ, Kenfield SA, Jimenez A. Exercise-induced biochemical changes and their potential influence on cancer: a scientific review. Br J Sports Med 2017;51:640–4.
10. Yu N, Ruan Y, Gao X, Sun J. Systematic review and meta-analysis of randomized, controlled trials on the effect of exercise on serum leptin and adiponectin in overweight and obese individuals. Horm Metab Res 2017;49:164–73.
11. Renehan AG, Frystyk J, Flyvbjerg A. Obesity and cancer risk: the role of the insulin–IGF axis. Trends Endocrinol Metab 2006;17:328–36.
12. Yin S, Bai H, Jing D. Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients: a systemic review and meta-analysis. Diagn Pathol 2014;9:91.
13. Renehan AG, Zwahlen M, Minder C, T O'Dwyer S, Shalet SM, Egger M. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. Lancet 2004;363:1346–53.
14. Joshi RK, Kim WJ, Lee S-A. Association between obesity-related adipokines and colorectal cancer: a case-control study and meta-analysis. World J Gastroenterol 2014;20:7941–9.
15. Yoon YS, Kwon AR, Lee YK, Oh SW. Circulating adipokines and risk of obesity related cancers: a systematic review and meta-analysis. Obes Res Clin Pract 2019;13:329–39.
16. Huang XF, Chen JZ. Obesity, the PI3K/Akt signal pathway and colon cancer. Obes Rev 2009;10:610–6.
17. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science 2010;330:1340–4.
18. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiother Oncol 2009;92:329–33.
19. Pham D-V, Park P-H. Tumor metabolic reprogramming by adipokines as a critical driver of obesity-associated cancer progression. Int J Mol Sci 2021;22:1444.
20. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nat Rev Cancer 2011;11:85–95
21. Warburg O. The metabolism of carcinoma

cells. Cancer Res 1925;9:148–63.
22. Schwartz L, Supuran CT, Alfarouk KO. The Warburg effect and the hallmarks of cancer. Anticancer Agents Med Chem 2017;17:164–70.
23. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011; 144:646–74.
24. van den Brandt PA, Goldbohm RA, Veer PVT, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990;43:285–95.
25. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986;73:1–11.
26. van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol 1990;19:553–8.
27. Goldbohm RA, van den Brandt PA, Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data; 1994.
28. van den Brandt PA. Molecular pathological epidemiology of lifestyle factors and colorectal and renal cell cancer risk. maastricht pathology 2018. 11th joint meeting of the british division of the international academy of pathology and the pathological society of great Britain & Ireland, 19–22 June 2018. J Pathol 2018; 246:S9.
29. Jenniskens JCA, Offermans K, Samarska I, Fazzi GE, Simons CCJM, Smits KM, et al. Validity and reproducibility of immunohistochemical scoring by trained non-pathologists on tissue microarrays. Cancer Epidemiol Biomarkers Prev 2021;30:1867–74.
30. Resnick MB, Routhier J, Konkin T, Sabo E, Pricolo VE. Epidermal growth factor receptor, c-MET, b-catenin, and p53 expression as prognostic indicators in stage II colon cancer: a tissue microarray study. Clin Cancer Res 2004;10:3069–75.
31. Richman SD, Adams R, Quirke P, Butler R, Hemmings G, Chambers P, et al. Pretrial inter-laboratory analytical validation of the FOCUS4 personalised therapy trial. J Clin Pathol 2016;69:35–41.
32. Cairns RA. Drivers of the Warburg phenotype. Cancer J 2015;21:56–61.
33. Hughes LA, Schouten LJ, Goldbohm RA, van den Brandt PA, Weijenberg MP. Self-reported clothing size as a proxy measure for body size. Epidemiology 2009;20:673–6.
34. Simons CCJM, Hughes LA, Van Engeland M, Goldbohm RA, van den Brandt PA, Weijenberg MP. Physical activity, occupational sitting time, and colorectal cancer risk in the Netherlands cohort study. Am J Epidemiol 2013;177:514–30.
35. Lin D, Wei L. The robust inference for the Cox proportional hazards model J Am Stat Assoc 1989;84:1074–8.
36. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982;69:239–41.
37. Janssen I, Heymsfield SB, Allison DB, Kotler DP, Ross R. Body mass index and waist circumference independently contribute to the prediction of nonabdominal, abdominal subcutaneous, and visceral fat. Am J Clin Nutr 2002;75:683–8.
38. De Vogel S, Bongaerts BW, Wouters KA, Kester AD, Schouten LJ, de Goeij AF, et al. Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. Carcinogenesis 2008;29:1765–73.
39. Wacholder S, Gail MH, Pee D, Brookmeyer R. Alternative variance and efficiency calculations for the case-cohort design. Biometrika 1989;76:117–23.
40. Guinney J, Dienstmann R, Wang X, De Reyniès A, Schlicker A, Soneson C, et al. The

consensus molecular subtypes of colorectal cancer. Nat Med 2015;21:1350–6.
41. Zhu J, Wang S, Bai H, Wang K, Hao J, Zhang J, et al. Identification of five glycolysis-related gene signature and risk score model for colorectal cancer. Front Oncol 2021;11:588811.
42. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Rohan TE, Manson JE, et al. Insulin, insulin-like growth factor-I, endogenous estradiol, and risk of colorectal cancer in postmenopausal women. Cancer Res 2008;68:329–37.
43. Park K-G, Park KS, Kim M-J, Kim H-S, Suh Y-S, Ahn JD, et al. Relationship between serum adiponectin and leptin concentrations and body fat distribution. Diabetes Res Clin Pract 2004;63:135–42.
44. Shirvani H, Bazgir B, Rahimi M, Isanejad A, Samadi M, Sobhani V, et al. Interval exercise training affects expression of MCT-1, GLUT-1, PFK-1 and p53 as key metabolic regulators on azoxymethane-induced experimental colon cancer. 2021. Available from: https://doi.org/10.21203/rs.3.rs-151527/v1.
45. Slattery ML, Curtin K, Ma K, Edwards S, Schaffer D, Anderson K, et al. Diet, activity, and lifestyle associations with p53 mutations in colon tumors. Cancer Epidemiol Biomarkers Prev 2002;11:541–8.
46. Song BK, Cho KO, Jo Y, Oh JW, Kim YS. Colon transit time according to physical activity level in adults. J Neurogastroenterol Motil 2012;18:64–9.
47. Friedenreich C, Norat T, Steindorf K, Boutron-Ruault M-C, Pischon T, Mazuir M, et al. Physical activity and risk of colon and rectal cancers: the European prospective investigation into cancer and nutrition. Cancer Epidemiol Biomarkers Prev 2006;15:2398–407.
48. Lee IM, Paffenbarger RS Jr, Hsieh C. Physical activity and risk of developing colorectal cancer among college alumni. J Natl Cancer Inst 1991;83:1324–9.
49. Tavani A, Braga C, La Vecchia C, Conti E, Filiberti R, Montella M, et al. Physical activity and risk of cancers of the colon and rectum: an Italian case-control study. Br J Cancer 1999;79:1912–6.
50. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut 2011;60:397–411.
51. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–7.

SUPPLEMENTARY METHODS

SCORING PROTOCOLS

Cytoplasmic LDHA expression in the tumor cells was scored as negative/weak cytoplasmic staining; 1-50% of tumor cells with strong cytoplasmic staining; or >50% of tumor cells with strong cytoplasmic staining. The percentage of tumor cells with membranous GLUT1 or MCT4 immunoreactivity was scored as negative; 1-10% positive; 11-50% positive; or >50% positive. Cytoplasmic PKM2 expression in the tumor cells was scored as negative/weak positive; moderate positive; 1-50% strong positive; or >50% strong positive. The percentage of nuclear P53 positive tumor cells was scored as negative (0%); 1-10% positive; 11-50% positive; 51-90% positive; or 91-100% positive. PTEN was scored as negative (no staining in tumor cytoplasm); weak (staining of tumor cytoplasm weaker than adjacent stroma); moderate (similar staining intensity in tumor cytoplasm and adjacent stroma); or strong (staining of tumor cytoplasm stronger than adjacent stroma), irrespective of the percentag e of stained tumor cells.

Supplementary Table S1 | Overview of proteins and their role in the Warburg-effect.

Protein	Full name	Cellular localization	Function	Role in Warburg-effect
PTEN	Phosphatase and tensin homolog	Cytosol, nucleus	Tumor suppressor	**Upstream regulator** PTEN is a regulator of the Warburg-effect via its effect on the PI3K/Akt-signaling pathway. Loss of PTEN tumor suppressor function leads to PI3K/Akt signaling activation, which induces the Warburg-effect via HIF-activation.
TP53	Tumor protein p53	Nucleus	Tumor suppressor	**Upstream regulator** TP53 is a transcriptional factor that regulates several glycolysis related proteins. Loss of TP53 tumor suppressor function leads to an increase of the Warburg-effect, and a decrease of the TCA cycle and oxidative phosphorylation.
GLUT1	Glucose transporter 1	Plasma membrane	Glucose transporter	**Glucose import** GLUT1 facilitates glucose uptake across the plasma membrane. Upregulation of GLUT1 provides the cell with the increased demand of glucose.
PKM2	Pyruvate Kinase M2	Cytosol	Enzyme	**Glycolytic enzyme** PKM2 catalyzes the conversion of phophoenolpyruvate and ADP into pyruvate and ATP, the final step of the glycolytic pathway.
LDHA	Lactate dehydrogenase A	Cytosol	Enzyme	**Glycolytic enzyme** LDHA facilitates conversion of pyruvate into lactate. This conversion is necessary to restore NAD+ levels and to avoid pyruvate accumulation.
MCT4	Monocarboxylate transporter 4	Plasma membrane	Monocarboxylate transporter	**Lactate export** MCT4 facilitates lactate removal across the plasma membrane. By upregulation of MCT4, the cancer cell eliminates the accumulated lactate, thereby preventing acidification of the cell.

Abbreviations: TCA, tricarboxylic acid; ADP, adenosine diphosphate; ATP, adenosine triphosphate; NAD+, nicotinamide adenine dinucleotide +.
References: the Human Protein Atlas (http://www.proteinatlas.org)[1], Feron et al[2], Levine et al[3], Bensinger et al[4].

1. Pontén F, Jirström K, Uhlen M. The Human Protein Atlas—a tool for pathology. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2008;216(4):387-93.
2. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiotherapy and oncology. 2009;92(3):329-33.
3. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330(6009):1340-4.
4. Bensinger SJ, Christofk HR, editors. New aspects of the Warburg effect in cancer cell biology. Seminars in cell & developmental biology; 2012: Elsevier.

Supplementary Table S2 | Overview staining protocols.

Antibody	Clone	Source	Dilution	Staining procedure	Antigen retrieval	Incubation time	Visualisation system	Chromogen
Pan-CK	AE1/AE3	DAKO (GA05361-2)	RTU	DAKO[a]	PT high[b]	10 min.	EnVision FLEX[f]	DAB
TP53	DO-7	DAKO (M700101-2)	RTU	DAKO[a]	PT high[b]	20 min.	EnVision FLEX[f]	DAB
PTEN	6H2.1	DAKO (M362729-2)	1:100	DAKO[a]	PT high[b]	20 min.	EnVision FLEX[f]	DAB
GLUT1	-	TFS (RB-9052-P)	1:200	DAKO[a]	PT low[c]	20 min.	EnVision FLEX[f]	DAB
LDHA	E-9	SCBT (sc-137243)	1:800	Manual	HIER high[d]	Overnight, 4^{O}C	REAL EnVision[g]	DAB
MCT4	D-1	SCBT (sc-376140)	1:100	Manual	HIER high[d]	Overnight, 4^{O}C	REAL EnVision[g]	DAB
PKM2	C-11	SCBT (sc-365684)	1:100	Manual	HIER low[e]	1 hour, 37^{O}C	LSAB2 Kit/HRP[h]	DAB

Abbreviations: TFS, Thermo Fisher Scientific; SBCT, Santa Cruz Biotechnology; RTU, ready to use; DAB, 3,3'-diaminobenzidine
[a]DAKO Autostainer Link 48
[b]High pH retrieval (K8004) for 20 minutes on the Dako PT link (Agilent Technologies)
[c]Low pH retrieval (K8005) for 20 minutes on the Dako PT link (Agilent Technologies)
[d]Heat-induced antigen retrieval using a solution of Tris/EDTA (pH 9.0)
[e]Heat-induced antigen retrieval using a solution of sodium citrate (pH 6.0)
[f]EnVision FLEX Visualization Kit (K8008, DAKO)
[g]REAL EnVision Detection System (K5007, DAKO)
[h]Universal LSAB2 kit/HRP (K0675, Agilent)

Supplementary Table S3 | Percentage of slides evaluated per assessor for the six immunohistochemical markers of proteins incorporated in the Warburg-subtypes.

	Assessor 1 Non-pathologist	**Assessor 2** Non-pathologist		**Assessor 3** Non-pathologist		**Assessor 4** Pathologist[b]
		Time point 1	Time point 2[a]	Time point 1	Time point 2[a]	
TP53	100%	100%	10%	100%	10%	10%
PTEN	100%	100%	10%	100%	10%	10%
GLUT1	25%[c]	100%	10%	100%	10%	10%
LDHA	-	100%	10%	100%	10%	10%
MCT4	-	100%	10%	100%	10%	10%
PKM2	-	100%	10%	100%	10%	10%

[a]10% randomly selected TMA sections per marker were scored for a second time after a period of at least five months to assess intra-observer reproducibility.
[b]10% randomly selected TMA sections per marker were scored by an experienced pathologist to assess inter-observer agreement between pathologist and non-pathologists.
[c]Assessor 1 left the project early because of an unforeseen work relocation.

Supplementary Table S4 | Scoring protocols of the six immunohistochemical stainings of proteins incorporated in the Warburg-subtypes, and kappa values with 95% confidence intervals for inter- and intra-observer agreement of these scoring protocols.

	P53 Nucleus	**PTEN** Cytoplasm[a]	**GLUT1** Membrane	**LDHA** Cytoplasm	**MCT4** Membrane	**PKM2** Cytoplasm
Scoring protocol[b]						
Low						
Category 1	(1) negative	(1) negative	(1) negative	(1) negative/weak	(1) negative	(1) negative/weak
Category 2	(2) 1-10% positive		(2) 1-10% positive		(2) 1-10% positive	
Moderate						
Category 2		(2) weak		(2) 1-50% strong positive		(2) moderate positive
Category 3	(3) 11-50% positive	(3) moderate	(3) 11-50% positive		(3) 11-50% positive	(3) 1-50% strong positive
High						
Category 3				(3) >50% strong positive		
Category 4	(4) 51-90% positive	(4) strong	(4) >50% positive		(4) >50% positive	(4) >50% strong positive
Category 5	(5) >90% positive					

This table continues on the next page

(continued)	**P53** Nucleus	**PTEN** Cytoplasm[b]	**GLUT1** Membrane	**LDHA** Cytoplasm	**MCT4** Membrane	**PKM2** Cytoplasm
Scoring agreement	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**
Interobserver agreement[c]						
Final score[d] vs pathologist						
Weighted kappa[e]	0.75 (0.72-0.79)	0.58 (0.53-0.62)	0.71 (0.67-0.74)	0.65 (0.60-0.69)	0.74 (0.71-0.77)	0.65 (0.61-0.69)
Non-weighted kappa	0.63 (0.58-0.68)	0.47 (0.41-0.52)	0.61 (0.57-0.66)	0.59 (0.54-0.64)	0.63 (0.59-0.68)	0.56 (0.51-0.60)
Intraobserver agreement[c,f]						
Non-pathologist assessor 1						
Weighted kappa[e]	0.83 (0.80-0.86)	0.69 (0.65-0.74)	0.82 (0.79-0.85)	0.78 (0.74-0.82)	0.86 (0.83-0.88)	0.70 (0.67-0.74)
Non-weighted kappa	0.73 (0.69-0.77)	0.63 (0.58-0.69)	0.75 (0.72-0.79)	0.76 (0.71-0.80)	0.79 (0.75-0.82)	0.58 (0.53-0.62)
Non-pathologist assessor 2						
Weighted kappa[e]	0.87 (0.84-0.90)	0.69 (0.64-0.74)	0.75 (0.72-0.78)	0.77 (0.73-0.81)	0.83 (0.81-0.86)	0.72 (0.68-0.75)
Non-weighted kappa	0.80 (0.76-0.84)	0.65 (0.60-0.70)	0.65 (0.60-0.69)	0.73 (0.69-0.78)	0.75 (0.71-0.79)	0.62 (0.58-0.67)

[a]For PTEN scoring, staining intensity of tumour cells was compared with that of stromal cells: (1) no PTEN staining in the tumour cells; (2) staining intensity in the tumour cells weaker than in the stromal cells; (3) similar staining intensity in tumour and stromal cells; (4) staining intensity in the tumour cells stronger than in the stromal cells.
[b]Only immunoreactivity in reported cellular localization was considered positive staining.
[c]Based on a random 10% of TMA sections
[d]The final score is based on at least two non-pathologists, with discrepancies replaced by a consensus score or pathologist's score.
[e]Weight of 0.5 for adjacent categories and 0 for non-adjacent categories.
[f]10% of TMA sections were scored for a second time after at least 2 months.

Supplementary Table S5 | Clinical characteristics [n(%)] of CRC cases, stratified on Warburg-subtypes; NLCS, 1986-2006.

	Men				Women			
	Total CRC	Warburg low	Warburg moderate	Warburg high	Total CRC	Warburg low	Warburg moderate	Warburg high
N	1131	338	405	388	841	223	287	331
Location								
Colon	772 (68.3)	215 (63.6)	280 (69.1)	277 (71.4)	655 (77.9)	170 (76.2)	216 (75.3)	269 (81.3)
Proximal	350 (31.0)	85 (25.2)	111 (27.4)	154 (39.7)	401 (47.7)	93 (41.7)	130 (45.3)	178 (53.8)
Distal	407 (36.0)	125 (37.0)	166 (41.0)	116 (29.9)	235 (27.9)	71 (31.8)	80 (27.9)	84 (25.4)
Rectosigmoid	132 (11.7)	47 (13.9)	49 (12.1)	36 (9.3)	59 (7.0)	16 (7.2)..	20 (7.0)..	23 (7.0)..
Rectum	227 (20.1)	76 (22.5)	76 (18.8)	75 (19.3)	127 (15.1)	37 (16.6)	51 (17.8)	39 (11.8)
TNM stage								
I	234 (20.7)	88 (26.0)	87 (21.5)	59 (15.2)	147 (17.5)	48 (21.5)	54 (18.8)	45 (13.6)
II	427 (37.8)	121 (35.8)	159 (39.3)	147 (37.9)	317 (37.7)	89 (39.9)	93 (32.4)	135 (40.8)
III	296 (26.2)	85 (25.2)	100 (24.7)	111 (28.6)	225 (26.8)	48 (21.5)	87 (30.3)	90 (27.2)
IV	145 (12.8)	35 (10.4)	48 (11.9)	62 (16.0)	132 (15.7)	31 (13.9)	46 (16.0)	55 (16.6)
Unknown	29 (2.6)	9 (2.7)	11 (2.7)	9 (2.3)	20 (2.4)	7 (3.1)	7 (2.4)	6 (1.8)
Differentiation grade								
Poor/undifferentiated	155 (13.7)	34 (10.1)	50 (12.4)	71 (18.3)	183 (21.8)	33 (14.8)	63 (22.0)	87 (26.3)
Moderate	782 (69.1)	236 (69.8)	285 (70.4)	261 (67.3)	526 (62.5)	151 (67.7)	170 (59.2)	205 (61.9)
Well	99 (8.8)	41 (12.1)	36 (8.9)	22 (5.7)	60 (7.1)	18 (8.1)	24 (8.4)	18 (5.4)
Unknown	95 (8.4)	27 (8.0)	34 (8.4)	34 (8.8)	72 (8.6)	21 (9.4)	30 (10.5)	21 (6.3)

Abbreviations: CRC, colorectal cancer; NLCS, Netherlands Cohort Study; TNM: tumor node metastasis

Supplementary Table S6 | Age-adjusted HRs[a] and 95% CIs for associations between BMI and Warburg-subtypes in CRC, by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
BMI quartiles (kg/m²)										
Men – colon										
< 23.4	22.2	7993	174	1.00 (ref.)	57	1.00 (ref.)	57	1.00 (ref.)	60	1.00 (ref.)
23.4-24.9	24.2	8343	199	1.08 (0.84-1.37)	55	0.91 (0.61-1.34)	81	1.34 (0.93-1.93)	63	0.99 (0.68-1.44)
25.0-26.6	25.7	7683	203	1.20 (0.94-1.53)	52	0.93 (0.62-1.39)	63	1.14 (0.78-1.67)	88	1.50 (1.05-2.14)
> 26.6	27.8	7003	196	1.28 (1.00-1.64)	51	1.01 (0.67-1.52)	79	1.58 (1.09-2.28)	66	1.25 (0.86-1.82)
P-trend				0.034		0.945		0.044		0.056
per 5 kg/m²		31022	772	1.22 (1.06-1.40)	215	1.03 (0.81-1.32)	280	1.24 (1.01-1.53)	277	1.35 (1.09-1.67)
Men – rectum										
< 23.4	22.2	7993	56	1.00 (ref.)	20	1.00 (ref.)	20	1.00 (ref.)	16	1.00 (ref.)
23.4-24.9	24.2	8343	53	0.89 (0.60-1.33)	14	0.66 (0.33-1.32)	17	0.80 (0.41-1.55)	22	1.31 (0.68-2.53)
25.0-26.6	25.7	7683	69	1.28 (0.88-1.87)	25	1.30 (0.71-2.37)	26	1.34 (0.74-2.44)	18	1.19 (0.60-2.35)
> 26.6	27.8	7003	49	1.00 (0.67-1.51)	17	0.97 (0.50-1.88)	13	0.75 (0.37-1.52)	19	1.37 (0.70-2.71)
P-trend				0.509		0.578		0.884		0.444
per 5 kg/m²		31022	227	1.07 (0.87-1.33)	76	1.09 (0.75-1.58)	76	1.02 (0.72-1.44)	75	1.11 (0.79-1.56)
Women – colon										
<22.8	21.5	9014	186	1.00 (ref.)	54	1.00 (ref.)	64	1.00 (ref.)	68	1.00 (ref.)
22.8-24.7	23.8	8914	147	0.78 (0.61-1.01)	27	0.49 (0.31-0.80)	51	0.79 (0.53-1.16)	69	1.01 (0.70-1.45)
24.8-27.0	25.7	8141	160	0.94 (0.73-1.21)	41	0.83 (0.54-1.27)	53	0.91 (0.61-1.33)	66	1.07 (0.74-1.53)
>27.0	29.2	8158	162	0.95 (0.74-1.21)	48	0.96 (0.64-1.45)	48	0.81 (0.55-1.21)	66	1.06 (0.73-1.52)
P-trend				0.974		0.749		0.435		0.704
per 5 kg/m²		34228	655	1.03 (0.91-1.16)	170	1.10 (0.89-1.35)	216	0.93 (0.77-1.13)	269	1.06 (0.90-1.25)

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Women – rectum										
<22.8	21.5	9014	35	1.00 (ref.)	11	1.00 (ref.)	12	1.00 (ref.)	12	1.00 (ref.)
22.8-24.7	23.8	8914	26	0.75 (0.44-1.26)	4	0.37 (0.12-1.18)	11	0.92 (0.40-2.11)	11	0.92 (0.40-2.10)
24.8-27.0	25.7	8141	31	0.98 (0.60-1.62)	10	1.02 (0.43-2.43)	15	1.38 (0.64-2.99)	6	0.55 (0.21-1.48)
>27.0	29.2	8158	35	1.10 (0.68-1.79)	12	1.23 (0.54-2.81)	13	1.19 (0.54-2.62)	10	0.91 (0.39-2.15)
P-trend				0.517		0.375		0.457		0.621
per 5 kg/m^2		34228	127	1.13 (0.92-1.39)	37	1.19 (0.83-1.71)	51	1.26 (0.93-1.71)	39	0.91 (0.61-1.37)

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.
[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.
[b]Median BMI per quartile based on the subcohort.

Supplementary Table S7 | Age-adjusted HRs[a] and 95% CIs for associations between lower body clothing size and Warburg-subtypes in CRC, by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Clothing size										
Men – colon										
≤50	50	10903	220	1.00 (ref.)	67	1.00 (ref.)	74	1.00 (ref.)	79	1.00 (ref.)
52	52	9750	257	1.27 (1.02-1.58)	80	1.29 (0.91-1.83)	89	1.32 (0.95-1.85)	88	1.20 (0.87-1.67)
54	54	5156	135	1.26 (0.97-1.64)	29	0.88 (0.56-1.40)	46	1.30 (0.87-1.93)	60	1.55 (1.07-2.24)
≥56	56	2618	90	1.68 (1.24-2.29)	24	1.46 (0.89-2.42)	35	1.98 (1.27-3.07)	31	1.61 (1.02-2.54)
P-trend				<0.001		0.430		0.005		0.008
per 2 sizes		28428	702	1.29 (1.13-1.46)	200	1.15 (0.93-1.41)	244	1.44 (1.18-1.75)	258	1.26 (1.05-1.52)
Men – rectum										
≤50	50	10903	77	1.00 (ref.)	28	1.00 (ref.)	28	1.00 (ref.)	25	1.00 (ref.)
52	52	9750	69	1.01 (0.71-1.42)	20	0.80 (0.45-1.44)	20	1.34 (0.77-2.33)	20	0.91 (0.50-1.66)
54	54	5156	46	1.29 (0.87-1.90)	17	1.29 (0.70-2.40)	17	1.32 (0.69-2.55)	14	1.24 (0.64-2.42)
≥56	56	2618	17	0.95 (0.55-1.66)	6	0.91 (0.37-2.26)	6	0.53 (0.16-1.78)	8	1.42 (0.63-3.21)
P-trend				0.548		0.749		0.831		0.355
per 2 sizes		28428	209	1.01 (0.83-1.23)	71	0.92 (0.66-1.30)	71	1.00 (0.75-1.33)	67	1.14 (0.82-1.60)
Women – colon										
≤40	40	6574	126	1.00 (ref.)	37	1.00 (ref.)	42	1.00 (ref.)	47	1.00 (ref.)
42	42	8582	162	0.97 (0.74-1.27)	32	0.65 (0.40-1.07)	67	1.20 (0.79-1.81)	63	1.01 (0.68-1.52)
44	44	9270	167	0.90 (0.68-1.18)	37	0.68 (0.42-1.09)	54	0.87 (0.56-1.33)	76	1.10 (0.75-1.63)
≥46	46	9454	188	0.98 (0.75-1.28)	61	1.08 (0.70-1.66)	48	0.75 (0.48-1.16)	79	1.12 (0.76-1.65)
P-trend				0.818		0.471		0.050		0.491
per 2 sizes		33880	643	1.08 (0.95-1.23)	167	1.27 (0.96-1.68)	211	0.85 (0.70-1.03)	265	1.17 (0.99-1.38)

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Women – rectum										
≤40	40	6574	21	1.00 (ref.)	3	1.00 (ref.)	9	1.00 (ref.)	9	1.00 (ref.)
42	42	8582	29	1.05 (0.59-1.88)	11	2.87 (0.80-10.3)	12	1.00 (0.42-2.40)	6	0.51 (0.18-1.44)
44	44	9270	36	1.20 (0.69-2.10)	14	3.44 (0.99-11.9)	10	0.76 (0.30-1.90)	12	0.93 (0.39-2.22)
≥46	46	9454	40	1.31 (0.75-2.27)	9	2.22 (0.60-8.16)	19	1.39 (0.62-3.14)	12	0.91 (0.37-2.20)
P-trend				0.275		0.277		0.467		0.809
per 2 sizes		33880	126	1.07 (0.87-1.32)	37	1.10 (0.76-1.60)	50	1.05 (0.77-1.44)	39	1.06 (0.74-1.54)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.
[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.
[b]Median clothing size per category based on the subcohort.

Supplementary Table S8 | Age-adjusted HRs[a] and 95% CIs for associations between non-occupational physical activity and Warburg-subtypes in CRC, by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Non-occupational physical activity (min/day)										
Men – colon										
≤30	21.4	4997	131	1.11 (0.86-1.44)	36	1.25 (0.81-1.94)	48	1.00 (0.69-1.45)	47	1.15 (0.78-1.69)
31-60	42.9	10100	247	1.00 (ref.)	61	1.00 (ref.)	100	1.00 (ref.)	86	1.00 (ref.)
61-90	73.6	6001	164	1.14 (0.89-1.45)	53	1.49 (1.00-2.22)	51	0.87 (0.60-1.25)	60	1.20 (0.83-1.72)
>90	130.0	9925	230	0.92 (0.74-1.15)	65	1.05 (0.72-1.52)	81	0.81 (0.59-1.12)	84	0.96 (0.69-1.34)
P-trend				0.258		0.850		0.167		0.571
per 30 min/day		31022	772	0.98 (0.94-1.02)	215	0.98 (0.92-1.05)	280	0.98 (0.92-1.04)	277	0.98 (0.92-1.04)
Men – rectum										
≤30	21.4	4997	19	0.56 (0.33-0.96)	10	0.87 (0.41-1.86)	9	0.75 (0.35-1.63)	0	-
31-60	42.9	10100	70	1.00 (ref.)	24	1.00 (ref.)	25	1.00 (ref.)	21	1.00 (ref.)
61-90	73.6	6001	58	1.39 (0.96-2.02)	21	1.47 (0.81-2.68)	14	0.95 (0.49-1.84)	23	1.83 (1.00-3.36)
>90	130.0	9925	80	1.18 (0.83-1.66)	21	0.89 (0.49-1.62)	28	1.13 (0.65-1.99)	31	1.58 (0.89-2.80)
P-trend				0.005		0.896		0.353		<0.001
per 30 min/day		31022	227	1.05 (0.99-1.10)	76	0.97 (0.88-1.07)	76	1.05 (0.96-1.16)	75	1.10 (1.02-1.18)
Women – colon										
≤30	19.3	7756	177	1.16 (0.92-1.48)	51	1.34 (0.89-2.00)	62	1.18 (0.82-1.70)	64	1.04 (0.73-1.48)
31-60	42.9	10923	212	1.00 (ref.)	53	1.00 (ref.)	73	1.00 (ref.)	86	1.00 (ref.)
61-90	75.0	8000	148	0.96 (0.75-1.23)	37	0.96 (0.62-1.49)	46	0.87 (0.59-1.28)	65	1.04 (0.73-1.46)
>90	115.7	7550	118	0.81 (0.63-1.05)	29	0.80 (0.50-1.28)	35	0.70 (0.46-1.07)	54	0.91 (0.64-1.32)
P-trend				0.011		0.039		0.014		0.585
per 30 min/day		34228	655	0.96 (0.91-1.02)	170	0.96 (0.85-1.09)	216	0.88 (0.79-0.97)	269	1.02 (0.94-1.11)

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Women – rectum										
≤30	19.3	7756	30	1.01 (0.62-1.64)	6	0.79 (0.29-2.15)	13	1.13 (0.54-2.37)	11	1.03 (0.47-2.28)
31-60	42.9	10923	42	1.00 (ref.)	11	1.00 (ref.)	16	1.00 (ref.)	15	1.00 (ref.)
61-90	75.0	8000	33	1.07 (0.67-1.72)	13	1.59 (0.71-3.55)	13	1.13 (0.54-2.39)	7	0.63 (0.26-1.57)
>90	115.7	7550	22	0.76 (0.45-1.29)	7	0.91 (0.35-2.34)	9	0.84 (0.36-1.92)	6	0.58 (0.22-1.51)
P-trend				0.416		0.516		0.594		0.159
per 30 min/day		34228	127	1.02 (0.89-1.16)	37	1.07 (0.89-1.30)	51	0.91 (0.78-1.08)	39	1.07 (0.84-1.37)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.
[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.
[b]Median daily minutes of physical activity per category based on the subcohort.

Supplementary Table S9 | Age-adjusted HRs[a] and 95% CIs for associations between occupational physical activity and Warburg-subtypes in CRC in men, by tumor location, NLCS 1986-2003.

	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Colon									
Energy expenditure	24037	532		150		204		178	
< 8 kJ/minute	14512	337	1.00 (ref.)	94	1.00 (ref.)	131	1.00 (ref.)	112	1.00 (ref.)
8-12 kJ/minute	6111	128	0.90 (0.71-1.14)	39	0.98 (0.66-1.46)	45	0.82 (0.57-1.17)	44	0.92 (0.63-1.33)
>12 kJ/minute	3414	67	0.82 (0.61-1.11)	17	0.74 (0.43-1.27)	28	0.89 (0.58-1.38)	22	0.80 (0.49-1.30)
P-trend			0.155		0.341		0.402		0.351
Sitting time	24037	532		150		204		178	
>6 hours/day	6240	166	1.00 (ref.)	48	1.00 (ref.)	68	1.00 (ref.)	50	1.00 (ref.)
2-6 hours/day	11139	233	0.76 (0.60-0.96)	60	0.68 (0.45-1.01)	86	0.69 (0.49-0.98)	87	0.94 (0.65-1.36)
<2 hours/day	6658	133	0.74 (0.57-0.97)	42	0.81 (0.52-1.25)	50	0.68 (0.46-1.01)	41	0.75 (0.48-1.16)
P-trend			0.027		0.354		0.056		0.194
Rectum									
Energy expenditure	24037	167		60		54		53	
< 8 kJ/minute	14512	95	1.00 (ref.)	37	1.00 (ref.)	31	1.00 (ref.)	27	1.00 (ref.)
8-12 kJ/minute	6111	51	1.29 (0.90-1.84)	15	0.98 (0.53-1.80)	17	1.29 (0.70-2.38)	19	1.70 (0.94-3.08)
>12 kJ/minute	3414	21	0.93 (0.57-1.53)	8	0.91 (0.42-1.99)	6	0.80 (0.33-1.94)	7	1.11 (0.48-2.56)
P-trend			0.756		0.822		0.938		0.350
Sitting time	24037	167		60		54		53	
>6 hours/day	6240	48	1.00 (ref.)	20	1.00 (ref.)	14	1.00 (ref.)	14	1.00 (ref.)
2-6 hours/day	11139	67	0.77 (0.52-1.14)	23	0.63 (0.34-1.16)	26	1.01 (0.52-1.97)	18	0.71 (0.35-1.45)
<2 hours/day	6658	52	1.01 (0.67-1.54)	17	0.80 (0.41-1.54)	14	0.92 (0.43-1.97)	21	1.41 (0.71-2.80)
P-trend			0.921		0.517		0.830		0.302

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study.
[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

Supplementary Table S10 | Multivariable-adjusted HRs[a] and 95% CIs for associations between BMI and Warburg-subtypes in CRC (additionally adjusted for clothing size), by sex and tumor location; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
BMI quartiles (kg/m²)											
Men – colon											
< 23.4	22.2	7356	160	1.00 (ref.)	54	1.00 (ref.)	52	1.00 (ref.)	54	1.00 (ref.)	
23.4-24.9	24.2	7756	187	0.98 (0.74-1.28)	54	0.87 (0.56-1.35)	73	1.08 (0.72-1.62)	60	0.97 (0.64-1.48)	
25.0-26.6	25.7	7056	185	0.97 (0.72-1.31)	49	0.80 (0.50-1.29)	53	0.76 (0.48-1.21)	83	1.40 (0.90-2.16)	
> 26.6	27.8	6260	170	0.96 (0.69-1.35)	43	0.79 (0.47-1.36)	66	0.93 (0.56-1.55)	61	1.16 (0.70-1.92)	0.067
P-trend				0.823		0.367		0.441		0.257	
per 5 kg/m²		28428	702	1.09 (0.90-1.32)	200	0.98 (0.71-1.34)	244	0.90 (0.67-1.20)	258	1.42 (1.07-1.88)	0.145
Men – rectum											
< 23.4	22.2	7356	52	1.00 (ref.)	19	1.00 (ref.)	19	1.00 (ref.)	14	1.00 (ref.)	
23.4-24.9	24.2	7756	50	0.82 (0.54-1.26)	14	0.70 (0.34-1.46)	16	0.68 (0.33-1.39)	20	1.15 (0.58-2.30)	
25.0-26.6	25.7	7056	66	1.19 (0.78-1.82)	23	1.35 (0.68-2.69)	25	1.17 (0.59-2.32)	18	1.06 (0.52-2.20)	
> 26.6	27.8	6260	41	0.81 (0.48-1.35)	15	1.04 (0.49-2.24)	11	0.52 (0.21-1.30)	15	0.92 (0.38-2.22)	0.709
P-trend				0.910		0.484		0.484		0.796	
per 5 kg/m²		28428	209	1.01 (0.76-1.34)	71	1.06 (0.66-1.70)	71	0.99 (0.61-1.59)	67	0.97 (0.61-1.53)	0.863
Women – colon											
<22.8	21.5	8964	184	1.00 (ref.)	54	1.00 (ref.)	63	1.00 (ref.)	67	1.00 (ref.)	
22.8-24.7	23.8	8874	146	0.78 (0.58-1.05)	27	0.42 (0.24-0.72)	50	1.03 (0.65-1.65)	69	0.92 (0.62-1.38)	
24.8-27.0	25.7	8101	155	0.92 (0.63-1.33)	39	0.59 (0.28-1.24)	52	1.46 (0.83-2.59)	64	0.88 (0.55-1.42)	
>27.0	29.2	7941	158	0.96 (0.57-1.61)	47	0.61 (0.19-1.91)	46	1.83 (0.88-3.84)	65	0.80 (0.42-1.51)	0.256
P-trend				0.987		0.475		0.080		0.499	
per 5 kg/m²		33880	643	1.03 (0.81-1.31)	167	0.93 (0.51-1.69)	211	1.23 (0.88-1.72)	265	0.95 (0.72-1.25)	0.276

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
<22.8	21.5	8964	35	1.00 (ref.)	11	1.00 (ref.)	12	1.00 (ref.)	12	1.00 (ref.)	
22.8-24.7	23.8	8874	26	0.82 (0.45-1.49)	4	0.47 (0.13-1.73)	11	1.02 (0.39-2.68)	11	0.86 (0.34-2.15)	
24.8-27.0	25.7	8101	31	1.13 (0.59-2.18)	10	1.54 (0.43-5.48)	15	1.68 (0.63-4.47)	6	0.49 (0.15-1.55)	
>27.0	29.2	7941	34	1.47 (0.61-3.54)	12	2.63 (0.47-14.9)	12	1.57 (0.44-5.68)	10	0.81 (0.18-3.64)	0.984
P-trend				0.321		0.187		0.302		0.560	
per 5 kg/m^2		33880	126	1.42 (0.96-2.09)	37	1.66 (0.81-3.40)	50	1.83 (1.07-3.14)	39	0.89 (0.43-1.84)	0.438

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), clothing size (size; continuous), non-occupational physical activity (minutes/day; continuous), height (cm; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes; no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Median BMI per quartile based on the subcohort.

Supplementary Table S11 | Multivariable-adjusted HRs[a] and 95% CIs for associations between lower body clothing size and Warburg-subtypes in CRC (additionally adjusted for BMI), by sex and tumor location; NLCS, 1986-2006.

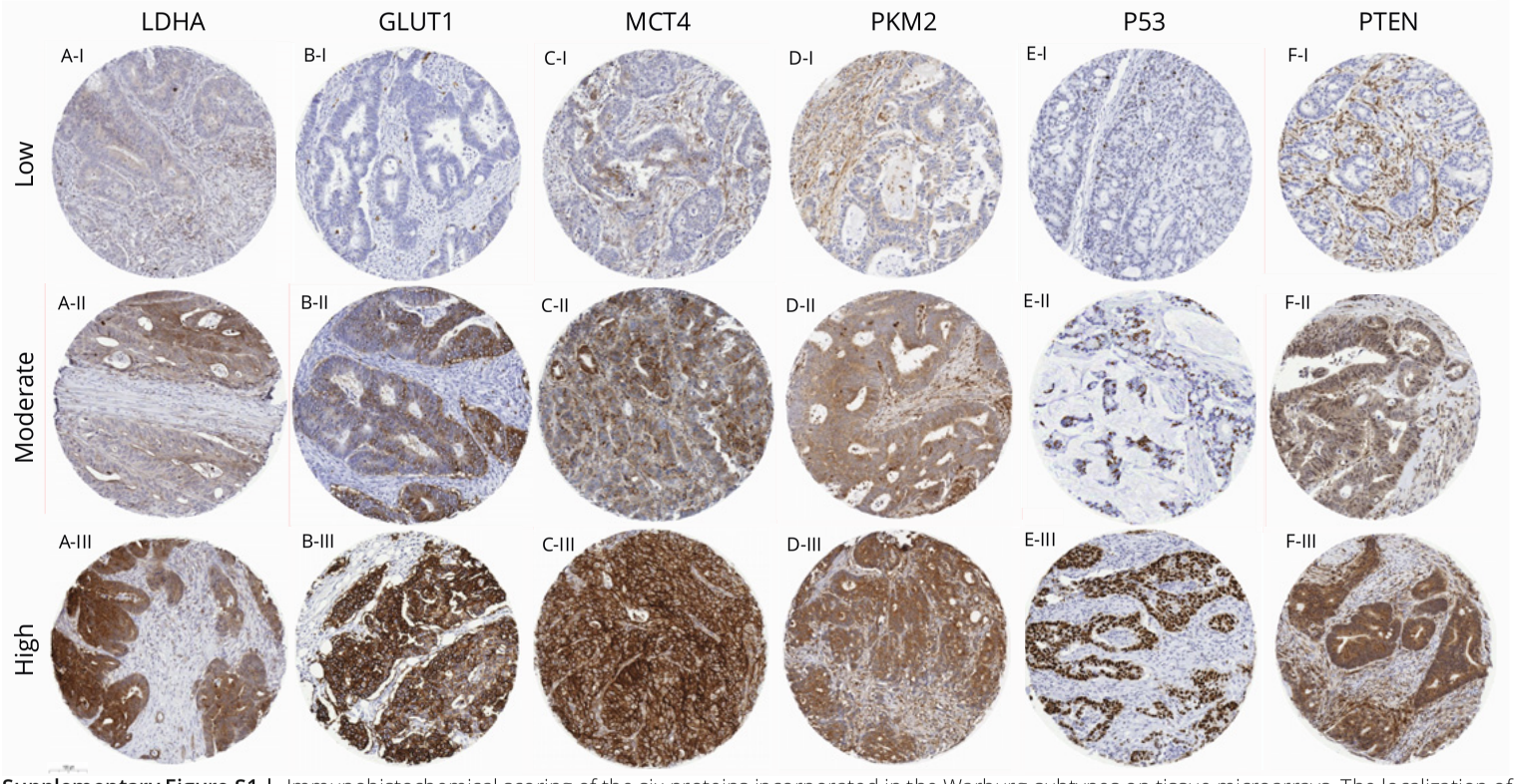
	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Clothing size											
Men – colon											
≤50	50	10903	220	1.00 (ref.)	67	1.00 (ref.)	74	1.00 (ref.)	79	1.00 (ref.)	
52	52	9750	257	1.24 (0.98-1.57)	80	1.34 (0.92-1.95)	89	1.28 (0.89-1.84)	88	1.13 (0.80-1.59)	
54	54	5156	135	1.20 (0.89-1.61)	29	0.91 (0.54-1.52)	46	1.23 (0.79-1.91)	60	1.39 (0.90-2.14)	
≥56	56	2618	90	1.52 (1.06-2.18)	24	1.54 (0.85-2.80)	35	1.76 (1.05-2.95)	31	1.30 (0.76-2.22)	0.345
P-trend				0.040		0.448		0.059		0.166	
per 2 sizes		28428	702	1.24 (1.07-1.45)	200	1.19 (0.93-1.53)	244	1.41 (1.11-1.78)	258	1.15 (0.93-1.42)	0.289
Men – rectum											
≤50	50	10903	77	1.00 (ref.)	28	1.00 (ref.)	28	1.00 (ref.)	25	1.00 (ref.)	
52	52	9750	69	1.02 (0.72-1.45)	20	0.81 (0.46-1.43)	20	1.43 (0.80-2.57)	20	0.88 (0.48-1.61)	
54	54	5156	46	1.34 (0.88-2.03)	17	1.30 (0.66-2.54)	17	1.53 (0.76-3.08)	14	1.21 (0.60-2.44)	
≥56	56	2618	17	1.00 (0.54-1.88)	6	0.89 (0.33-2.44)	6	0.67 (0.17-2.61)	8	1.39 (0.55-3.53)	0.871
P-trend				0.435		0.769		0.700		0.439	
per 2 sizes		28428	209	1.02 (0.83-1.26)	71	0.88 (0.62-1.25)	71	1.11 (0.79-1.57)	67	1.11 (0.78-1.58)	0.694
Women – colon											
≤40	40	6574	126	1.00 (ref.)	37	1.00 (ref.)	42	1.00 (ref.)	47	1.00 (ref.)	
42	42	8582	162	0.99 (0.74-1.32)	32	0.69 (0.41-1.17)	67	1.20 (0.77-1.85)	63	1.02 (0.66-1.56)	
44	44	9270	167	0.92 (0.67-1.26)	37	0.75 (0.44-1.28)	54	0.84 (0.51-1.39)	76	1.11 (0.70-1.75)	
≥46	46	9454	188	1.06 (0.71-1.57)	61	1.38 (0.73-2.59)	48	0.72 (0.37-1.41)	79	1.16 (0.65-2.06)	0.018
P-trend				0.940		0.314		0.177		0.557	
per 2 sizes		33880	643	1.24 (0.99-1.54)	167	1.69 (1.13-2.54)	211	0.85 (0.63-1.14)	265	1.36 (1.06-1.74)	0.007

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
≤40	40	6574	21	1.00 (ref.)	3	1.00 (ref.)	9	1.00 (ref.)	9	1.00 (ref.)	
42	42	8582	29	1.08 (0.57-2.03)	11	2.82 (0.78-10.2)	12	0.96 (0.36-2.54)	6	0.60 (0.18-1.95)	
44	44	9270	36	1.32 (0.69-2.51)	14	3.93 (1.26-12.2)	10	0.71 (0.25-2.03)	12	1.26 (0.39-4.08)	
≥46	46	9454	40	1.54 (0.65-3.64)	9	2.58 (0.60-11.0)	19	1.26 (0.33-4.87)	12	1.58 (0.33-7.63)	0.546
P-trend				0.267		0.175		0.851		0.408	
per 2 sizes		33880	126	1.10 (0.78-1.56)	37	1.19 (0.67-2.10)	50	0.85 (0.52-1.39)	39	1.48 (0.74-2.98)	0.998

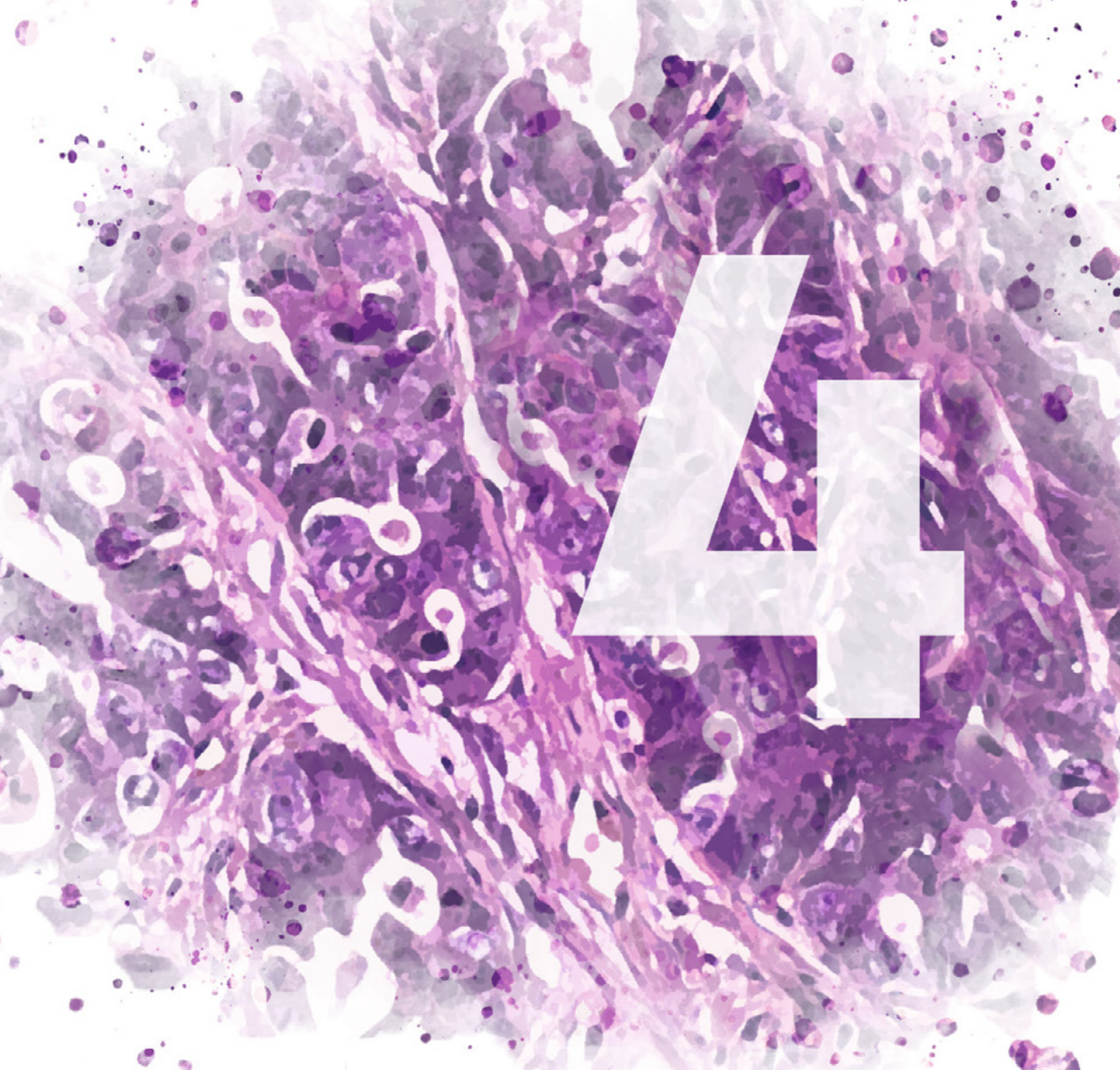
Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; BMI, body mass index; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), BMI (kg/m2; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes; no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Median clothing size per category based on the subcohort.



Supplementary Figure S1 | Immunohistochemical scoring of the six proteins incorporated in the Warburg-subtypes on tissue microarrays. The localization of the antigen-antibody complex is visualized in brown (DAB) and counterstaining is shown in blue (haematoxylin). **LDHA: A-I)** low, (any) weak cytoplasmic staining; **A-II)** moderate, 1-50% strong cytoplasmic staining; **A-III)** high, >50% strong cytoplasmic staining. **GLUT1: B-I)** low, <10% membranous staining; **B-II)** moderate, 11-50% membranous staining; **B-III)** high, >50% membranous staining. **MCT4: C-I)** low, <10% membranous staining; **C-II)** moderate, 11-50% membranous staining; **C-III)** high, >50% membranous staining. **PKM2: D-I)** low, (any) weak positive cytoplasmic staining; **D-II)** moderate, 1-50% strong cytoplasmic staining; **D-III)** high, >50% cytoplasmic staining. **P53: E-I)** low, <10% nuclear staining; **E-II)** moderate, 11-50% nuclear staining; **E-III)** high, >50% nuclear staining. **PTEN** (staining intensity of tumor cells was compared with that of stromal cells): **F-I)** low, completely negative cytoplasm; **F-II)** moderate, weak or moderate positive cytoplasm; **F-III)** high, strong cytoplasmic staining.



4

ENERGY BALANCE-RELATED FACTORS IN CHILDHOOD AND ADOLESCENCE AND RISK OF COLORECTAL CANCER EXPRESSING DIFFERENT LEVELS OF PROTEINS INVOLVED IN THE WARBURG-EFFECT

Josien C. A. Jenniskens, Kelly Offermans, Colinda C. J. M. Simons, Iryna Samarska, Gregorio E. Fazzi, Kim M. Smits, Leo J. Schouten, Matty P. Weijenberg, Heike I. Grabsch, Piet A. van den Brandt

International Journal of Cancer. 2022;150:1812-1824

ABSTRACT

Early-life (childhood to adolescence) energy balance-related factors (height, energy restriction, BMI) have been associated with adult colorectal cancer (CRC) risk. Warburg-effect activation via PI3K/Akt-signaling might explain this link. We investigated whether early-life energy balance-related factors were associated with risk of Warburg-subtypes in CRC. We used immunohistochemistry for six proteins involved in the Warburg-effect (LDHA, GLUT1, MCT4, PKM2, P53, and PTEN) on tissue microarrays of 2399 incident CRC cases from the prospective Netherlands Cohort Study (NLCS). Expression levels of all proteins were combined into a pathway-based sum score and categorized into three Warburg-subtypes (Warburg-low/-moderate/-high). Multivariable Cox-regression analyses were used to estimate associations of height, energy restriction proxies (exposure to Dutch Hunger Winter; Second World War [WWII]; Economic Depression) and adolescent BMI with Warburg-subtypes in CRC. Height was positively associated with colon cancer in men, regardless of Warburg-subtypes, and with Warburg-low colon and Warburg-moderate rectal cancer in women. Energy restriction during the Dutch Hunger Winter was inversely associated with colon cancer in men, regardless of Warburg-subtypes. In women, energy restriction during the Hunger Winter and WWII was inversely associated with Warburg-low colon cancer, whereas energy restriction during the Economic Depression was positively associated with Warburg-high colon cancer. Adolescent BMI was positively associated with Warburg-high colon cancer in men, and Warburg-moderate rectal cancer in women. In conclusion, the Warburg-effect seems to be involved in associations of adolescent BMI with colon cancer in men, and of energy restriction during the Economic Depression with colon cancer in women. Further research is needed to validate these results.

INTRODUCTION

Energy balance–related factors are known to influence risk of colorectal cancer. Factors related to early-life (childhood to adolescence) energy balance have been reported to have long-term effects on colorectal cancer (CRC) risk. Increased adult-attained height, as a proxy for fetal and early-life (nutritional) exposures, has been associated with an increased risk of CRC.[1,2] Early-life energy restriction seems to decrease risk of adult CRC,[3-5] though positive associations have been reported as well.[6,7] Childhood or adolescent body mass index (BMI) has been associated with an increased risk of CRC.[8] The mechanism(s) behind these long-term effects of early-life energy balance-related factors remain to be elucidated.

A proposed common effect of energy balance-related factors is aberrant signaling of insulin, insulin-like growth factor (IGF)-1, and IGF binding proteins (IGFBP), as well as adipokine (i.e. leptin and adiponectin) signaling.[9-11] These signaling molecules have previously been associated with CRC risk in adults,[12-15] but it is not clear whether aberrant insulin, IGF-1, or adipokine signaling early in life are associated with adult risk of CRC. Nevertheless, it has been proposed that early-life alterations of signaling molecules might persist until adulthood due to an accumulation effect over the years, by epigenetic changes, or by alterations in the gut microbiota.[16-20]

A common downstream effect of insulin, IGF-1, and leptin is the activation of the PI3K/Akt-signaling pathway, whereas adiponectin counteracts activation of this pathway.[21] Besides its well-known oncogenic effects, activation of the PI3K/Akt-signaling pathway has been shown to induce aerobic glycolysis by upregulation of several transporter proteins and (glycolytic) enzymes.[22,23] This metabolic phenotype is often referred to as the "Warburg-effect", named after its discovery by Otto Warburg and colleagues.[24] It has been suggested that the Warburg-effect is a cause rather than an effect of cancer,[25] which is supported by the addition of metabolic reprogramming as one of the emerging hallmarks of cancer.[26]

We have previously shown that adult BMI and clothing size are associated with colon cancer expressing high levels of proteins involved in the Warburg-effect.[27] However, studies investigating whether the Warburg-effect might be involved in the long-term effects of early-life energy balance-related factors on CRC risk are currently lacking. We therefore aimed to investigate whether early-life energy balance-related factors were associated with risk of CRC tumors expressing different levels of proteins involved in the Warburg-effect.

We aimed to capture the Warburg-effect by ensuring that the different steps of the pathway were represented by at least one protein (Table S1). These steps include upstream regulation of the Warburg-effect (PTEN, P53), glucose import (GLUT1), glycolysis (PKM2), conversion of pyruvate into lactate (LDHA), and lactate secretion (MCT4). The expression levels of these six proteins (PTEN, P53, GLUT1, PKM2, LDHA, and MCT4) were combined into a sum score, which was divided into three subgroups,

representing tumors with a low, moderate, or high likelihood of the presence of the Warburg-effect, hereafter referred to as the Warburg-subtypes (Warburg-low, Warburg-moderate, Warburg-high, respectively).

We hypothesized that associations between early-life energy balance-related factors (height, energy restriction, and adolescent BMI) and adult risk of CRC differ across Warburg-subtypes.

MATERIALS AND METHODS

DESIGN AND STUDY POPULATION

The Netherlands Cohort Study (NLCS) is a population-based prospective cohort study initiated in 1986, which included 120 852 subjects aged 55 to 69 years old.[28] At baseline, all participants filled in a mailed, self-administered questionnaire on cancer risk factors. A case-cohort design was used for data processing and analysis.[29] A subcohort (n = 5000) was randomly selected immediately after baseline, providing an estimate of accumulated person-years for the total cohort. Vital status information of subcohort members was checked by biennial active follow-up and by linkage with municipal population registries afterward. Only one male subcohort member was lost to follow-up. Incident cancer cases were drawn from the entire cohort via annual record linkage with the Netherlands Cancer Registry and PALGA, the nationwide Dutch Pathology Registry,[30] covering 20.3 years of follow-up (17 September 1986 until 1 January 2007). The completeness of cancer follow-up by the Netherlands Cancer Registry and PALGA was estimated to be over 96%.[31] A total of 4597 incident CRC and 4774 subcohort members were available after excluding prevalent cancer cases (except skin cancer) at baseline, as described previously.[27]

Formalin-fixed paraffin-embedded (FFPE) CRC tissue blocks from 3872 CRC cases were collected as part of the Rainbow-TMA project from 2012 to 2017.[32] CRC cases were selected based on available linkage to a PALGA-record (which provides access to pathology labs) and surgical specimen with pathology report, or coloscopic resection. Cases treated with neo-adjuvant therapy were excluded. Tissue blocks from 3021 CRC cases were successfully collected from 43 pathology laboratories throughout the Netherlands (78% retrieval rate).

For tissue microarray (TMA) construction, pathologists reviewed scanned Hematoxylin&Eosin (H&E)-stained sections and identified areas with the highest tumor density, from which three 0.6 mm diameter cores were sampled per case along with three normal tissue cores (TMA-Grandmaster, 3D-Histech, Hungary). In total, 78 TMAs were constructed comprising 2694 CRC cases.

IMMUNOHISTOCHEMISTRY

Five micrometers thick sections were cut from all 78 TMA blocks, H&E-stained according to standard protocol, and subjected to immunohistochemistry (IHC). An automated immunostainer (DAKO Autostainer Link 48, Glostrup, Denmark) was used for GLUT1, P53, and PTEN, whereas LDHA, PKM2, and MCT4 were stained manually. Detailed information on primary antibodies and staining protocols are shown in Table S2. All TMA sections were scanned using an Aperio scanner (Leica Microsystems, Milton Keynes, UK) at 40x magnification at the University of Leeds (UK) Scanning Facility or at the Department of Pathology, Aachen University Hospital (Germany).

H&E-stained TMA sections were reviewed, in combination with pan-cytokeratin stained sections if necessary, to confirm the presence of adenocarcinoma for each core. Requiring at least one core with adenocarcinoma per case, 2497 cases passed quality control. Scoring of IHC-stained TMAs was performed by three non-pathologists (G.E. Fazzi: senior histology technician; K. Offermans: PhD-student; J.C.A. Jenniskens: PhD-student; Table S3 shows the contribution of each assessor) after appropriate training, as described previously.[33] IHC scoring protocols for all markers are described in the Supporting Information Methods and shown in Figure S1. Kappa values on inter- and intra-observer scoring agreement are shown in Table S4. Complete IHC protein expression information for all six proteins was available for 2399 cases.

Establishment of Warburg-subtypes from core-level protein expression has previously been described in detail.[27] In short, (i) core-based scores from multiple assessors were combined into a combination score if the same score was assigned by at least two assessors; (ii) scoring discrepancies were resolved by consensus agreement of two non-pathologist assessors or by an experienced pathologist; (iii) final scores from all available cores per case were averaged and rounded to the nearest scoring category; (iv) the average scores per case were categorized as low, moderate, or high expression (Table S4 shows cut-offs per protein); (v) a pathway-based sum score (range: 0-12) was used to combine expression levels of all six proteins; (vi) Warburg-subtypes were established by taking tertiles from the sum score. Cases with sum scores 0 to 3 were classed as "Warburg-low" (n = 698, 29.1%), sum scores 4 to 5 as "Warburg-moderate" (n = 859, 35.8%), and sum scores 6 to 12 as "Warburg-high" (n = 842, 35.1%).

EARLY-LIFE ENERGY BALANCE-RELATED FACTORS

Proxy variables were used to assess exposure to energy restriction during childhood to adolescence, as previously described[16,34]: (i) place of residence during the Dutch Hunger Winter (1944-1945); (ii) place of residence in 1942, reflecting World War II (WWII; 1940-1944); and (iii) employment status of the father during the Dutch Economic Depression (1932-1940). Living in a city in the western part of the Netherlands during the Hunger Winter indicated severe energy restriction, with caloric intake of 400 to 800 kcal per day at the height of the famine.[35,36] Living in a Dutch city in 1942 (WWII) with more than 40 000 inhabitants was used as an indicator for energy restriction. Unemployment of the father during the Economic Depression was used as an indicator of lack of variation in the food pattern, though sufficient calories were available.

Participants of the NLCS were 12 to 28 years old during the Hunger Winter, 8 to 28 years old during WWII, and 0 to 23 years old during the Economic Depression.

Height was self-reported at baseline (cm), adolescent BMI was calculated by using self-reported weight at age 20 years and height at baseline (kg/m^2).

COX REGRESSION MODELS

After excluding participants with incomplete or inconsistent data on exposure variables or confounders, 3911 subcohort members and 1972 CRC cases were available for analyses. Cox proportional hazards models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for the association between CRC and early-life energy restriction measures (place of residence during the Hunger Winter; place of residence during WWII; employment status of the father during the Economic Depression), height (according to sex-specific quartiles, and per 5 cm increase), and adolescent BMI (according to sex-specific quartiles, and per 5 kg/m2 increase). All associations were investigated stratified on sex, tumor location, and Warburg-subtypes. Standard errors of the HRs were estimated using the Huber-White sandwich estimator to account for additional variance introduced by sampling the subcohort from the total cohort.[37] The proportional hazards assumption was tested using the scaled Schoenfeld residuals[38] and by introducing time-covariate interactions into the models.

All multivariable models were adjusted for age, family history of CRC (yes/no), alcohol intake (0; 0.1-4; 5-14; >15 g/day), energy intake at baseline (kcal/day), and non-occupational physical activity (minutes/day). Models on early-life energy restriction and adult-attained height were additionally adjusted for BMI at baseline (kg/m2). Models on adolescent BMI were additionally adjusted for height (cm). Potential additional confounders were weight change since adolescence (kg), smoking status (never/former/current), level of education (primary or lower vocational education; secondary or medium vocational education; higher vocational education or university), red meat consumption (g/day), and processed meat consumption (g/day). These potential confounders were included in multivariable models if they introduced a ≥10% change in HRs.

Heterogeneity in associations between risk factors and Warburg-subtypes was tested to evaluate differences across tumors expressing different levels of proteins involved in the Warburg-effect. This was done using an adapted version of the competing risks procedure in Stata developed for the case-cohort design, as described previously.[39,40] In sensitivity analyses, we used two instead of three Warburg-subtypes (Warburg-low: sum score 0-4; Warburg-high: sum score 5-12) to increase power. Since our analyses were hypothesis-driven and exposures reflect different aspects of (early-life) energy balance, we did not correct for multiple testing. All analyses were conducted in Stata Statistical Software: Release 16 (StataCorp., College Station, TX).

RESULTS

Descriptive results on early-life energy balance-related factors are shown in Table 1, stratified on sex, tumor location, and Warburg-subtypes. First, cases were taller compared to subcohort members. Cases with a Warburg-low tumor were generally tallest, except for male colon cancer cases, which showed similar mean height across Warburg-subtypes. Second, cases were generally less often exposed to early-life energy restriction compared to subcohort members. This was especially the case for male colon cancer cases with Warburg-high tumors, and for female colon cancer cases with Warburg-low tumors. Third, cases were more often overweight at age 20 years compared to subcohort members, except for male rectal cancer cases. This was especially the case for Warburg-high tumors, except for rectal cancer in women, where Warburg-moderate tumors showed the highest percentage of overweight cases.

Multivariable-adjusted Cox regression models on early-life energy balance-related factors are shown in Tables 2 to 4. Age-adjusted Cox regression models are shown in Tables S5 to S7. All models were stratified on sex, tumor location, and Warburg-subtypes, and included age as a time-varying covariate because of violation of the proportional hazards assumption.

ADULT-ATTAINED HEIGHT

In men, height was positively associated with risk of colon cancer [HR_{5cm} (95% CI): 1.12 (1.05-1.19); P-$trend_{quartiles}$ = 0.004], but not rectal cancer (Table 2). After stratification, similar associations were found across all Warburg-subtypes. In women, height was positively associated with risk of both colon [HR_{5cm} (95% CI): 1.09 (1.01-1.17); P-$trend_{quartiles}$ = 0.056] and rectal cancer [HR_{5cm} (95% CI): 1.16 (0.99-1.37); P-$trend_{quartiles}$ = 0.047] (Table 2). After stratification on Warburg-subtypes, height was positively associated with Warburg-low colon cancer [HR_{5cm} (95% CI): 1.16 (1.02-1.32); P-$trend_{quartiles}$ = 0.033]. For rectal cancer, positive associations were found for the Warburg-low, although not statistically significant [HR_{5cm} (95% CI): 1.21 (0.92-1.59); P-$trend_{quartiles}$ = 0.224], and the Warburg-moderate subtype [HR_{5cm} (95% CI): 1.22 (0.95-1.55); P-$trend_{quartiles}$ = 0.037].

PROXIES FOR EARLY-LIFE ENERGY RESTRICTION

Living in the western part of the Netherlands during the Hunger Winter (1944-1945) was associated with a decreased risk of both colon and rectal cancer in men, compared to participants living in a nonwestern part [colon: $HR_{western\ rural}$ (95% CI): 0.69 (0.51-0.92); $HR_{western\ city}$: 0.63 (0.50-0.81); rectum: $HR_{western\ rural}$: 0.72 (0.46-1.12); $HR_{western\ city}$: 0.60 (0.40-0.89)] (Table 3). After stratification, similar associations were found for all Warburg-subtypes. Place of residence during WWII or employment status of the father during the Economic Depression were not associated with adult colon or rectal cancer risk in men (Table 3).

Table 1 | Characteristics [mean (SD) or %] of subcohort members and colorectal cancer cases; NLCS, 1986–2006.

	Subcohort	Colon				Rectum			
		Total	Warburg low	Warburg moderate	Warburg high	Total	Warburg low	Warburg moderate	Warburg high
Men									
N[a]	1971	772	215	280	277	227	76	76	75
Age at baseline (years)	61.3 (4.2)	61.6 (4.2)	61.8 (4.2)	61.3 (4.1)	61.8 (4.2)	60.7 (3.9)	60.8 (3.6)	61.1 (4.3)	60.2 (3.6)
Height (cm)	176.6 (6.7)	177.3 (6.8)	177.1 (6.6)	177.5 (6.8)	177.3 (6.9)	176.9 (6.6)	177.7 (7.0)	176.4 (5.5)	176.7 (7.0)
Hunger Winter (Western city %)	21.4	16.3	16.1	18.3	14.4	15.9	15.8	15.8	16.0
Wolrd War II (city %)	49.7	46.5	49.7	47.9	42.4	45.6	40.3	46.6	50.0
Economic depression (unemployed %)	10.8	10.6	13.3	9.6	9.7	10.5	10.8	9.3	11.4
Overweight at age 20 years (%)	7.6	8.5	7.8	8.5	9.1	6.0	4.4	5.9	7.9
Weight change since age 20 (kg)	10.3 (9.3)	11.1 (9.5)	11.5 (9.3)	11.4 (9.7)	10.6 (9.6)	9.7 (9.2)	10.4 (8.9)	9.0 (8.8)	9.7 (10.0)
Women									
N[a]	1940	655	170	216	269	127	37	51	39
Age at baseline (years)	61.4 (4.3)	61.9 (4.1)	62.0 (4.1)	62.0 (4.1)	61.8 (4.1)	61.4 (4.2)	60.5 (4.1)	62.1 (4.2)	61.5 (4.3)
Height (cm)	165.3 (6.1)	166.0 (6.2)	166.5 (5.8)	165.8 (6.3)	165.8 (6.3)	166.5 (6.4)	166.9 (6.0)	166.8 (6.2)	165.8 (7.1)
Hunger Winter (Western city %)	27.8	24.6	22.3	23.7	26.8	25.0	22.2	31.4	18.9
Wolrd War II (city %)	52.3	49.1	42.5	50.6	52.4	51.1	56.7	46.3	52.2
Economic depression (unemployed %)	11.3	11.8	9.3	8.8	15.7	9.8	5.7	10.0	13.2
Overweight at age 20 years (%)	7.3	7.9	5.1	6.2	11.3	7.8	5.9	13.0	2.9
Weight change since age 20 (kg)	9.9 (9.9)	10.1 (10.4)	10.8 (10.7)	9.6 (10.3)	10.1 (10.2)	9.7 (10.0)	10.4 (9.7)	9.5 (10.9)	9.4 (9.4)

Abbreviations: SD, standard deviation; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; WWII, World War II

[a]Total number based on the most complete variable (height). Numbers of other variables might not add up to the same total because of missing values.

In women, none of the proxy variables for early-life energy restriction showed statistically significant associations with overall colon or rectal cancer risk (Table 3). However, after stratification on Warburg-subtypes, living in the western part of the Netherlands during the Hunger Winter was associated with a lower risk of Warburg-low colon cancer [$HR_{western\ rural}$ (95% CI): 0.57 (0.33-0.99); $HR_{western\ city}$: 0.68 (0.46-1.00)]. Similarly, living in an urban area during WWII was inversely associated with risk of Warburg-low colon cancer [$HR_{urban\ area}$ (95% CI): 0.66 (0.46-0.94)]. In contrast, unemployment of the father during the Economic Depression was associated with a higher risk of Warburg-high colon cancer [HR (95% CI): 1.51 (1.03-2.20)]. For rectal cancer, no clear associations were found after stratification on Warburg-subtypes.

ADOLESCENT BMI

A high adolescent BMI was non-significantly associated with an increased risk of overall colon cancer in men [highest compared to lowest quartile: HR (95% CI): 1.26 (0.96-1.66); $P\text{-trend}_{quartiles}$ = 0.069] (Table 4). After stratification on Warburg-subtypes, this association was stronger for the Warburg-high subtype [highest compared to lowest quartile: HR (95% CI): 1.62 (1.08-2.43); $P\text{-trend}_{quartiles}$ = 0.022], whereas no associations were found for Warburg-low and Warburg-moderate subtypes. Adolescent BMI was not associated with rectal cancer risk in men.

In women, adolescent BMI was not associated with either colon or rectal cancer risk (Table 4). After stratification on Warburg-subtypes, a positive association was found with risk of Warburg-moderate rectal cancer [$HR_{5kg/m2}$ (95% CI): 1.64 (1.12-2.39); $P\text{-trend}_{quartiles}$ = 0.092].

TESTS FOR HETEROGENEITY

Heterogeneity tests did not show any statistically significant differences between Warburg-subtypes for any of the associations.

SENSITIVITY ANALYSES

The use of two instead of three Warburg-subtypes generally led to similar conclusions. Associations that were found for the Warburg-moderate subtype when using three Warburg-subtypes resulted in similar associations for Warburg-low and Warburg-high subtypes when two Warburg-subtypes were used (*data not shown*).

Table 2 | Multivariable-adjusted HRs[a] and 95% CIs for associations between adult-attained height and CRC, stratified on sex, tumor location, and Warburg-subtypes; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Quartiles of height (cm)											
Men – colon											
<173	170	8935	196	1.00 (ref.)	55	1.00 (ref.)	69	1.00 (ref.)	72	1.00 (ref.)	
173-176	175	7680	191	1.19 (0.93-1.52)	52	1.14 (0.75-1.73)	71	1.24 (0.86-1.79)	68	1.18 (0.82-1.71)	
177-181	179	7097	178	1.20 (0.93-1.55)	53	1.29 (0.85-1.95)	66	1.23 (0.85-1.79)	59	1.11 (0.76-1.62)	
>181	185	7310	207	1.46 (1.14-1.87)	55	1.39 (0.92-2.09)	74	1.43 (1.00-2.05)	78	1.54 (1.07-2.22)	0.978
P-trend				0.004		0.092		0.064		0.035	
Per 5 cm		31022	772	1.12 (1.05-1.19)	215	1.09 (0.98-1.21)	280	1.13 (1.03-1.24)	277	1.13 (1.02-1.24)	0.736
Men – rectum											
<173	170	8935	61	1.00 (ref.)	17	1.00 (ref.)	23	1.00 (ref.)	21	1.00 (ref.)	
173-176	175	7680	51	0.93 (0.62-1.40)	16	1.06 (0.53-2.13)	18	0.87 (0.45-1.68)	17	0.90 (0.46-1.75)	
177-181	179	7097	66	1.35 (0.92-1.98)	24	1.69 (0.88-3.22)	21	1.17 (0.63-2.19)	21	1.26 (0.65-2.42)	
>181	185	7310	49	0.97 (0.64-1.48)	19	1.30 (0.65-2.59)	14	0.76 (0.38-1.51)	16	0.94 (0.46-1.90)	0.927
P-trend				0.609		0.232		0.673		0.881	
Per 5 cm		31022	227	1.04 (0.93-1.15)	76	1.07 (0.90-1.28)	76	1.01 (0.87-1.18)	75	1.03 (0.85-1.25)	0.749
Women – colon											
<162	158	8764	140	1.00 (ref.)	30	1.00 (ref.)	51	1.00 (ref.)	59	1.00 (ref.)	
162-165	164	9216	185	1.26 (0.97-1.63)	54	1.75 (1.09-2.81)	57	1.04 (0.70-1.55)	74	1.20 (0.83-1.74)	
166-169	168	7771	152	1.21 (0.92-1.59)	34	1.30 (0.78-2.19)	58	1.22 (0.81-1.83)	60	1.15 (0.77-1.71)	
>169	172	8477	178	1.34 (1.03-1.75)	52	1.92 (1.19-3.11)	50	0.97 (0.64-1.47)	76	1.38 (0.94-2.01)	0.315
P-trend				0.056		0.033		0.918		0.137	
Per 5 cm		34228	655	1.09 (1.01-1.17)	170	1.16 (1.02-1.32)	216	1.05 (0.93-1.18)	269	1.08 (0.96-1.20)	0.649

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
<162	158	8764	28	1.00 (ref.)	8	1.00 (ref.)	9	1.00 (ref.)	11	1.00 (ref.)	
162-165	164	9216	23	0.78 (0.45-1.37)	7	0.86 (0.31-2.35)	7	0.77 (0.28-2.08)	9	0.73 (0.30-1.78)	
166-169	168	7771	38	1.52 (0.91-2.55)	11	1.60 (0.64-4.03)	21	2.70 (1.20-6.10)	6	0.60 (0.22-1.64)	
>169	172	8477	38	1.42 (0.84-2.41)	11	1.50 (0.60-3.78)	14	1.72 (0.70-4.23)	13	1.16 (0.49-2.74)	0.935
P-trend				0.047		0.224		0.037		0.778	
Per 5 cm		34228	127	1.16 (0.99-1.37)	37	1.21 (0.92-1.59)	51	1.22 (0.95-1.55)	39	1.08 (0.79-1.47)	0.880

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.
[a]Hazard Ratios were adjusted for age (years; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), BMI at baseline (kg/m^2; continuous), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.
[b]Median height per quartile based on the subcohort.

Table 3 | Multivariable-adjusted HRs[a] and 95% CIs for associations between early-life energy restriction and CRC, stratified on sex, tumor location, and Warburg-subtypes; NLCS, 1986-2006.

	Person years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Place of residence during the Dutch Hunger Winter (1944-45)									
Men – colon	30174	754		211		273		270	
Non-west	15188	446	1.00 (ref.)	124	1.00 (ref.)	169	1.00 (ref.)	153	1.00 (ref.)
Western rural	4129	78	0.69 (0.51-0.92)	22	0.68 (0.42-1.11)	24	0.56 (0.35-0.89)	32	0.83 (0.55-1.25)
Western city	6524	123	0.63 (0.50-0.81)	34	0.62 (0.41-0.94)	50	0.69 (0.49-0.97)	39	0.58 (0.40-0.85)
P-het									0.606
Men – rectum	30174	227		76		76		75	
Non-west	15188	141	1.00 (ref.)	44	1.00 (ref.)	49	1.00 (ref.)	48	1.00 (ref.)
Western rural	4129	27	0.72 (0.46-1.12)	11	0.94 (0.47-1.88)	10	0.75 (0.37-1.51)	6	0.48 (0.20-1.13)
Western city	6524	36	0.60 (0.40-0.89)	12	0.64 (0.33-1.23)	12	0.57 (0.30-1.09)	12	0.60 (0.31-1.15)
P-het									0.983
Women – colon	33722	646		166		215		265	
Non-west	18083	373	1.00 (ref.)	108	1.00 (ref.)	125	1.00 (ref.)	140	1.00 (ref.)
Western rural	4851	81	0.81 (0.61-1.08)	16	0.57 (0.33-0.99)	27	0.79 (0.50-1.24)	38	1.02 (0.69-1.50)
Western city	9234	159	0.83 (0.67-1.03)	37	0.68 (0.46-1.00)	51	0.81 (0.57-1.14)	71	0.97 (0.71-1.32)
P-het									0.238
Women – rectum	33722	124		36		51		37	
Non-west	18083	74	1.00 (ref.)	23	1.00 (ref.)	25	1.00 (ref.)	26	1.00 (ref.)
Western rural	4851	11	0.59 (0.31-1.13)	4	0.71 (0.23-2.20)	5	0.73 (0.28-1.93)	2	0.32 (0.08-1.39)
Western city	9234	31	0.86 (0.55-1.34)	8	0.71 (0.31-1.62)	16	1.25 (0.66-2.36)	7	0.59 (0.24-1.44)
P-het									0.803
Place of residence during World War II (1942)									
Men – colon	23793	572		167		209		196	
Rural area	11327	287	1.00 (ref.)	78	1.00 (ref.)	104	1.00 (ref.)	105	1.00 (ref.)
Urban area	11713	266	0.90 (0.73-1.11)	83	1.00 (0.71-1.41)	100	0.96 (0.71-1.30)	83	0.77 (0.56-1.05)
P-het									0.614
Men – rectum	23973	182		62		58		62	
Rural area	11327	95	1.00 (ref.)	37	1.00 (ref.)	29	1.00 (ref.)	29	1.00 (ref.)
Urban area	11713	83	0.88 (0.63-1.22)	25	0.65 (0.37-1.11)	27	0.95 (0.54-1.66)	31	1.13 (0.66-1.96)
P-het									0.381
Women – colon	26164	505		139		158		208	
Rural area	11882	243	1.00 (ref.)	78	1.00 (ref.)	74	1.00 (ref.)	91	1.00 (ref.)
Urban area	13562	248	0.90 (0.73-1.11)	59	0.66 (0.46-0.94)	80	0.97 (0.68-1.37)	109	1.05 (0.78-1.43)
P-het									0.111
Women – rectum	26164	94		30		41		23	
Rural area	11882	42	1.00 (ref.)	13	1.00 (ref.)	19	1.00 (ref.)	10	1.00 (ref.)
Urban area	13562	48	1.02 (0.66-1.58)	17	1.20 (0.57-2.54)	19	0.87 (0.45-1.65)	12	1.17 (0.47-2.93)
P-het									0.592

(continued)	Person years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Employment status of the father during the Dutch Economic Depression (1932-40)									
Men – colon	29841	743		203		271		269	
Employed	26697	664	1.00 (ref.)	176	1.00 (ref.)	245	1.00 (ref.)	243	1.00 (ref.)
Unemployed	3145	79	0.95 (0.71-1.26)	27	1.23 (0.79-1.91)	26	0.85 (0.55-1.33)	26	0.84 (0.54-1.30)
P-het									0.214
Men – rectum	29841	219		74		75		70	
Employed	26697	196	1.00 (ref.)	66	1.00 (ref.)	68	1.00 (ref.)	62	1.00 (ref.)
Unemployed	3145	23	0.99 (0.63-1.57)	8	1.01 (0.47-2.15)	7	0.86 (0.40-1.89)	8	1.09 (0.51-2.34)
P-het									0.756
Women – colon	32597	627		161		204		262	
Employed	29046	553	1.00 (ref.)	146	1.00 (ref.)	186	1.00 (ref.)	221	1.00 (ref.)
Unemployed	3552	74	1.09 (0.81-1.47)	15	0.85 (0.48-1.49)	18	0.79 (0.47-1.32)	41	1.51 (1.03-2.20)
P-het									0.872
Women – rectum	32597	123		35		50		38	
Employed	29046	111	1.00 (ref.)	33	1.00 (ref.)	45	1.00 (ref.)	33	1.00 (ref.)
Unemployed	3552	12	0.85 (0.46-1.57)	2	0.41 (0.09-1.80)	5	0.88 (0.35-2.21)	5	1.29 (0.51-3.27)
P-het									0.936

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity

[a]Hazard Ratios were adjusted for age (years; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), BMI at baseline (kg/m^2; continuous), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

Table 4 | Multivariable-adjusted HRs[a] and 95% CIs for associations between adolescent BMI (age 20 years) and CRC, stratified on sex, tumor location, and Warburg-subtypes; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Quartiles of BMI at age 20 years (kg/m²)											
Men – colon											
<20.2	19.2	6103	159	1.00 (ref.)	48	1.00 (ref.)	61	1.00 (ref.)	50	1.00 (ref.)	
20.2-21.6	21.0	6458	149	0.91 (0.69-1.20)	48	0.97 (0.63-1.50)	47	0.74 (0.49-1.12)	54	1.05 (0.69-1.59)	
21.7-23.3	22.4	6308	157	1.01 (0.77-1.34)	41	0.88 (0.55-1.39)	62	1.06 (0.71-1.57)	54	1.09 (0.72-1.67)	
>23.3	24.3	6012	181	1.26 (0.96-1.66)	42	0.95 (0.60-1.51)	65	1.21 (0.82-1.79)	74	1.62 (1.08-2.43)	0.301
P-trend				0.069		0.734		0.169		0.022	
Per 5 kg/m²		24881	646	1.06 (0.87-1.29)	179	0.93 (0.67-1.30)	235	1.03 (0.77-1.40)	232	1.19 (0.89-1.59)	0.400
Men – rectum											
<20.2	19.2	6103	40	1.00 (ref.)	15	1.00 (ref.)	14	1.00 (ref.)	11	1.00 (ref.)	
20.2-21.6	21.0	6458	63	1.50 (0.98-2.30)	21	1.34 (0.67-2.68)	20	1.39 (0.69-2.78)	22	1.82 (0.85-3.90)	
21.7-23.3	22.4	6308	46	1.14 (0.71-1.82)	15	1.06 (0.49-2.26)	19	1.33 (0.62-2.87)	12	1.01 (0.43-2.36)	
>23.3	24.3	6012	50	1.29 (0.81-2.05)	17	1.32 (0.63-2.77)	15	1.07 (0.48-2.38)	18	1.51 (0.69-3.31)	0.890
P-trend				0.573		0.627		0.921		0.675	
Per 5 kg/m²		24881	199	1.14 (0.87-1.51)	68	1.26 (0.83-1.89)	68	1.05 (0.65-1.69)	63	1.12 (0.69-1.82)	0.877
Women – colon											
<19.6	18.4	7795	148	1.00 (ref.)	39	1.00 (ref.)	52	1.00 (ref.)	57	1.00 (ref.)	
19.6-21.2	20.5	7731	152	1.05 (0.80-1.38)	41	1.09 (0.68-1.75)	45	0.87 (0.56-1.33)	66	1.21 (0.82-1.78)	
21.3-23.0	22.0	7964	151	1.03 (0.79-1.36)	43	1.16 (0.73-1.84)	55	1.04 (0.69-1.57)	53	0.94 (0.63-1.41)	
>23.0	24.2	7683	141	1.06 (0.80-1.40)	34	1.01 (0.62-1.66)	43	0.87 (0.56-1.35)	64	1.26 (0.85-1.87)	0.577
P-trend				0.752		0.880		0.744		0.490	
Per 5 kg/m²		31173	592	1.07 (0.91-1.26)	157	1.02 (0.79-1.33)	195	0.95 (0.75-1.21)	240	1.21 (0.95-1.53)	0.319

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high		P-het
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Women – rectum											
<19.6	18.4	7795	22	1.00 (ref.)	8	1.00 (ref.)	7	1.00 (ref.)	7	1.00 (ref.)	
19.6-21.2	20.5	7731	37	1.77 (1.01-3.12)	11	1.45 (0.55-3.83)	13	2.05 (0.78-5.37)	13	1.89 (0.73-4.86)	
21.3-23.0	22.0	7964	28	1.29 (0.72-2.33)	7	0.83 (0.29-2.36)	12	1.85 (0.71-4.87)	9	1.34 (0.48-3.72)	
>23.0	24.2	7683	28	1.39 (0.76-2.54)	8	1.01 (0.37-2.78)	14	2.40 (0.91-6.30)	6	0.93 (0.30-2.91)	0.970
P-trend				0.529		0.720		0.092		0.724	
Per 5 kg/m^2		31173	115	1.15 (0.86-1.53)	34	0.81 (0.46-1.42)	46	1.64 (1.12-2.39)	35	0.97 (0.60-1.58)	0.095

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), height (cm; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Median adolescent BMI per quartile based on the subcohort.

DISCUSSION

In this large prospective cohort study, we investigated the associations of early-life energy balance-related factors with adult risk of Warburg-subtypes in CRC using a molecular pathological epidemiology approach. Height was positively associated with an increased risk of colon cancer in men, irrespective of Warburg-subtypes. In women, height was positively associated with risk of Warburg-low colon cancer, and Warburg-low and Warburg-moderate rectal cancer. Place of residence during the Dutch Hunger Winter, as a proxy for energy restriction, was inversely associated with colon cancer risk in men, regardless of Warburg-subtypes. None of the other energy restriction proxies were associated with colon or rectal cancer risk in men. In women, energy restriction during the Hunger Winter or during WWII showed an inverse association with Warburg-low colon cancer. In contrast, unemployment of the father during the Dutch Economic Depression was positively associated with risk of Warburg-high colon cancer in women. Adolescent BMI was associated with an increased risk of Warburg-high colon cancer in men, and Warburg-moderate rectal cancer in women. None of the heterogeneity tests showed statistically significant differences between Warburg-subtypes.

Studies investigating potential etiological differences across CRC cases expressing different levels of the Warburg-effect in the tumor are currently lacking. However, proposed precursors (i.e. insulin, IGF-1, leptin, and adiponectin) of the Warburg-effect have been investigated in relation to CRC risk,[12-15] reporting positive associations for leptin, insulin, and IGF-1, and inverse associations for adiponectin. However, these associations were studied in adults. It has been proposed that early-life alterations in these signaling molecules, for example by exposure to energy balance-related factors, can have long-term detrimental effects.[16-18,41] We were the first, to the best of our knowledge, to investigate whether early-life energy-balance-related factors are associated with CRC subtypes expressing different levels of proteins involved in the Warburg-effect. Even though heterogeneity tests were not statistically significant, we will discuss observed differences between Warburg-subtypes.

HEIGHT

The current results do not suggest a role of the Warburg-effect in the link between height and CRC. In men, similar associations with height were found for all Warburg-subtypes of colon cancer, suggesting that a pathway other than the Warburg-effect dominates this risk-enhancement. In women, height was associated with Warburg-low colon cancer and Warburg-moderate rectal cancer, suggesting that the Warburg-effect does not play a role, at least not to a great extent, in women either.

Height is often not considered to be the cause of cancer itself, but rather a marker for various dietary and lifestyle exposures early in life, from conception until puberty, that affect growth.[42] The various early-life exposures associated with height might affect adult CRC risk via different pathways. Apart from the pathway under investigation in the current study, involving IGF-1, insulin, leptin, and/or adiponectin signaling, it has been suggested that increased cancer incidence in tall people might be related to the higher

number of cells and cell divisions.[9,43] In particular, it has been suggested that height is associated with intestinal length,[44] potentially resulting in a higher risk of cellular alterations leading to malignancies.[1] Further research is needed to decipher which mechanisms might be involved in the link between height and CRC risk.

ENERGY RESTRICTION

In men, exposure to the Dutch Hunger Winter, which was used as an indicator of severe energy restriction early in life, was associated with a decreased risk of colon and rectal cancer, irrespective of Warburg-subtypes. In women, exposure to the Dutch Hunger Winter as well as during WWII was inversely associated with Warburg-low colon cancer. This might suggest that energy restriction hinders oncogenic mechanisms other than the Warburg-effect. In contrast, exposure to unemployment of the father during the Dutch Economic Depression was associated with an increased risk of Warburg-high colon cancer in women, suggesting a role of the Warburg-effect in this risk-enhancement.

Previous studies on early-life energy restriction in relation to adult CRC risk show inconsistent results, reporting both inverse and positive associations.[3-7,45] The use of different proxies for early-life energy restriction entails variation in exposure, for example, timing, duration, or severity. This variation in exposure potentially explains the differential associations we observed. First, participants were older during the Hunger Winter and WWII (12-28 and 8-28 years, respectively) compared to the Economic Depression (0-23 years). It has previously been proposed that particularly exposure early in childhood increases risk of CRC.[7] Second, place of residence during the Hunger Winter and WWII are indicators of (severe) energy restriction, whereas unemployment of the father during the Economic Depression indicates a lack of dietary variety.[34] It has been proposed that lack of vitamins or important (micro)nutrients are associated with an increased risk of CRC.[6,7]

The potential involvement of the Warburg-effect in the positive association between energy restriction during the Economic Depression and colon cancer in women might be supported by a study of Elias et al,[18] which showed a long-term effect of early-life energy restriction on the IGF-axis in adult women, with increased plasma levels of IGF-1 and IGFBP-3. As mentioned, IGF-1 is considered a potential instigator of the Warburg-effect in tumor cells.[21-23] Other studies investigating the relationship between energy restriction and the Warburg-effect were performed in adults and entailed rather short-term effects.[10,46]

ADOLESCENT BMI

Our results suggest a role of the Warburg-effect in men for the association between high adolescent BMI and increased adult risk of CRC. This effect was only shown for the highest quartile, which included cases with a BMI >23.3 kg/m^2, and was thus the only one encompassing overweight cases according to standards of the World Health Organization, defining a BMI ≥25 kg/m^2 as overweight. The variation of adolescent BMI across cases from the NLCS was thus limited, which might explain the lack of significant associations for example for colon cancer in women, which showed a similar HR per 5

kg/m^2 as men. It would thus be interesting to replicate the current study in a population with more variation in adolescent BMI across individuals, as increased contrasts might give further insights.

In line with the current results, we have previously shown that adult BMI, reported at baseline (age range: 55-69), was associated with Warburg-moderate and Warburg-high colon cancer in men.[27] This suggests that the same pathway, involving the Warburg-effect, is involved in both the long-term and shorter-term effect of a high BMI on colon cancer risk in men.

The positive association we found for adolescent BMI and risk of Warburg-moderate rectal cancer in women suggests that the Warburg-effect is not involved, at least not to a great extent, in this relation. In our previous study on adult BMI,[27] a positive association was found for Warburg-moderate rectal cancer as well, but this association did not reach statistical significance. As proposed previously, a potential mechanism other than the Warburg-effect in the relationship between adiposity and CRC risk in women might be related to sex hormones,[27] though this was investigated in adult women instead of adolescents.[47] Further research is needed to investigate potential mechanisms involved in the link between (adolescent) BMI and rectal cancer in women.

STRENGTHS AND LIMITATIONS

A major strength of the current study is the large prospective cohort design with long follow-up (20.3 years). However, heterogeneity in associations based on sex and tumor location limited the number of cases in final statistical analyses. Furthermore, this potentially resulted in chance findings due to multiple testing. Our results should thus be interpreted with caution since validation of the current findings is needed.

Another important strength of the current study is the unique opportunity to investigate exposure to early-life energy restriction in relation to adult CRC risk in a large prospective cohort. Nevertheless, the use of proxy measures for energy restriction during the Dutch Hunger Winter, WWII and the Dutch Economic Depression might have resulted in some exposure misclassification, since individual data on dietary exposures during those times were not available. However, a previous validation study among female subcohort members concluded that the proxy measure for energy restriction in the Hunger Winter was adequate.[48] In addition, since the timing of exposure to energy restriction might influence its association with cancer risk,[7,45] the wide age-span (0-28 years) of our early-life energy restriction proxies might have affected the current results. An additional limitation for the current study is the retrospectively obtained information on early-life energy-balance-related factors, which could introduce information bias, especially for adolescent weight.

A third important strength of the current study is the availability of tumor material for a large number of incident CRC cases. Still, assessment of protein expression by IHC on TMAs from FFPE tissue might potentially entail some concerns. First, the age of FFPE tissue blocks from CRC resections might be a concern for the reliability of IHC results,

though a study by Grillo et al[49] showed that most antigens are well preserved and available for analysis after several decades. Furthermore, we used internal controls when possible as a quality control measure to ensure adequate staining of the tissue. Second, the usage of TMAs instead of full sections might have resulted in misclassification due to the subsampling of the tumor. However, we tried to minimize this problem by sampling the three cores from different regions to capture potential tumor heterogeneity. Third, IHC-scoring by non-pathologists might be conceived as a potential problem. However, we have previously shown that non-pathologists can generate reproducible IHC-scoring results, similar to those of a pathologist, provided they were sufficiently trained by an experienced pathologist.[33] Fourth, the semi-quantitative assessment of protein expression entails a margin of error, potentially leading to outcome misclassification. To minimize this issue, we used multiple observers to assess protein expression levels based on IHC-stained tissue.

To enable replication of the current results, we used a relatively simple sum score of six proteins involved in the Warburg-effect. We acknowledge that this also entails some disadvantages, as it probably does not reflect all factors involved in the Warburg-effect. However, capturing the complete pathway is nearly impossible considering time and budgetary constraints. We aimed to capture the Warburg-effect by incorporating proteins from different levels of the pathway in this sum score (i.e. from upstream regulators, glucose import, glycolysis, and lactate secretion), aiming to provide a comprehensive view of the Warburg-effect. Additional important considerations for the choice of the specific proteins included: (i) the availability and quality of scientific literature on the protein, preferably in relation to the Warburg-effect in CRC; and (ii) whether the protein was already an established marker in the routine diagnostic histopathology lab to ensure reliable IHC staining results. We have previously reported differences in prognosis, independent of well-known prognostic factors, for the current Warburg-subtypes.[50]

CONCLUSION

In this large prospective cohort study, we found different associations for early-life energy balance-related factors and CRC risk between Warburg-subtypes defined by expression levels of six key proteins related to the Warburg-effect. The Warburg-effect seems to be involved in the positive association between adolescent BMI and colon cancer risk in men. The Warburg-effect is probably involved neither in the association between height and colon or rectal cancer risk in men or women, nor in the association between adolescent BMI and rectal cancer risk in women. Results on early-life energy restriction proxies in relation to risk of Warburg-subtypes in CRC did not show clear patterns. Further research is needed to validate current results and investigate potential additional mechanisms.

REFERENCES

1. Abar L, Vieira AR, Aune D, et al. Height and body fatness and colorectal cancer risk: an update of the WCRF-AICR systematic review of published prospective studies. Eur J Nutr. 2018;57(5):1701-1720.
2. Gunnell D, Okasha M, Davey Smith G, Oliver S, Sandhu J, Holly J. Height, leg length, and cancer risk: a systematic review. Epidemiol Rev. 2001;23(2):313-342.
3. Hughes LA, van den Brandt PA, Goldbohm RA, et al. Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from The Netherlands cohort study. Int J Epidemiol. 2010;39(5):1333-1344.
4. 4. Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age–period–cohort analysis. Eur J Cancer Prev. 2002;11(5):489-495.
5. Svensson E, Møller B, Tretli S, et al. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. Cancer Causes Control. 2005;16(3):215-223.
6. Keinan-Boker L, Vin-Raviv N, Liphshitz I, Linn S, Barchana M. Cancer incidence in Israeli jewish survivors of world war II. J Natl Cancer Inst. 2009;101(21):1489-1500.
7. Xie SH, Lagergren J. A possible link between famine exposure in early life and future risk of gastrointestinal cancers: implications from age-period-cohort analysis. Int J Cancer. 2017;140(3):636-645.
8. Garcia H, Song M. Early-life obesity and adulthood colorectal cancer risk: a meta-analysis. Rev Panam Salud Publica. 2019;43:e3.
9. Stefan N, Häring H-U, Hu FB, Schulze MB. Divergent associations of height with cardiometabolic disease and cancer: epidemiology, pathophysiology, and global implications. Lancet Diab Endocrinol. 2016;4(5):457-467.
10. Castejón M, Plaza A, Martinez-Romero J, Fernandez-Marcos PJ, de Cabo R, Diaz-Ruiz A. Energy restriction and colorectal cancer: a call for additional research. Nutrients. 2020;12(1):114.
11. Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. Nat Rev Endocrinol. 2014; 10(8):455-465.
12. Yin S, Bai H, Jing D. Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients: a systemic review and meta-analysis. Diagn Pathol. 2014;9(1):1-6.
13. Renehan AG, Zwahlen M, Minder C, O'Dwyer TS, Shalet SM, Egger M. Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. The Lancet. 2004;363(9418):1346-1353.
14. Joshi RK, Kim WJ, Lee S-A. Association between obesity-related adipokines and colorectal cancer: a case-control study and meta-analysis. World J Gastroenterol. 2014;20(24):7941-7949.
15. Yoon YS, Kwon AR, Lee YK, Oh SW. Circulating adipokines and risk of obesity related cancers: a systematic review and meta-analysis. Obes Res Clin Pract. 2019;13(4):329-339.
16. Simons CCJM, van den Brandt PA, Stehouwer CDA, van Engeland M, Weijenberg MP. Body size, physical activity, early-Life energy restriction, and associations with methylated insulin-like growth factor-binding protein genes in colorectal cancer. Cancer Epidemiol Biomarkers Prev. 2014;23(9):1852-1862. https://doi.org/10.1158/1055-9965.epi13-1285.
17. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated

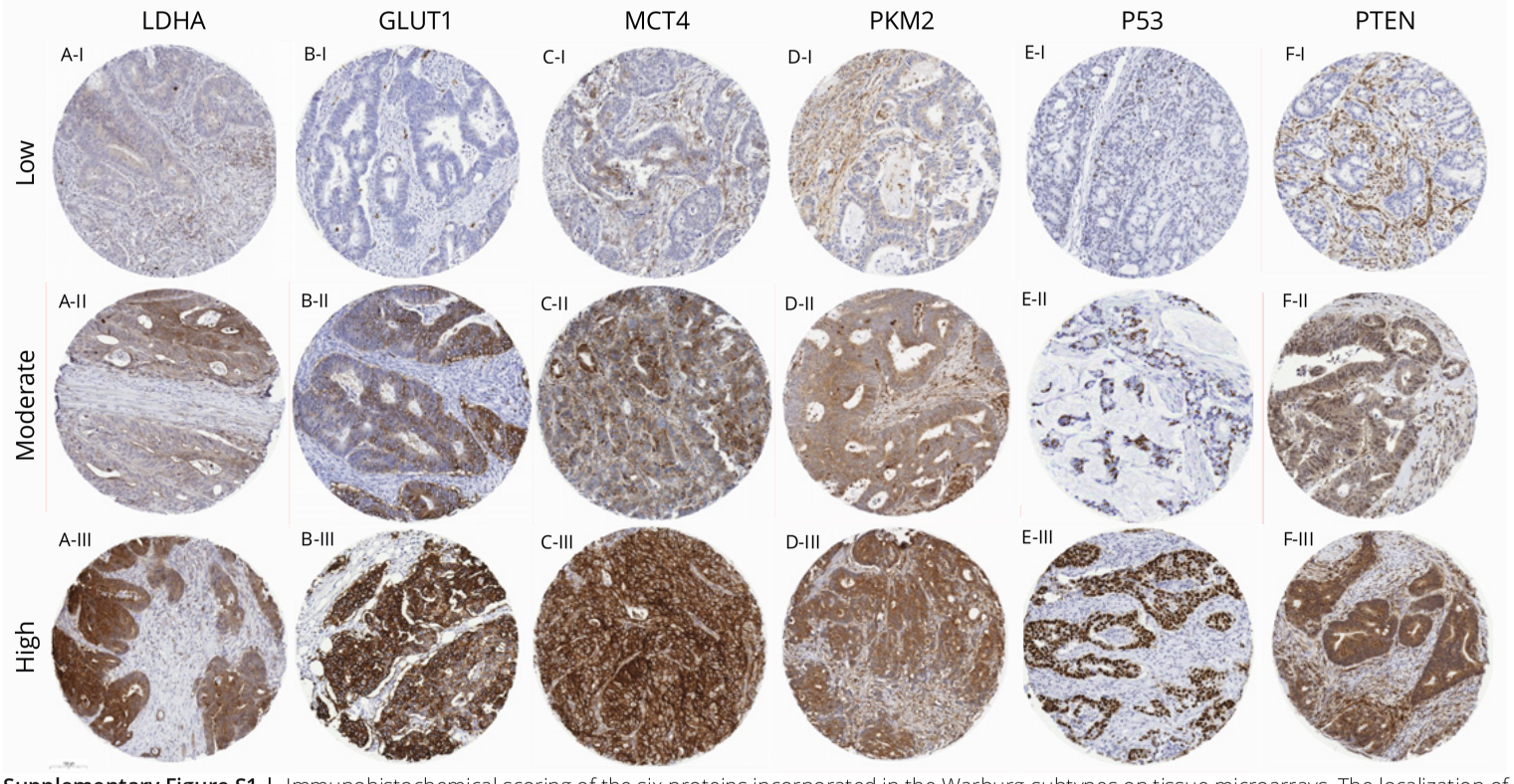
with prenatal exposure to famine in humans. Proc Natl Acad Sci USA. 2008;105(44):17046-17049.
18. Elias SG, Keinan-Boker L, Peeters PHM, et al. Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. Int J Cancer. 2004;108(4):628-630. https://doi.org/10. 1002/ijc.11584.
19. Waterland RA. Epigenetic mechanisms and gastrointestinal development. J Pediatr. 2006;149(5):S137-S142.
20. Li J, Wei Z, Cao S, et al. A pilot study on clinicopathological features and intestinal microflora changes in colorectal cancer patients born over a nine-year period encompassing three years before and after the great Chinese famine. Cancer Epidemiol. 2019;59:166-172.
21. Huang XF, Chen JZ. Obesity, the PI3K/Akt signal pathway and colon cancer. Obes Rev. 2009;10(6):610-616.
22. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330(6009):1340-1344.
23. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiother Oncol. 2009; 92(3):329-333.
24. Warburg O. The metabolism of carcinoma cells. J Cancer Res. 1925; 9(1):148-163. https://doi.org/10.1158/jcr.1925.148.
25. Schwartz L, Supuran CT, Alfarouk KO. The Warburg effect and the hallmarks of cancer. Anti-Cancer Agents Med Chem. 2017;17(2):164-170.
26. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-674.
27. Jenniskens JCA, Offermans K, Simons CCJM, et al. Energy Balance–Related factors and risk of colorectal cancer expressing different levels of proteins involved in the warburg effect. Cancer Epidemiol Biomark Prev. 2021. https://doi.org/10.1158/1055-9965 epi-21-0678.
28. van den Brandt PA, Goldbohm RA, Veer PVT, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol. 1990;43(3):285-295.
29. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika. 1986;73(1):1-11. https://doi.org/10.1093/biomet/73.1.1
30. van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch cancer registry for epidemiological research. Int J Epidemiol. 1990;19(3):553-558.
31. Goldbohm RA, van den Brandt PA, Dorant E. Estimation of the Coverage of Dutch Municipalities by Cancer Registries and PALGA Based on Hospital Discharge Data. Tijdschrift voor Soc Gezondheidszorg. 1994;72:80-84.
32. van den Brandt PA. Molecular pathological epidemiology of lifestyle factors and colorectal and renal cell cancer risk. Maastricht Pathology; 2018. 11th Joint Meeting of the British Division of the International Academy of Pathology and the Pathological Society of Great Britain & Ireland, 19-22 June 2018. J Pathol. 2018;246(Suppl 1):S9.
33 Jenniskens JCA, Offermans K, Samarska I, et al. Validity and reproducibility of immunohistochemical scoring by
. trained non-pathologists on tissue microarrays. Cancer Epidemiol Biomark Prev. 2021;30(10):1867-1874. https://doi.org/10.1158/1055-9965.epi-21-0295
34. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey L. Energy restriction early in life and colon carcinoma risk. Cancer. 2003;97(1):46-55.
35. Trienekens GM Tussen ons volk en de honger: de voedselvoorziening, 1940-1945; 1985.
36. Burger G, Drummond J, Sandstead H.

Malnutrition and Starvation in Western Netherlands. Part I. The Hague, Netherlands: General State Printing Office; 1948.
37. Lin DY, Wei LJ. The robust inference for the cox proportional hazards model. J Am Stat Assoc. 1989;84(408):1074-1078. https://doi.org/10.1080/01621459.1989.10478874.
38. Schoenfeld D Partial residuals for the proportional hazards regression model. Biometrika. 1982;69(1):239-241. https://doi.org/10.1093/biomet/69.1.239
39. De Vogel S, Bongaerts BW, Wouters KA, et al. Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. Carcinogenesis. 2008;29(9):1765-1773.
40. Wacholder S, Gail MH, Pee D, Brookmeyer R Alternative variance and efficiency calculations for the case-cohort design. Biometrika. 1989;76(1):117-123. https://doi.org/10.1093/biomet/76.1.117
41. Xu J, Galley JD, Bailey MT, Thomas-Ahner JM, Clinton SK, OlivoMarston SE. The impact of dietary energy intake early in life on the colonic microbiota of adult mice. Sci Rep. 2016;6(1):1-12.
42. Perkins JM, Subramanian SV, Davey Smith G, Özaltin E. Adult height, nutrition, and population health. Nutr Rev. 2016;74(3):149-165.
43. Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? J Natl Cancer Inst. 1988;80(10):772-775.
44. Hirsch J, Ahrens EH Jr, Blankenhorn DH. Measurement of the human intestinal length in vivo and some causes of variation. Gastroenterology. 1956;31(3):274-284.
45. Elands RJ, Simons CCJM, Dongen MV, et al. A systematic literature review and meta-regression analysis on early-life energy restriction and cancer risk in humans. PLoS One. 2016;11(9):e0158003.
46. Bianchi G, Martella R, Ravera S, et al. Fasting induces anti-Warburg effect that increases respiration but reduces ATP-synthesis to promote apoptosis in colon cancer models. Oncotarget. 2015;6(14):11806-11819.
47. Gunter MJ, Hoover DR, Yu H, et al. Insulin, insulin-like growth factor-I, endogenous estradiol, and risk of colorectal cancer in postmenopausal women. Cancer Res. 2008;68(1):329-337.
48. Dirx MJ, Van Den Brandt PA, Goldbohm RA, Lumey L. Diet in adolescence and the risk of breast cancer: results of The Netherlands cohort study. Cancer Causes Control. 1999;10(3):189-199.
49. Grillo F, Bruzzone M, Pigozzi S, et al. Immunohistochemistry on old archival paraffin blocks: is there an expiry date? J Clin Pathol. 2017;70(11):988-993.
50. Offermans K, Jenniskens JCA, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Expression of proteins associated with the Warburg-effect and survival in colorectal cancer. J Pathol: Clin Res. 2021. https://doi.org/10.1002/cjp2.250.

SUPPLEMENTARY METHODS

STAINING PROTOCOLS

All cores were scored by at least two assessors, independently and blinded for case characteristics. IHC scoring protocols are shown in Supplementary Figure S1. Cytoplasmic LDHA expression in the tumor cells was scored as (1) negative/weak cytoplasmic staining (in the absence of strong cytoplasmic staining); (2) 1-50% of tumor cells with strong cytoplasmic staining; or (3) >50% of tumor cells with strong cytoplasmic staining. The percentage of tumor cells with membranous GLUT1 or MCT4 immunoreactivity was scored as (1) negative; (2) 1-10% positive; (3) 11-50% positive; or (4) >50% positive. Cytoplasmic PKM2 expression in the tumor cells was scored as (1) negative/weak cytoplasmic staining (in the absence of moderate or strong cytoplasmic staining); (2) moderate cytoplasmic staining (in the absence of strong cytoplasmic staining); (3) 1-50% of tumor cells with strong cytoplasmic staining; or (4) strong cytoplasmic staining in ≥50% of tumor cells. The percentage of nuclear P53 positive tumor cells was scored as (1) negative; (2) 1-10% positive; (3) 11-50% positive; (4) 51-90% positive; or (5) 91-100% positive. PTEN was scored as (1) negative (no staining in tumor cytoplasm); (2) weak (staining of tumor cytoplasm weaker than adjacent stroma); (3) moderate (similar staining intensity in tumor cytoplasm and adjacent stroma); or (4) strong (staining of tumor cytoplasm stronger than adjacent stroma), irrespective of the percentage of stained tumor cells.



Supplementary Figure S1 | Immunohistochemical scoring of the six proteins incorporated in the Warburg-subtypes on tissue microarrays. The localization of the antigen-antibody complex is visualized in brown (DAB) and counterstaining is shown in blue (haematoxylin). **LDHA: A-I)** low, (any) weak cytoplasmic staining; **A-II)** moderate, 1-50% strong cytoplasmic staining; **A-III)** high, >50% strong cytoplasmic staining. **GLUT1: B-I)** low, <10% membranous staining; **B-II)** moderate, 11-50% membranous staining; **B-III)** high, >50% membranous staining. **MCT4: C-I)** low, <10% membranous staining; **C-II)** moderate, 11-50% membranous staining; **C-III)** high, >50% membranous staining. **PKM2: D-I)** low, (any) weak positive cytoplasmic staining; **D-II)** moderate, 1-50% strong cytoplasmic staining; **D-III)** high, >50% cytoplasmic staining. **P53: E-I)** low, <10% nuclear staining; **E-II)** moderate, 11-50% nuclear staining; **E-III)** high, >50% nuclear staining. **PTEN** (staining intensity of tumor cells was compared with that of stromal cells): **F-I)** low, completely negative cytoplasm; **F-II)** moderate, weak or moderate positive cytoplasm; **F-III)** high, strong cytoplasmic staining.

Supplementary Table S1 | Overview of proteins and their role in the Warburg-effect.

Protein	Full name	Cellular localization	Function	Role in Warburg-effect
PTEN	Phosphatase and tensin homolog	Cytosol, nucleus	Tumor suppressor	**Upstream regulator** PTEN is a regulator of the Warburg-effect via its effect on the PI3K/Akt-signaling pathway. Loss of PTEN tumor suppressor function leads to PI3K/Akt signaling activation, which induces the Warburg-effect via HIF-activation.
TP53	Tumor protein p53	Nucleus	Tumor suppressor	**Upstream regulator** TP53 is a transcriptional factor that regulates several glycolysis related proteins. Loss of TP53 tumor suppressor function leads to an increase of the Warburg-effect, and a decrease of the TCA cycle and oxidative phosphorylation.
GLUT1	Glucose transporter 1	Plasma membrane	Glucose transporter	**Glucose import** GLUT1 facilitates glucose uptake across the plasma membrane. Upregulation of GLUT1 provides the cell with the increased demand of glucose.
PKM2	Pyruvate Kinase M2	Cytosol	Enzyme	**Glycolytic enzyme** PKM2 catalyzes the conversion of phophoenolpyruvate and ADP into pyruvate and ATP, the final step of the glycolytic pathway.
LDHA	Lactate dehydrogenase A	Cytosol	Enzyme	**Glycolytic enzyme** LDHA facilitates conversion of pyruvate into lactate. This conversion is necessary to restore NAD+ levels and to avoid pyruvate accumulation.
MCT4	Monocarboxylate transporter 4	Plasma membrane	Monocarboxylate transporter	**Lactate export** MCT4 facilitates lactate removal across the plasma membrane. By upregulation of MCT4, the cancer cell eliminates the accumulated lactate, thereby preventing acidification of the cell.

Abbreviations: TCA, tricarboxylic acid; ADP, adenosine diphosphate; ATP, adenosine triphosphate; NAD+, nicotinamide adenine dinucleotide +.
References: the Human Protein Atlas (http://www.proteinatlas.org)[1], Feron et al[2], Levine et al[3], Bensinger et al[4].

1. Pontén F, Jirström K, Uhlen M. The Human Protein Atlas—a tool for pathology. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2008;216(4):387-93.
2. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiotherapy and oncology. 2009;92(3):329-33.
3. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330(6009):1340-4.
4. Bensinger SJ, Christofk HR, editors. New aspects of the Warburg effect in cancer cell biology. Seminars in cell & developmental biology; 2012: Elsevier.

Supplementary Table S2 | Overview staining protocols.

Antibody	Clone	Source	Dilution	Staining procedure	Antigen retrieval	Incubation time	Visualisation system	Chromogen
Pan-CK	AE1/AE3	DAKO (GA05361-2)	RTU	DAKO[a]	PT high[b]	10 min.	EnVision FLEX[f]	DAB
TP53	DO-7	DAKO (M700101-2)	RTU	DAKO[a]	PT high[b]	20 min.	EnVision FLEX[f]	DAB
PTEN	6H2.1	DAKO (M362729-2)	1:100	DAKO[a]	PT high[b]	20 min.	EnVision FLEX[f]	DAB
GLUT1	-	TFS (RB-9052-P)	1:200	DAKO[a]	PT low[c]	20 min.	EnVision FLEX[f]	DAB
LDHA	E-9	SCBT (sc-137243)	1:800	Manual	HIER high[d]	Overnight, 4^{O}C	REAL EnVision[g]	DAB
MCT4	D-1	SCBT (sc-376140)	1:100	Manual	HIER high[d]	Overnight, 4^{O}C	REAL EnVision[g]	DAB
PKM2	C-11	SCBT (sc-365684)	1:100	Manual	HIER low[e]	1 hour, 37^{O}C	LSAB2 Kit/HRP[h]	DAB

Abbreviations: TFS, Thermo Fisher Scientific; SBCT, Santa Cruz Biotechnology; RTU, ready to use; DAB, 3,3'-diaminobenzidine
[a]DAKO Autostainer Link 48
[b]High pH retrieval (K8004) for 20 minutes on the Dako PT link (Agilent Technologies)
[c]Low pH retrieval (K8005) for 20 minutes on the Dako PT link (Agilent Technologies)
[d]Heat-induced antigen retrieval using a solution of Tris/EDTA (pH 9.0)
[e]Heat-induced antigen retrieval using a solution of sodium citrate (pH 6.0)
[f]EnVision FLEX Visualization Kit (K8008, DAKO)
[g]REAL EnVision Detection System (K5007, DAKO)
[h]Universal LSAB2 kit/HRP (K0675, Agilent)

Supplementary Table S3 | Percentage of slides evaluated per assessor for the six immunohistochemical markers of proteins incorporated in the Warburg-subtypes.

	Assessor 1 Non-pathologist	**Assessor 2** Non-pathologist		**Assessor 3** Non-pathologist		**Assessor 4** Pathologist[b]
		Time point 1	Time point 2[a]	Time point 1	Time point 2[a]	
TP53	100%	100%	10%	100%	10%	10%
PTEN	100%	100%	10%	100%	10%	10%
GLUT1	25%[c]	100%	10%	100%	10%	10%
LDHA	-	100%	10%	100%	10%	10%
MCT4	-	100%	10%	100%	10%	10%
PKM2	-	100%	10%	100%	10%	10%

[a]10% randomly selected TMA sections per marker were scored for a second time after a period of at least five months to assess intra-observer reproducibility.
[b]10% randomly selected TMA sections per marker were scored by an experienced pathologist to assess inter-observer agreement between pathologist and non-pathologists.
[c]Assessor 1 left the project early because of an unforeseen work relocation.

Supplementary Table S4 | Scoring protocols of the six immunohistochemical stainings of proteins incorporated in the Warburg-subtypes, and kappa values with 95% confidence intervals for inter- and intra-observer agreement of these scoring protocols.

	P53 Nucleus	**PTEN** Cytoplasm[a]	**GLUT1** Membrane	**LDHA** Cytoplasm	**MCT4** Membrane	**PKM2** Cytoplasm
Scoring protocol[b]						
Low						
Category 1	(1) negative	(1) negative	(1) negative	(1) negative/weak	(1) negative	(1) negative/weak
Category 2	(2) 1-10% positive		(2) 1-10% positive		(2) 1-10% positive	
Moderate						
Category 2		(2) weak		(2) 1-50% strong positive		(2) moderate positive
Category 3	(3) 11-50% positive	(3) moderate	(3) 11-50% positive		(3) 11-50% positive	(3) 1-50% strong positive
High						
Category 3				(3) >50% strong positive		
Category 4	(4) 51-90% positive	(4) strong	(4) >50% positive		(4) >50% positive	(4) >50% strong positive
Category 5	(5) >90% positive					

(continued)	**P53** Nucleus	**PTEN** Cytoplasm[b]	**GLUT1** Membrane	**LDHA** Cytoplasm	**MCT4** Membrane	**PKM2** Cytoplasm
Scoring agreement	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**	**κ (95% CI)**
Interobserver agreement[c]						
Final score[d] vs pathologist						
Weighted kappa[e]	0.75 (0.72-0.79)	0.58 (0.53-0.62)	0.71 (0.67-0.74)	0.65 (0.60-0.69)	0.74 (0.71-0.77)	0.65 (0.61-0.69)
Non-weighted kappa	0.63 (0.58-0.68)	0.47 (0.41-0.52)	0.61 (0.57-0.66)	0.59 (0.54-0.64)	0.63 (0.59-0.68)	0.56 (0.51-0.60)
Intraobserver agreement[c,f]						
Non-pathologist assessor 1						
Weighted kappa[e]	0.83 (0.80-0.86)	0.69 (0.65-0.74)	0.82 (0.79-0.85)	0.78 (0.74-0.82)	0.86 (0.83-0.88)	0.70 (0.67-0.74)
Non-weighted kappa	0.73 (0.69-0.77)	0.63 (0.58-0.69)	0.75 (0.72-0.79)	0.76 (0.71-0.80)	0.79 (0.75-0.82)	0.58 (0.53-0.62)
Non-pathologist assessor 2						
Weighted kappa[e]	0.87 (0.84-0.90)	0.69 (0.64-0.74)	0.75 (0.72-0.78)	0.77 (0.73-0.81)	0.83 (0.81-0.86)	0.72 (0.68-0.75)
Non-weighted kappa	0.80 (0.76-0.84)	0.65 (0.60-0.70)	0.65 (0.60-0.69)	0.73 (0.69-0.78)	0.75 (0.71-0.79)	0.62 (0.58-0.67)

[a]For PTEN scoring, staining intensity of tumour cells was compared with that of stromal cells: (1) no PTEN staining in the tumour cells; (2) staining intensity in the tumour cells weaker than in the stromal cells; (3) similar staining intensity in tumour and stromal cells; (4) staining intensity in the tumour cells stronger than in the stromal cells.
[b]Only immunoreactivity in reported cellular localization was considered positive staining.
[c]Based on a random 10% of TMA sections
[d]The final score is based on at least two non-pathologists, with discrepancies replaced by a consensus score or pathologist's score.
[e]Weight of 0.5 for adjacent categories and 0 for non-adjacent categories.
[f]10% of TMA sections were scored for a second time after at least 2 months.

Supplementary Table S5 | Age-adjusted HRs[a] and 95% CIs for associations between adult-attained height and CRC, stratified on sex, tumor location, and Warburg-subtypes; NLCS, 1986-2006.

	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Quartiles of height (cm)										
Men – colon										
<173	170	8935	196	1.00 (ref.)	55	1.00 (ref.)	69	1.00 (ref.)	72	1.00 (ref.)
173-176	175	7680	191	1.16 (0.91-1.47)	52	1.13 (0.75-1.69)	71	1.21 (0.85-1.74)	68	1.13 (0.79-1.61)
177-181	179	7097	178	1.16 (0.91-1.48)	53	1.23 (0.83-1.84)	66	1.22 (0.85-1.75)	59	1.05 (0.72-1.52)
>181	185	7310	207	1.32 (1.04-1.67)	55	1.26 (0.84-1.87)	74	1.33 (0.93-1.89)	78	1.36 (0.96-1.93)
P-trend				0.028		0.220		0.129		0.126
Per 5 cm		31022	772	1.09 (1.02-1.16)	215	1.06 (0.96-1.17)	280	1.11 (1.01-1.22)	277	1.09 (1.00-1.20)
Men – rectum										
<173	170	8935	61	1.00 (ref.)	17	1.00 (ref.)	23	1.00 (ref.)	21	1.00 (ref.)
173-176	175	7680	51	0.97 (0.66-1.45)	16	1.10 (0.55-2.21)	18	0.92 (0.49-1.73)	17	0.93 (0.48-1.78)
177-181	179	7097	66	1.36 (0.94-1.98)	24	1.79 (0.95-3.37)	21	1.15 (0.63-2.11)	21	1.24 (0.67-2.30)
>181	185	7310	49	0.97 (0.65-1.45)	19	1.36 (0.70-2.65)	14	0.75 (0.38-1.46)	16	0.91 (0.47-1.76)
P-trend				0.622		0.162		0.588		0.973
Per 5 cm		31022	227	1.04 (0.94-1.14)	76	1.09 (0.92-1.29)	76	1.00 (0.87-1.16)	75	1.02 (0.85-1.21)
Women – colon										
<162	158	8764	140	1.00 (ref.)	30	1.00 (ref.)	51	1.00 (ref.)	59	1.00 (ref.)
162-165	164	9216	185	1.25 (0.97-1.61)	54	1.70 (1.07-2.71)	57	1.06 (0.71-1.58)	74	1.19 (0.82-1.71)
166-169	168	7771	152	1.22 (0.94-1.59)	34	1.27 (0.77-2.12)	58	1.28 (0.86-1.91)	60	1.14 (0.78-1.68)
>169	172	8477	178	1.33 (1.03-1.72)	52	1.82 (1.14-2.91)	50	1.03 (0.68-1.55)	76	1.34 (0.93-1.94)
P-trend				0.047		0.048		0.652		0.146
Per 5 cm		34228	655	1.08 (1.01-1.17)	170	1.14 (1.00-1.29)	216	1.06 (0.94-1.19)	269	1.07 (0.96-1.19)

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Women – rectum										
<162	158	8764	28	1.00 (ref.)	8	1.00 (ref.)	9	1.00 (ref.)	11	1.00 (ref.)
162-165	164	9216	23	0.78 (0.44-1.38)	7	0.84 (0.30-2.33)	7	0.74 (0.27-2.00)	9	0.78 (0.32-1.88)
166-169	168	7771	38	1.53 (0.93-2.54)	11	1.55 (0.62-3.89)	21	2.66 (1.21-5.88)	6	0.61 (0.22-1.67)
>169	172	8477	38	1.41 (0.85-2.34)	11	1.40 (0.56-3.51)	14	1.65 (0.71-3.86)	13	1.23 (0.54-2.77)
P-trend				0.040		0.276		0.030		0.710
Per 5 cm		34228	127	1.16 (0.99-1.35)	37	1.18 (0.90-1.54)	51	1.20 (0.96-1.50)	39	1.09 (0.80-1.48)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study
[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.
[b]Median height per quartile based on the subcohort.

Supplementary Table S6 | Age-adjusted HRs[a] and 95% CIs for associations between early-life energy restriction and CRC, stratified on sex, tumor location, and Warburg-subtypes; NLCS, 1986-2006.

	Person years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Place of residence during the Dutch Hunger Winter (1944-45)									
Men – colon	30174	754		211		273		270	
Non-west	15188	446	1.00 (ref.)	124	1.00 (ref.)	169	1.00 (ref.)	153	1.00 (ref.)
Western rural	4129	78	0.67 (0.50-0.88)	22	0.68 (0.42-1.10)	24	0.54 (0.34-0.84)	32	0.80 (0.53-1.20)
Western city	6524	123	0.64 (0.50-0.81)	34	0.64 (0.43-0.95)	50	0.68 (0.49-0.96)	39	0.59 (0.40-0.85)
Men – rectum	30174	227		76		76		75	
Non-west	15188	141	1.00 (ref.)	44	1.00 (ref.)	49	1.00 (ref.)	48	1.00 (ref.)
Western rural	4129	27	0.71 (0.46-1.09)	11	0.92 (0.47-1.81)	10	0.77 (0.38-1.54)	6	0.45 (0.19-1.07)
Western city	6524	36	0.59 (0.40-0.87)	12	0.63 (0.33-1.21)	12	0.57 (0.30-1.08)	12	0.57 (0.30-1.10)
Women – colon	33722	646		166		215		265	
Non-west	18083	373	1.00 (ref.)	108	1.00 (ref.)	125	1.00 (ref.)	140	1.00 (ref.)
Western rural	4851	81	0.81 (0.61-1.08)	16	0.57 (0.33-0.99)	27	0.79 (0.50-1.24)	38	1.02 (0.69-1.50)
Western city	9234	159	0.83 (0.67-1.03)	37	0.68 (0.46-1.00)	51	0.81 (0.57-1.14)	71	0.97 (0.71-1.32)
Women – rectum	33722	124		36		51		37	
Non-west	18083	74	1.00 (ref.)	23	1.00 (ref.)	25	1.00 (ref.)	26	1.00 (ref.)
Western rural	4851	11	0.59 (0.31-1.13)	4	0.71 (0.23-2.20)	5	0.73 (0.28-1.93)	2	0.32 (0.08-1.39)
Western city	9234	31	0.86 (0.55-1.34)	8	0.71 (0.31-1.62)	16	1.25 (0.66-2.36)	7	0.59 (0.24-1.44)
Place of residence during World War II (1942)									
Men – colon	23793	572		167		209		196	
Rural area	11327	287	1.00 (ref.)	78	1.00 (ref.)	104	1.00 (ref.)	105	1.00 (ref.)
Urban area	11713	266	0.89 (0.73-1.09)	83	1.02 (0.73-1.42)	100	0.93 (0.69-1.25)	83	0.76 (0.56-1.04)
Men – rectum	23973	182		62		58		62	
Rural area	11327	95	1.00 (ref.)	37	1.00 (ref.)	29	1.00 (ref.)	29	1.00 (ref.)
Urban area	11713	83	0.85 (0.62-1.17)	25	0.66 (0.39-1.11)	27	0.90 (0.53-1.55)	31	1.06 (0.63-1.78)
Women – colon	26164	505		139		158		208	
Rural area	11882	243	1.00 (ref.)	78	1.00 (ref.)	74	1.00 (ref.)	91	1.00 (ref.)
Urban area	13562	248	0.90 (0.73-1.11)	59	0.66 (0.46-0.94)	80	0.97 (0.68-1.37)	109	1.05 (0.78-1.43)
Women – rectum	26164	94		30		41		23	
Rural area	11882	42	1.00 (ref.)	13	1.00 (ref.)	19	1.00 (ref.)	10	1.00 (ref.)
Urban area	13562	48	1.02 (0.66-1.58)	17	1.20 (0.57-2.54)	19	0.87 (0.45-1.65)	12	1.17 (0.47-2.93)

(continued)	Person years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Employment status of the father during the Dutch Economic Depression (1932-40)									
Men – colon	29841	743		203		271		269	
Employed	26697	664	1.00 (ref.)	176	1.00 (ref.)	245	1.00 (ref.)	243	1.00 (ref.)
Unemployed	3145	79	0.98 (0.74-1.31)	27	1.27 (0.82-1.97)	26	0.89 (0.57-1.37)	26	0.88 (0.57-1.35)
Men – rectum	29841	219		74		75		70	
Employed	26697	169	1.00 (ref.)	66	1.00 (ref.)	68	1.00 (ref.)	62	1.00 (ref.)
Unemployed	3145	23	1.00 (0.63-1.58)	8	1.03 (0.49-2.19)	7	0.86 (0.39-1.91)	8	1.10 (0.52-2.34)
Women – colon	32597	627		161		204		262	
Employed	29046	553	1.00 (ref.)	146	1.00 (ref.)	186	1.00 (ref.)	221	1.00 (ref.)
Unemployed	3552	74	1.09 (0.81-1.47)	15	0.85 (0.48-1.49)	18	0.79 (0.47-1.32)	41	1.51 (1.03-2.20)
Women – rectum	32597	123		35		50		38	
Employed	29046	111	1.00 (ref.)	33	1.00 (ref.)	45	1.00 (ref.)	33	1.00 (ref.)
Unemployed	3552	12	0.85 (0.46-1.57)	2	0.41 (0.09-1.80)	5	0.88 (0.35-2.21)	5	1.29 (0.51-3.27)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study

[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

Supplementary Table S7 | Age-adjusted HRs[a] and 95% CIs for associations between adolescent BMI (age 20 years) and CRC, stratified on sex, tumor location, and Warburg-subtypes; NLCS, 1986-2006.

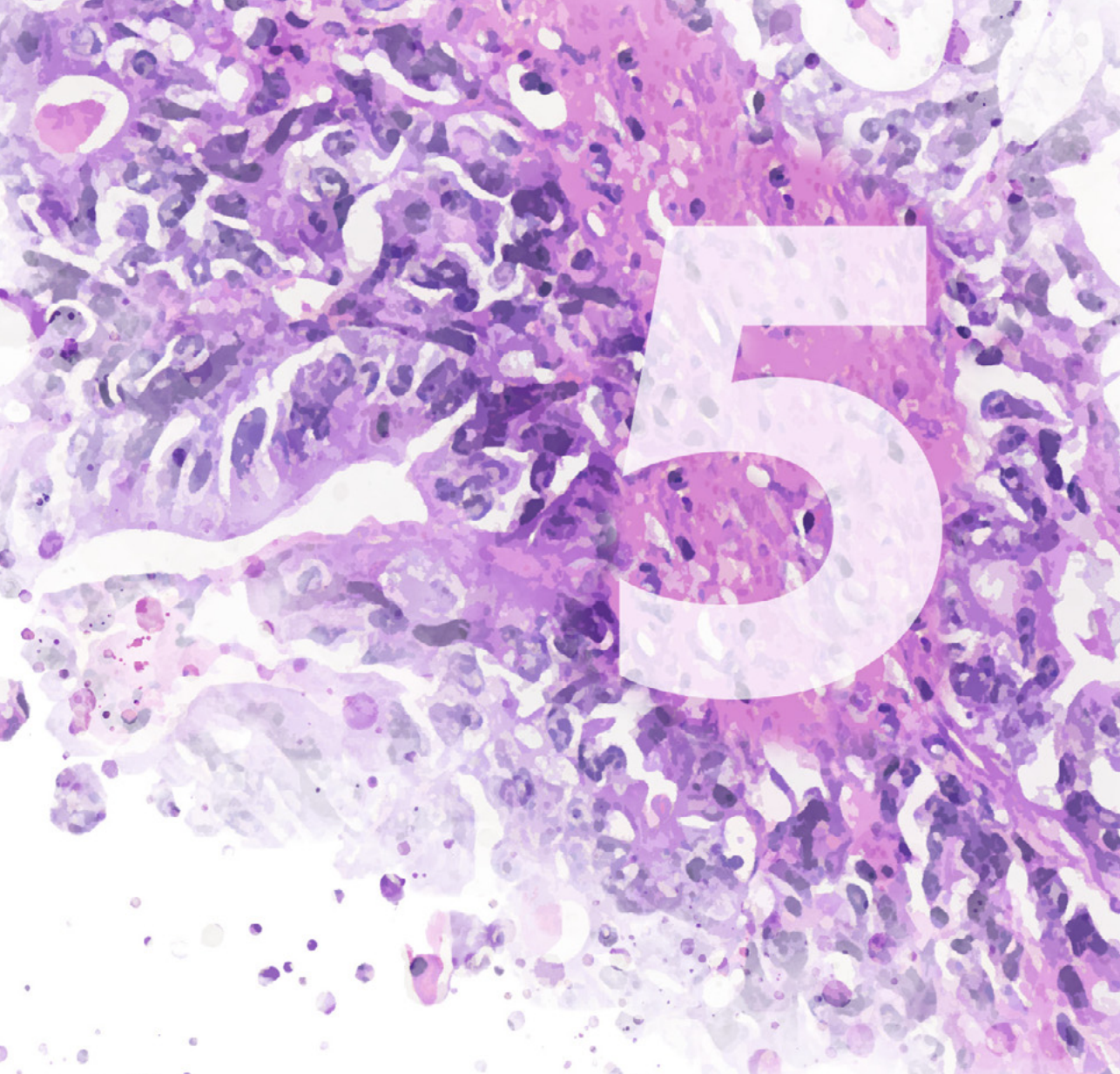
	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Quartiles of BMI at age 20 years (kg/m²)										
Men – colon										
<20.2	19.2	6103	159	1.00 (ref.)	48	1.00 (ref.)	61	1.00 (ref.)	50	1.00 (ref.)
20.2-21.6	21.0	6458	149	0.89 (0.68-1.16)	48	0.95 (0.62-1.46)	47	0.72 (0.48-1.09)	54	1.03 (0.68-1.56)
21.7-23.3	22.4	6308	157	0.95 (0.73-1.24)	41	0.82 (0.53-1.28)	62	0.98 (0.66-1.43)	54	1.04 (0.69-1.58)
>23.3	24.3	6012	181	1.18 (0.91-1.54)	42	0.91 (0.59-1.42)	65	1.09 (0.75-1.60)	74	1.55 (1.05-2.29)
P-trend				0.178		0.553		0.387		0.034
Per 5 kg/m²		24881	646	1.03 (0.85-1.25)	179	0.90 (0.65-1.25)	235	1.00 (0.75-1.34)	232	1.17 (0.88-1.56)
Men – rectum										
<20.2	19.2	6103	40	1.00 (ref.)	15	1.00 (ref.)	14	1.00 (ref.)	11	1.00 (ref.)
20.2-21.6	21.0	6458	63	1.46 (0.96-2.23)	21	1.31 (0.67-2.58)	20	1.34 (0.67-2.69)	22	1.80 (0.85-3.81)
21.7-23.3	22.4	6308	46	1.10 (0.70-1.73)	15	0.96 (0.46-2.00)	19	1.30 (0.64-2.62)	12	1.04 (0.45-2.39)
>23.3	24.3	6012	50	1.27 (0.81-1.98)	17	1.16 (0.57-2.36)	15	1.10 (0.52-2.32)	18	1.62 (0.75-3.51)
P-trend				0.610		0.929		0.843		0.422
Per 5 kg/m²		24881	199	1.12 (0.86-1.46)	68	1.16 (0.79-1.72)	68	1.04 (0.67-1.62)	63	1.17 (0.73-1.86)
Women – colon										
<19.6	18.4	7795	148	1.00 (ref.)	39	1.00 (ref.)	52	1.00 (ref.)	57	1.00 (ref.)
19.6-21.2	20.5	7731	152	1.04 (0.80-1.36)	41	1.07 (0.67-1.69)	45	0.88 (0.58-1.35)	66	1.18 (0.80-1.72)
21.3-23.0	22.0	7964	151	1.01 (0.77-1.31)	43	1.09 (0.69-1.71)	55	1.04 (0.70-1.56)	53	0.92 (0.62-1.37)
>23.0	24.2	7683	141	1.00 (0.76-1.31)	34	0.91 (0.57-1.48)	43	0.87 (0.56-1.33)	64	1.17 (0.80-1.72)
P-trend				0.920		0.762		0.724		0.708
Per 5 kg/m²		31173	592	1.03 (0.89-1.20)	157	0.96 (0.75-1.23)	195	0.95 (0.76-1.20)	240	1.15 (0.92-1.44)

(continued)	Median[b]	Person-years at risk	Total		Warburg-low		Warburg-moderate		Warburg-high	
			n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Women – rectum										
<19.6	18.4	7795	22	1.00 (ref.)	8	1.00 (ref.)	7	1.00 (ref.)	7	1.00 (ref.)
19.6-21.2	20.5	7731	37	1.72 (1.00-2.96)	11	1.40 (0.56-3.52)	13	1.93 (0.76-4.89)	13	1.88 (0.74-4.74)
21.3-23.0	22.0	7964	28	1.25 (0.71-2.22)	7	0.86 (0.31-2.38)	12	1.70 (0.66-4.36)	9	1.26 (0.47-3.40)
>23.0	24.2	7683	28	1.31 (0.74-2.32)	8	1.01 (0.37-2.71)	14	2.11 (0.84-5.32)	6	0.87 (0.29-2.61)
P-trend				0.663		0.746		0.151		0.568
Per 5 kg/m^2		31173	115	1.11 (0.85-1.45)	34	0.83 (0.48-1.43)	46	1.49 (1.06-2.10)	35	0.93 (0.59-1.46)

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), height (cm; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Median adolescent BMI per quartile based on the subcohort.



5

ENERGY BALANCE-RELATED FACTORS AND RISK OF COLORECTAL CANCER BASED ON *KRAS*, *PIK3CA* AND *BRAF* MUTATIONS AND MMR STATUS

Josien C. A. Jenniskens, Kelly Offermans, Colinda C. J. M. Simons, Iryna Samarska, Gregorio E. Fazzi, Jaleesa R. M. van der Meer, Kim M. Smits, Leo J. Schouten, Matty P. Weijenberg, Heike I. Grabsch, Piet A. van den Brandt

Journal of Cancer Research and Clinical Oncology. 2022;148:2723-2742

ABSTRACT

INTRODUCTION

KRAS mutations (*KRAS*$_{mut}$), *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, and mismatch repair deficiency (dMMR) have been associated with the Warburg-effect. We previously observed differential associations between energy balance-related factors (BMI, clothing size, physical activity) and colorectal cancer (CRC) subtypes based on the Warburg-effect. We now investigated whether associations between energy balance-related factors and risk of CRC differ between subgroups based on mutation and MMR status.

METHODS

Information on molecular features was available for 2,349 incident CRC cases within the Netherlands Cohort Study (NLCS), with complete covariate data available for 1,934 cases and 3,911 subcohort members. Multivariable-adjusted Cox-regression was used to estimate associations of energy balance-related factors with risk of CRC based on individual molecular features (*KRAS*$_{mut}$; *PIK3CA*$_{mut}$; *BRAF*$_{mut}$; dMMR) and combinations thereof (all-wild-type+MMR-proficient (pMMR); any-mutation/dMMR).

RESULTS

In men, BMI and clothing size were positively associated with risk of colon, but not rectal cancer, regardless of molecular features subgroups; the strongest associations were observed for *PIK3CA*$_{mut}$ colon cancer. In women, however, BMI and clothing size were only associated with risk of *KRAS*$_{mut}$ colon cancer (p-heterogeneity$_{\textit{KRAS}\text{mut versus all-wild-type+pMMR}}$=0.008). Inverse associations of non-occupational physical activity with risk of colon cancer were strongest for any-mutation/dMMR tumors in men and women, and specifically for *PIK3CA*$_{mut}$ tumors in women. Occupational physical activity was inversely associated with both combination subgroups of colon cancer in men.

CONCLUSION

In men, associations did not vary according to molecular features. In women, a role of *KRAS* mutations in the etiological pathway between adiposity and colon cancer is suggested, and of *PIK3CA* mutations between physical activity and colon cancer.

INTRODUCTION

Colorectal cancer (CRC) risk was shown to be affected by energy balance-related factors[1-4]. Adiposity measures, such as body mass index (BMI) and waist circumference, have been associated with an increased risk of CRC[1, 2], whereas physical activity has been associated with a decreased risk of CRC[2-4]. One of the proposed mechanisms underlying these associations is activation of the so-called Warburg-effect through upregulated PI3K/Akt-signaling[5-9]. We have previously observed differential associations between energy balance-related factors (i.e. BMI; clothing size, as a proxy for waist circumference; physical activity) and CRC subtypes expressing different levels of proteins involved in the Warburg-effect[10].

The Warburg-effect is a metabolic phenotype first discovered in the 1920s by Otto Warburg and colleagues[11]. This phenotype is characterized by increased aerobic glycolysis[6, 7] and is considered an important step in carcinogenesis[8, 9]. Mutations in well-known oncogenes *KRAS*, *PIK3CA*, and *BRAF* have been reported to drive metabolic reprogramming towards the Warburg-effect[6, 12-14]. Furthermore, we have previously shown in CRC that DNA mismatch repair deficiency (dMMR), a surrogate for microsatellite instability (MSI), was associated with the Warburg-effect[15].

MSI and *KRAS*, *PIK3CA*, and *BRAF* mutations ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, respectively) are common molecular features in CRC[16-18]. Associations between energy balance-related factors (i.e. BMI, waist circumference, physical activity) and risk of CRC in relation to $KRAS_{mut}$, $BRAF_{mut}$, and MSI/MMR status have been reported previously[19-30]. However, results thus far are inconsistent. To the best of our knowledge, there are no studies that have investigated associations between energy balance-related factors and risk of CRC in relation to $PIK3CA_{mut}$ status.

The aim of the current study was to investigate the associations of BMI, lower body clothing size (as a proxy for waist circumference), and physical activity with risk of CRC subgroups based on $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status. First, we compared CRC subgroups based on a combination of these molecular features: I) all-wild-type + pMMR — cases wild-type for all genes (*KRAS*, *PIK3CA*, and *BRAF*) and MMR-proficient (pMMR); II) any-mutation/dMMR — cases with a mutation in any of the genes (*KRAS*, *PIK3CA*, and/or *BRAF*) and/or dMMR. Second, we investigated subgroups of these molecular features individually: $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR. The all-wild-type+pMMR subgroup served as the reference group for all other subgroups.

We hypothesized that associations between energy balance-related factors and risk of CRC differ between subgroups based on $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status, which could indicate involvement of the Warburg-effect in etiological associations. We reasoned that associations with subgroups of individual molecular features ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, or dMMR) and/or with the any-mutation/dMMR subgroup, but not the all-wildtype + pMMR subgroup, give an indication of involvement of the Warburg-effect in the etiological pathway between the

exposure of interest and CRC.

MATERIALS AND METHODS

DESIGN AND STUDY POPULATION

Data from the Netherlands Cohort Study (NLCS), a large prospective cohort study, was used. At baseline (1986), 120,852 subjects aged 55–69 years completed a mailed, selfadministered questionnaire on cancer risk factors [31]. By completing and returning the questionnaire, participants agreed to participate in the study. The NLCS was approved by institutional review boards from Maastricht University and the Netherlands Organization for Applied Scientific Research. Ethical approval was obtained from the Medical Ethical Committee of Maastricht University Medical Center+ . For data processing and analysis, a case-cohort approach was used[32]. A subcohort (n = 5,000) was randomly sampled from the total cohort immediately after baseline, and accumulated person-years were estimated from this subcohort. Vital status information of subcohort members was obtained biennially by active follow-up and by linkage with municipal population registries. Incident cancer cases from the total cohort were detected through annual record linkage with the Netherlands Cancer Registry and PALGA, the nationwide Dutch Pathology Registry[33], covering 20.3 years of follow-up (September 17, 1986 until January 1, 2007). Completeness of cancer follow-up by the Netherlands Cancer Registry and PALGA was estimated to be over 96%[34]. After excluding cases and subcohort members who reported a history of cancer (except skin cancer) at baseline, a total of 4,597 incident CRC cases and 4,774 subcohort members were available (Figure 1). As described previously[10], formalin-fixed paraffin-embedded (FFPE) tissue blocks from primary tumor and matched normal colon tissue from 3,872 CRC cases were requested from participating laboratories as part of the Rainbow-TMA project during 2012–2017. Tissue blocks from 3,021 CRC cases were successfully collected from 43 pathology laboratories throughout the Netherlands (78% retrieval rate) (Figure 1).

MISMATCH REPAIR STATUS

From the FFPE blocks, 78 tissue microarrays (TMAs) were constructed sampling three 0.6 mm tumor cores from 2,694 CRC cases (Figure 1). Information on TMA construction has been published previously[10]. Five µm thick sections were cut from all TMA blocks, stained with Hematoxylin & Eosin (H&E) according to a standard protocol, and subjected to immunohistochemistry (IHC) using an automated immunostainer (DAKO Autostainer Link 48, Glostrup, Denmark). MMR status, a surrogate for the presence or absence of MSI, was assessed using IHC staining of MLH1 and MSH2 as described previously[15]. All TMA sections were scanned using an Aperio scanner (Leica Microsystems, Milton Keynes, UK) at 40× magnification at the University of Leeds (UK) Scanning Facility or at the Department of Pathology, Aachen University Hospital (Germany).

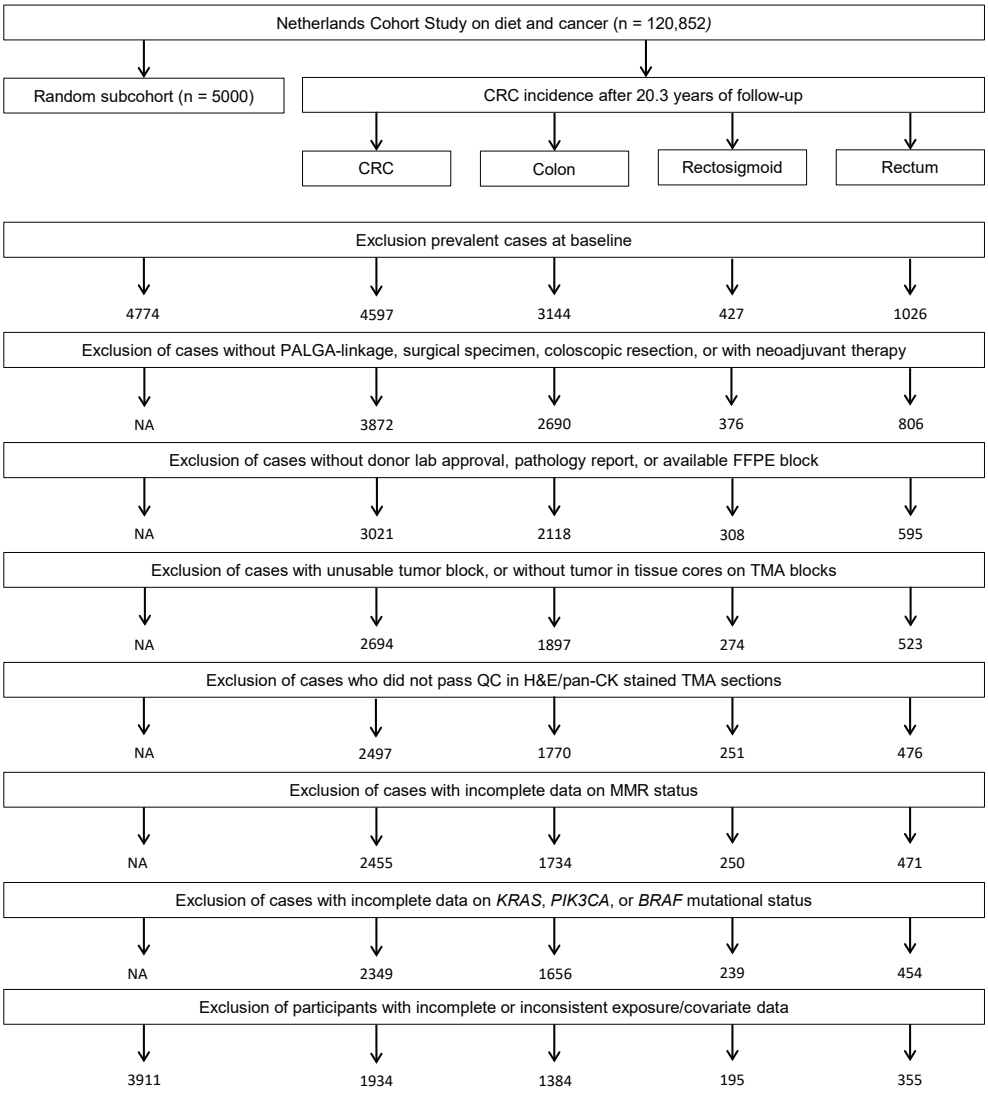


Figure 1 | Flow diagram of the number of CRC cases and subcohort members; NLCS, 1986-2006. *Abbreviations: CRC, colorectal cancer; NA, not applicable; PALGA, Dutch Pathology Registry; FFPE, formalin-fixed paraffin-embedded; TMA, tissue microarray; QC, quality control; H&E, Hematoxylin & Eosin; pan-CK, pan-cytokeratin; MMR, mismatch repair.*

H&E-stained TMA sections combined with pan-cytokeratin stained sections (if necessary) were reviewed to confirm presence of adenocarcinoma for each core. Requiring at least one core per case with adenocarcinoma, 2,497 cases passed quality control (Figure 1). IHC scoring of MLH1 and MSH2 was performed according to the protocol published by Richman et al[35] by an experienced histopathologist (HG) as well as by three trained[36] non-pathologists (G.E. Fazzi: histology technician; K. Offermans: PhD student; J.C.A. Jenniskens: PhD student). Tumors with complete loss of either MLH1 or MSH2 expression were classified as MMR-deficient (dMMR), and those expressing both MLH1 and MSH2 were classified as MMR-proficient (pMMR). MMR status information was available for 2,455 CRC cases (Figure 1).

DNA ISOLATION AND MUTATION DETECTION

For DNA extraction, two 20 µm thick sections were cut from FFPE blocks containing primary tumor. Sections were deparaffinized manually using the Buffer ATL (Cat. No. 939011, Qiagen, Hilden, Germany), Proteinase K (Cat. No. 19131, Qiagen), and the Deparaffinization Solution (Cat. No. 19093, Qiagen), using an adapted version of the manufacturer's protocol (Supplementary Methods). The QIAsymphony® DSP DNA Mini Kit (Cat. No. 937236, Qiagen) and the QIAsymphony® (Qiagen) instrument were used for DNA isolation following the manufacturer's protocol (Tissue_ HC_200 protocol). The Quantus™ Fluorometer (Promega, Madison, WI, USA) with a QuantiFluor® dsDNA system (Promega) was used to determine the double-stranded DNA concentrations. Mutations in tumor DNA were analyzed at Institut für Immunologie und Genetik (Kaiserslautern, Germany) with the ColoCarta panel (Agena Bioscience, Hamburg), which screens for 32 mutations in 6 genes (*BRAF*, *HRAS*, *KRAS*, *MET*, *NRAS*, *PIK3CA*; see Supplementary Table S1 for specific mutations) using Matrix Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) mass spectrometry. To ensure valid mutation information, the following cut-offs were used: Z-score ≥ 4.00; spectrum quality ≥ 0.750; typer peak probability ≥ 0.850; primer extension rate cut-off ≥ 0.200. Detection of mutations at a frequency of ≥ 7.5% for any of the alleles was considered evidence of a mutation in the corresponding gene. A failed reaction at a single nucleotide position resulted in missing data for the corresponding gene status only if the reactions at all other positions were wild-type.

No mutations were observed in *HRAS*, and *NRAS* mutations were found in a total of 86 cases. *NRAS* mutations were not included in the current analyses as after stratification on sex and tumor location, subgroups would have less than 50 cases (range 10–42 cases). This would have led to empty cells or cells with less than five cases for models based on categories of exposures. Complete information on *KRAS*, *PIK3CA*, and *BRAF* mutation status as well as MMR status was available for 2,349 CRC cases (Figure 1). Supplementary Table S2 shows baseline characteristics of CRC cases by availability of mutation and MMR status.

SUBGROUPS OF MOLECULAR FEATURES

The following subgroups were used for statistical analyses: (I) all-wild-type+pMMR — cases wild-type for all genes (*KRAS*, *PIK3CA*, and *BRAF*) and pMMR; (II) any-mutation/dMMR — cases with a mutation in any of the genes (*KRAS*, *PIK3CA*, and *BRAF*) and/or dMMR; (III) $KRAS_{mut}$ — cases with a (non-exclusive) *KRAS* mutation; (IV)

BRAF$_{mut}$; (V) *PIK3CA*$_{mut}$; and (VI) dMMR. Note: subgroups of individual mutation and MMR status might overlap since multiple mutations and/or dMMR can occur within the same tumor.

ENERGY BALANCE-RELATED FACTORS

Baseline questionnaires provided information on anthropometry, physical activity, diet, and other risk factors[31]. BMI at baseline (kg/m^2) was calculated using baseline weight (kg) divided by height squared (m^2). Lower body clothing size (trouser/skirt) was used as a proxy for waist circumference[37]. Non-occupational physical activity included leisure activities like walking, cycling, or doing sports, as described in more detail previously[38]. Occupational energy expenditure and sitting time were estimated for the longest held job, which was self-reported at baseline. Jobs were classified as low, moderate, or high activity, as described previously[38]. Energy expenditure was classified as <8, 8–12, and >12 kJ/minute, and sitting time as sitting for >6, 2–6, and <2 working hours/day. Data on occupational physical activity were only available for the subcohort and for cases until 17.3 years of follow-up, since funding for later data-entry and classification of occupations was unavailable. Furthermore, we did not analyze occupational physical activity measures in women because many did not have paid jobs[38].

STATISTICAL ANALYSES

After exclusion of participants with incomplete or inconsistent data on exposure variables or confounders, 3,911 subcohort members and 1,934 CRC cases were available for analyses (Figure 1). Descriptive statistics and frequency distributions were calculated for subgroups based on molecular features and cohort characteristics. Differences of molecular features between men and women and between colon and rectum were evaluated using Chi-square. Associations between energy balance-related factors and CRC subgroups based on molecular features were investigated stratified on sex and tumor location. Cox proportional hazard models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for the associations between CRC and BMI (according to sex-specific quartiles, and per 5 kg/m^2 increase), clothing size (according to sex-specific quartiles, and per 2 sizes increase), non-occupational physical activity (in categories of <30, 30–60, 60–90, >90 min per day, and per 30 min/day increase), and, for men, occupational physical activity (energy expenditure in categories of <8, 8–12, >12 kJ/minute; sitting time in categories of >6, 2–6, and <2 working hours/day). Standard errors of the HRs were estimated using the Huber-White sandwich estimator to account for additional variance introduced by sampling from the cohort[39]. The proportional hazard assumption was tested using the scaled Schoenfeld residuals[40] and by introducing time-covariate interactions into the models.

All multivariable models were adjusted for age, family history of CRC (yes/no), alcohol intake (0; 0.1–4; 5–14; >15 g/day), energy intake at baseline (kcal/day), red meat consumption (g/day), and processed meat consumption (g/day), as used previously[10]. In addition, BMI and clothing size models were adjusted for non-occupational physical activity (minutes/day), and BMI models for height (cm). All physical activity models were adjusted for BMI. Moreover, an additional analysis was conducted with

mutual adjustment for clothing size and BMI, where clothing size adjusted for BMI represents a proxy for abdominal fatness, and BMI adjusted for clothing size a proxy for subcutaneous fatness [37, 41]. Sensitivity analyses were performed excluding the first two years of follow-up.

Heterogeneity in associations between energy balancerelated factors and CRC subgroups based on molecular features was evaluated using an adapted version of the competing risks procedure in Stata developed specifically for the case-cohort design[42]. The original procedure assumes independence of both estimated HRs, which underestimates the standard error and thus overestimates the p-values for their difference. Therefore, the p-values and associated CIs were estimated based on a bootstrapping method developed specifically for the case-cohort design[43]. Each bootstrap analysis was based on 1000 replications. The all-wild-type+pMMR subgroup was the reference group for heterogeneity tests of all subgroups. Since our analyses were hypothesis-driven and exposures reflect different aspects of energy balance, we did not correct for multiple testing. All analyses were conducted in Stata Statistical Software: Release 15 (StataCorp., 2017, College Station, TX).

RESULTS

FREQUENCIES OF MOLECULAR FEATURES

In total, 1142 (59.1%) tumors had a mutation in at least one of the genes (*KRAS*, *PIK3CA*, or *BRAF*) and/or were classified as dMMR (Table 1, Figure 2a). The overall frequency of mutations and/or presence of dMMR was higher in women compared to men (66.4% vs 53.6%, respectively; p-value: <0.001), and higher in tumors located in the colon compared to the rectum (64.7% vs 43.9%, respectively; p-value: <0.001) (Table 1).

$KRAS_{mut}$-tumors were observed in 673 (34.8%) cases, $PIK3CA_{mut}$-tumors in 334 (17.3%) cases, $BRAF_{mut}$-tumors in 298 (15.4%) cases, and dMMR-tumors in 206 (10.7%) cases (Table 1, Figure 2b). The frequency of $BRAF_{mut}$-tumors and dMMR-tumors was higher in women compared to men ($BRAF_{mut}$: 22.1 vs 10.5%, p-value: <0.001; dMMR: 16.3% vs 6.5%, p-value: <0.001, respectively). $PIK3CA_{mut}$-, $BRAF_{mut}$-, and dMMR-tumors were more often observed in colon compared to rectum ($PIK3CA_{mut}$: 19.2% vs 12.7%, p-value: 0.004; $BRAF_{mut}$: 20.1% vs 3.9%, p-value: <0.001; dMMR: 14.5% vs 0.9%, p-value: <0.001, respectively) (Table 1).

Within the any-mutation/dMMR subgroup, exclusive $KRAS_{mut}$-tumors were observed in 505 (44.2%), exclusive $PIK3CA_{mut}$-tumors in 125 (11.0%), exclusive $BRAF_{mut}$-tumors in 132 (11.6%), and exclusive dMMR-tumors in 44 (3.9%) cases (Figure 2c). Combinations of $KRAS_{mut}$ and $PIK3CA_{mut}$ and of $BRAF_{mut}$ and dMMR were most common (13.0% and 10.3%, respectively). Other combinations of mutations and/or dMMR were relatively rare (i.e. <5%) (Figure 2c).

Table 1 | Frequencies[a] [n (%)] of subgroups based on mutation and MMR status in CRC cases, by tumor location and sex; NLCS, 1986–2006.

	CRC			p[d]	Colon			Rectum			p[e]
	Total n = 1,934	Men n = 1,113	Women n = 821		Total n = 1,384	Men n = 754	Women n = 630	Total n = 355	Men n = 224	Women n = 131	
All-wild-type+pMMR[b]	792 (41.0)	516 (46.4)	276 (33.6)	<0.001	488 (35.3)	309 (41.0)	179 (28.4)	199 (56.1)	135 (60.3)	64 (48.9)	<0.001
Any-mutation/dMMR[c]	1142 (59.1)	597 (53.6)	545 (66.4)		896 (64.7)	445 (59.0)	451 (71.6)	156 (43.9)	89 (39.7)	67 (51.2)	
KRAS$_{mut}$	673 (34.8)	376 (33.8)	297 (36.2)	0.275	478 (34.5)	256 (34.0)	222 (35.2)	123 (34.7)	68 (30.4)	55 (42.0)	0.969
PIK3CA$_{mut}$	334 (17.3)	196 (17.6)	138 (16.8)	0.645	266 (19.2)	150 (19.9)	116 (18.4)	45 (12.7)	30 (13.4)	15 (11.5)	0.004
BRAF$_{mut}$	298 (15.4)	117 (10.5)	181 (22.1)	<0.001	278 (20.1)	105 (13.9)	173 (27.5)	14 (3.9)	8 (3.6)	6 (4.6)	<0.001
dMMR	206 (10.7)	72 (6.5)	134 (16.3)	<0.001	201 (14.5)	70 (9.3)	131 (20.8)	3 (0.9)	1 (0.5)	2 (1.5)	<0.001

Abbreviations: (d/p)MMR, mismatch repair (deficient/proficient); CRC, colorectal cancer; NLCS, Netherlands Cohort Study; mut, mutated.
[a]Percentages might not add up because multiple molecular characteristics (e.g. *BRAF* mutation and MMR deficiency) can occur per individual.
[b]This group excludes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*), as well as MMR deficient cases.
[c]This group includes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*) and/or cases that are MMR deficient.
[d]Difference between men and women, based on total CRC, evaluated using Chi-square.
[e]Difference between colon and rectum, based on men and women combined, evaluated using Chi-square.

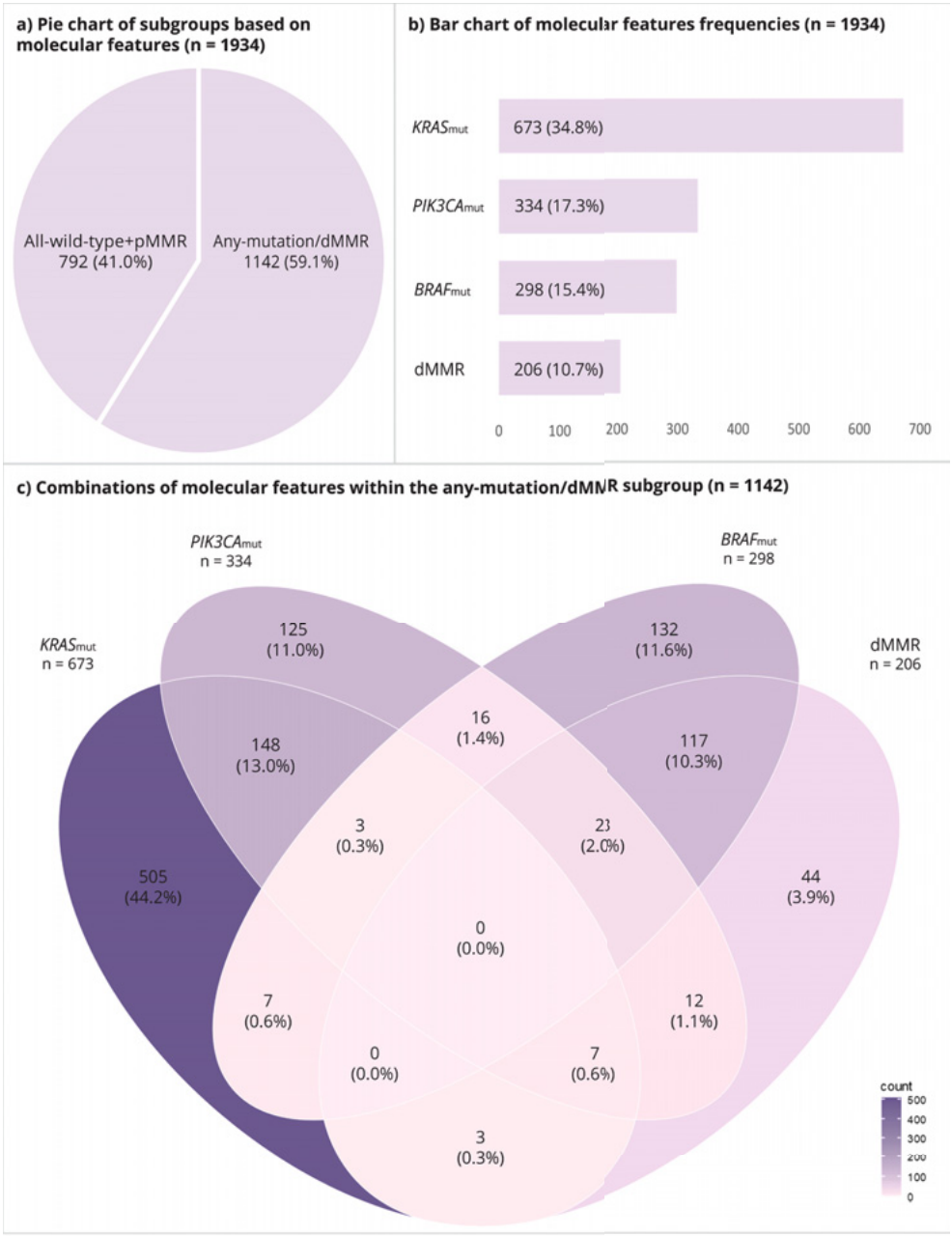


Figure 2 | Graphical presentation of $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$ and MMR status in CRC cases from the NLCS. a) Pie chart showing the distribution of the all-wild-type+pMMR and any-mutation/dMMR subgroups (based on all CRC cases; n=1,934). b) Bar chart showing frequencies of $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR (based on all CRC cases; n=1,934). c) Venn diagram showing combinations of $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR (based on any-mutation/dMMR subgroup; n=1,142). The color intensity indicates the frequency: a darker color indicates more cases; a lighter color indicates fewer cases. *Abbreviations – d/pMMR, mismatch repair deficiency/proficiency; mut, mutation; CRC, colorectal cancer; NLCS, Netherlands Cohort Study.*

COHORT CHARACTERISTICS IN SUBGROUPS OF MOLECULAR FEATURES

Information on cohort characteristics of CRC cases, overall and according to subgroups based on molecular features, is provided in Table 2. Cases in the any-mutation/dMMR subgroup were older than those in the all-wildtype+pMMR subgroup. Furthermore, cases in the any-mutation/dMMR subgroup were more often overweight compared to those in the all-wild-type+pMMR subgroup, with the exception of men with colon cancer. In general, overweight was most frequently observed amongst cases with $KRAS_{mut}$- and/or $PIK3CA_{mut}$-tumors. Similarly, the any-mutation/dMMR subgroup showed a larger mean clothing size compared to the all-wild-type+pMMR subgroup, with the exception of men with colon cancer. The mean clothing size was largest for the $KRAS_{mut}$ subgroup, again with the exception of men with colon cancer. Non-occupational physical activity was higher amongst the all-wild-type+pMMR subgroup than amongst the any-mutation/dMMR subgroup, with the exception of women with rectal cancer. In men, cases with a $PIK3CA_{mut}$-tumor in the colon were least physically active, whereas in women cases with dMMR- or $BRAF_{mut}$-tumors in the colon were least physically active. Colon cancer cases in the any-mutation/dMMR subgroup showed a higher occupational energy expenditure than those in the all-wild-type+pMMR subgroup. In particular, dMMR colon cancer cases showed the highest occupational energy expenditure and lowest occupational sitting time. In contrast, rectal cancer cases in the any-mutation/dMMR subgroup showed lower occupational energy expenditure compared to those in the allwild-type+pMMR subgroup.

ASSOCIATIONS OF ENERGY BALANCE-RELATED FACTORS AND COLORECTAL CANCER SUBGROUPS BASED ON MOLECULAR FEATURES

Multivariable-adjusted Cox-regression models on energy balance-related factors and risk of CRC subgroups based on molecular features are shown in Tables 3-6. Age-adjusted Cox-regression models are shown in Supplementary Tables S3–S6. Results of associations between energy balance-related factors and risk of CRC wild-type and MMR proficient subgroups separately are additionally presented in Supplementary Tables S7-S8. Age was included as a time-varying covariate in all models, because of violation of the proportional hazards assumption.

Adiposity

BMI and clothing size were both associated with an increased risk of overall colon cancer in men (Table 3). Associations were similarly positive for the all-wild-type+pMMR subgroup [BMI: $HR_{5kg/m2}$ (95% CI): 1.34 (1.08–1.67), p-$trend_{quartiles}$: 0.038; clothing size: $HR_{two\ sizes}$: 1.34 (1.12–1.61), p-$trend_{quartiles}$: 0.008] and the any-mutation/dMMR subgroup [BMI: $HR_{5kg/m2}$ (95% CI): 1.28 (1.07–1.53), p-$trend_{quartiles}$: 0.027; clothing size: $HR_{two\ sizes}$: 1.32 (1.11–1.55), p-$trend_{quartiles}$: 0.002]. Although positive associations were found across all subgroups of individual molecular features (Table 4), associations were strongest for the $PIK3CA_{mut}$ subgroup [BMI: $HR_{5kg/m2}$ (95% CI): 1.38 (1.05–1.82), p-$trend_{quartiles}$: 0.007; clothing size: $HR_{two\ sizes}$: 1.31 (1.01–1.70), p-$trend_{quartiles}$: 0.094], and weakest for the $BRAF_{mut}$ subgroup [BMI: $HR_{5kg/m2}$ (95% CI): 1.23 (0.87–1.72), p-$trend_{quartiles}$: 0.603; clothing size: $HR_{two\ sizes}$: 1.18 (0.85–1.64), p-$trend_{categories}$: 0.360]. In women, BMI and clothing size were not associated with risk of overall colon cancer, nor with the all-wild-type+pMMR

Table 2 | Characteristics [mean (SD) or %] of CRC cases in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Total	Wild-type+ pMMR	Any-mutation/ dMMR	*KRAS*$_{mut}$	*PIK3CA*$_{mut}$[f]	*BRAF*$_{mut}$[f]	dMMR[f]
Men - colon							
N	754	309	445	256	150	105	70
Age (years)	61.6 (4.2)	61.2 (4.2)	61.9 (4.2)	62.0 (4.2)	61.5 (4.3)	62.2 (4.1)	62.7 (4.1)
Overweight/obesity[a] (%)	52.4	52.8	52.1	52.7	56.7	48.6	54.3
Clothing size[b]	52.2 (2.6)	52.2 (2.5)	52.2 (2.7)	52.1 (2.7)	52.1 (2.7)	52.2 (2.9)	52.3 (2.8)
Non-occupational PA >60min/day (%)	50.4	53.7	48.1	51.6	42.0	49.5	55.7
Occ. energy expenditure (>12 kJ/min)[c]	11.6	8.8	13.4	13.2	13.9	11.2	17.0
Occ. sitting time (<2 hrs/day)[c]	23.2	23.5	22.9	23.5	27.1	20.2	28.8
Men - rectum							
N	224	135	89	68	30	8	1
Age (years)	60.8 (3.9)	60.4 (4.0)	61.4 (3.9)	61.7 (3.9)			
Overweight/obesity[a] (%)	48.2	46.7	50.6	52.9			
Clothing size[b]	51.7 (2.5)	51.6 (2.3)	51.8 (2.8)	52.3 (2.8)			
Non-occupational PA >60min/day (%)	59.4	60.0	58.4	54.4			
Occ. energy expenditure (>12 kJ/min)[c]	11.0	12.1	9.3	8.8			
Occ. sitting time (<2 hrs/day)[c]	30.4	31.0	29.3	29.8			

(continued)	Total	Wild-type+ pMMR	Any-mutation/ dMMR	*KRAS*$_{mut}$	*PIK3CA*$_{mut}$[f]	*BRAF*$_{mut}$[f]	dMMR[f]
Women - colon							
N	630	179	451	222	116	173	131
Age (years)	62.0 (4.1)	61.1 (3.9)	62.3 (4.1)	62.2 (4.1)	62.2 (4.2)	62.6 (4.1)	62.3 (4.0)
Overweight/obesity[a] (%)	44.8	39.7	46.8	52.7	49.1	43.4	38.2
Clothing size[b]	43.6 (3.4)	43.4 (4.1)	43.6 (3.0)	43.9 (3.2)	43.5 (2.9)	43.6 (2.8)	43.3 (3.0)
Non-occupational PA >60min/day (%)	41.8	46.4	39.9	40.1	38.8	38.7	41.2
Women - rectum							
N	131	64	67	55	15	6	2
Age (years)	61.5 (4.2)	60.9 (4.3)	62.0 (4.0)	61.7 (4.0)			
Overweight/obesity[a] (%)	49.6	48.4	50.8	52.7			
Clothing size[b]	43.5 (2.7)	43.3 (2.7)	43.8 (2.7)	43.9 (2.8)			
Non-occupational PA >60min/day (%)	42.8	37.5	47.8	47.3			

Abbreviations: SD, standard deviation; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; PA, physical activity; Occ, occupational.
[a]Body mass index ≥25.
[b]Lower body clothing size. Based on fewer participants due to extra missings.
[c]Based on fewer participants due to shorter follow-up (17.3 years), only available for men.
[d]This group excludes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*), as well as MMR deficient cases.
[e]This group includes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*) and/or cases that are MMR deficient.
[f]Analyses for subgroups with <50 cases were not performed

or anymutation/dMMR subgroups (Table 3). For individual molecular features, both BMI and clothing size were associated with an increased risk of $KRAS_{mut}$ [BMI: $HR_{5kg/m2}$ (95% CI): 1.31 (1.10–1.57), p-$trend_{quartiles}$: 0.031; clothing size: $HR_{two\ sizes}$: 1.26 (1.03–1.53), p-$trend_{quartiles}$: 0.229], but not with $PIK3CA_{mut}$, $BRAF_{mut}$, or dMMR colon cancer in women (Table 4). No associations between BMI or clothing size and risk of overall rectal cancer were observed in men or in women, and stratification on subgroups did not lead to clear associations (Tables 3-4). None of the models with mutual adjustment for BMI and clothing size showed clear associations of BMI or clothing size with CRC subgroups based on molecular features (Supplementary Tables S9–S10).

Non-occupational physical activity

Non-occupational physical activity was not associated with overall colon cancer risk in men (Table 5). However, a borderline significant inverse association was found between non-occupational physical activity and risk of the anymutation/dMMR subgroup [$HR_{30min/day}$ (95% CI): 0.97 (0.92–1.02), p-$trend_{categories}$: 0.050], whereas no association was found for the all-wild-type+pMMR subgroup. Other subgroups of molecular features in colon cancer did not show clear associations (Table 6). In contrast, non-occupational physical activity was associated with an increased risk of overall rectal cancer in men, which was stronger for the any-mutation/dMMR subgroup [$HR_{>90\ vs\ \leq 30\ min/day}$ (95% CI): 3.32 (1.28–8.60), p-$trend_{categories}$: 0.033], whereas no clear association was found for the all-wild-type+pMMR or $KRAS_{mut}$ subgroups (Tables 5, 6). However, it should be noted that the reference group (≤30 min/day) in the any-mutation/dMMR and $KRAS_{mut}$ subgroups had a limited number of cases (n = 5). In women, non-occupational physical activity was associated with a decreased risk of overall colon cancer (Table 5). Although inverse associations were found for all subgroups, most did not reach statistical significance (Tables 5, 6). Only the any-mutation/dMMR subgroup [$HR_{>90\ vs\ \leq 30\ min/day}$ (95% CI): 0.71 (0.51–0.98), p-$trend_{categories}$: 0.024] and the subgroup with a $PIK3CA_{mut}$-tumor [$HR_{>90vs\leq 30\ min/day}$ (95% CI): 0.51 (0.28–0.93), p-$trend_{categories}$: 0.042] showed statistically significant inverse associations. Non-occupational physical activity was not associated with overall rectal cancer in women, and stratification on subgroups did not lead to clear associations (Tables 5-6).

Occupational physical activity

Occupational energy expenditure was associated with a decreased risk of overall colon cancer in men (Table 5). Even though inverse associations were observed for both combination subgroups, only the association with the all-wild-type+pMMR subgroup reached statistical significance [$HR_{>12\ kJ/min}$ (95% CI): 0.51 (0.30–0.84), p-$trend_{categories}$: 0.006]. Furthermore, lower occupational sitting time was associated with a decreased risk of overall colon cancer in men (Table 5), and associations were slightly stronger for the all-wild-type+pMMR subgroup [$HR_{<2\ hrs/day}$ (95% CI): 0.56 (0.38–0.81), p-$trend_{categories}$: 0.003] compared to the any-mutation/dMMR subgroup [$HR_{<2\ hrs/day}$ (95% CI): 0.70 (0.50–0.97), p-$trend_{categories}$: 0.034]. No associations were observed for occupational physical activity measures and subgroups of individual molecular features in colon cancer (Table 6). Occupational physical activity measures were not associated with risk of rectal cancer in men, and stratification on subgroups

did not lead to clear associations (Tables 5-6).

HETEROGENEITY TESTING

For heterogeneity analyses, the all-wild-type+pMMR subgroup served as the reference group for all other subgroups (i.e. any-mutation/dMMR, $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR). Statistically significant heterogeneity was observed only for BMI associations between $KRAS_{mut}$ versus all-wild-type+pMMR colon cancer in women ($p = 0.008$), but not for any other subgroup.

SENSITIVITY ANALYSES

Sensitivity analyses excluding the first two years of follow-up did not lead to essential changes (*data not shown*).

DISCUSSION

In this large prospective cohort study, we investigated associations between energy balance-related factors and risk of CRC subgroups based on $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status. Associations between energy balancerelated factors and risk of CRC varied by abovementioned molecular features, as well by sex and tumor location. A statistically significant difference in associations was only found between all-wild-type+pMMR and $KRAS_{mut}$ subgroups of colon cancer in women regarding BMI associations. In women, we observed positive associations for BMI and clothing size with risk of $KRAS_{mut}$ colon cancer, but not with any other subgroup. In men, BMI and clothing size were positively associated with risk of colon, but not rectal cancer, regardless of molecular features subgroups. While positive associations of BMI and clothing size with risk of colon cancer were observed in men for all individual molecular features, associations were strongest for $PIK3CA_{mut}$-tumors and weakest for $BRAF_{mut}$-tumors. Non-occupational physical activity was inversely associated with any-mutation/dMMR colon cancer in men and women, but not with all-wild-type+pMMR colon cancer. In men, no clear associations were observed between non-occupational physical activity and individual molecular features in colon cancer. In women, inverse associations were observed for all individual molecular features, but associations were strongest for $PIK3CA_{mut}$ colon cancer. Occupational physical activity was associated with a decreased risk of colon cancer for both combination subgroups in men, but associations were strongest for all-wild-type+pMMR tumors.

Table 3 | Multivariable-adjusted HRs[a] and 95% CIs for associations between adiposity measures and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
BMI quartiles (kg/m²): range (median)								
Men – colon								
< 23.4 (22.2)	7993	167	1.00 (ref.)	66	1.00 (ref.)	101	1.00 (ref.)	
23.4-24.9 (24.2)	8343	188	1.06 (0.82-1.37)	79	1.10 (0.76-1.57)	109	1.04 (0.77-1.42)	
25.0-26.6 (25.7)	7683	200	1.21 (0.93-1.56)	77	1.13 (0.78-1.64)	123	1.26 (0.92-1.72)	
> 26.6 (27.8)	7003	199	1.41 (1.09-1.83)	87	1.47 (1.03-2.11)	112	1.37 (0.99-1.88)	0.710
p-trend			0.005		0.038		0.027	
per kg/m	31022	754	1.30 (1.12-1.51)	309	1.34 (1.08-1.67)	445	1.28 (1.07-1.53)	0.454
Men – rectum								
< 23.4 (22.2)	7993	58	1.00 (ref.)	34	1.00 (ref.)	24	1.00 (ref.)	
23.4-24.9 (24.2)	8343	54	0.86 (0.57-1.28)	35	0.95 (0.58-1.57)	19	0.73 (0.39-1.37)	
25.0-26.6 (25.7)	7683	65	1.15 (0.78-1.69)	42	1.29 (0.80-2.09)	23	0.95 (0.51-1.74)	
> 26.6 (27.8)	7003	47	0.93 (0.61-1.43)	24	0.82 (0.47-1.44)	23	1.09 (0.59-2.01)	0.458
p-trend			0.851		0.870		0.636	
per kg/m	31022	224	1.02 (0.81-1.28)	135	0.95 (0.71-1.26)	89	1.14 (0.80-1.62)	0.387
Women – colon								
<22.8 (21.5)	9014	181	1.00 (ref.)	56	1.00 (ref.)	125	1.00 (ref.)	
22.8-24.7 (23.8)	8914	146	0.81 (0.63-1.05)	43	0.78 (0.51-1.19)	103	0.83 (0.62-1.11)	
24.8-27.0 (25.7)	8141	147	0.92 (0.71-1.20)	36	0.73 (0.47-1.16)	111	1.01 (0.75-1.36)	
>27.0 (29.2)	8158	156	1.01 (0.77-1.31)	44	0.91 (0.59-1.41)	112	1.05 (0.78-1.43)	0.601
p-trend			0.805		0.595		0.516	
per 5 kg/m²	34228	630	1.04 (0.92-1.18)	179	0.88 (0.69-1.11)	451	1.11 (0.96-1.27)	0.081
Women – rectum								
<22.8 (21.5)	9014	37	1.00 (ref.)	16	1.00 (ref.)	21	1.00 (ref.)	
22.8-24.7 (23.8)	8914	26	0.69 (0.41-1.17)	15	0.90 (0.43-1.89)	11	0.53 (0.25-1.13)	
24.8-27.0 (25.7)	8141	32	0.91 (0.55-1.53)	18	1.16 (0.58-2.34)	14	0.75 (0.36-1.56)	
>27.0 (29.2)	8158	36	1.04 (0.63-1.72)	15	0.93 (0.45-1.90)	21	1.20 (0.61-2.38)	0.491
p-trend			0.691		0.973		0.548	
per 5 kg/m²	34228	131	1.08 (0.86-1.34)	64	1.05 (0.77-1.42)	67	1.12 (0.82-1.53)	0.872
Clothing size: range (median)								
Men - colon								
≤50 (50)	10903	211	1.00 (ref.)	90	1.00 (ref.)	121	1.00 (ref.)	
52 (52)	9750	247	1.30 (1.03-1.62)	104	1.29 (0.94-1.77)	143	1.30 (0.98-1.71)	
54 (54)	5156	136	1.36 (1.04-1.77)	53	1.26 (0.86-1.84)	83	1.43 (1.03-1.98)	
≥56 (56)	2619	90	1.80 (1.31-2.46)	39	1.87 (1.22-2.86)	51	1.75 (1.19-2.57)	0.897
p-trend			<0.001		0.008		0.002	
per 2 sizes	28428	684	1.33 (1.16-1.52)	286	1.34 (1.12-1.61)	398	1.32 (1.11-1.55)	0.983

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Clothing size: range (median)								
Men – rectum								
≤50 (50)	10903	78	1.00 (ref.)	46	1.00 (ref.)	32	1.00 (ref.)	
52 (52)	9750	69	1.00 (0.70-1.41)	46	1.12 (0.73-1.72)	23	0.80 (0.46-1.42)	
54 (54)	5156	43	1.20 (0.80-1.80)	28	1.17 (0.70-1.96)	18	1.23 (0.67-2.25)	
≥56 (56)	2619	16	0.90 (0.51-1.59)	7	0.67 (0.29-1.51)	9	1.23 (0.57-2.66)	0.454
p-trend			0.801		0.760		0.470	
per 2 sizes	28428	206	0.98 (0.81-1.20)	124	0.95 (0.75-1.21)	82	1.02 (0.74-1.41)	0.711
Women – colon								
≤40 (40)	6574	128	1.00 (ref.)	46	1.00 (ref.)	82	1.00 (ref.)	
42 (42)	8582	150	0.88 (0.67-1.17)	34	0.58 (0.36-0.93)	116	1.05 (0.76-1.46)	
44 (44)	9270	159	0.83 (0.63-1.10)	48	0.74 (0.48-1.15)	111	0.89 (0.64-1.23)	
≥46 (46)	9454	182	0.95 (0.72-1.26)	50	0.78 (0.50-1.20)	132	1.06 (0.77-1.46)	0.104
p-trend			0.764		0.537		0.979	
per 2 sizes	33880	619	1.08 (0.95-1.24)	178	1.07 (0.82-1.40)	441	1.09 (0.94-1.26)	0.759
Women – rectum								
≤40 (40)	6574	23	1.00 (ref.)	11	1.00 (ref.)	12	1.00 (ref.)	
42 (42)	8582	30	0.94 (0.53-1.67)	17	1.13 (0.52-2.49)	13	0.77 (0.34-1.73)	
44 (44)	9270	35	1.01 (0.58-1.75)	20	1.28 (0.62-2.65)	15	0.79 (0.35-1.77)	
≥46 (46)	9454	42	1.14 (0.66-1.97)	16	0.96 (0.44-2.08)	26	1.31 (0.62-2.78)	0.319
p-trend			0.532		0.941		0.355	
per 2 sizes	33880	130	0.99 (0.80-1.23)	64	0.96 (0.71-1.28)	66	1.02 (0.76-1.39)	0.562

Abbreviations: HR, hazard ratio; CI, confidence i nterval; C RC, c olorectal c ancer; (d/p)MMR, m ismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; BMI, body mass index; p-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate. BMI models were additionally adjusted for height (cm; continuous).

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*), as well as MMR deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*) and/or cases that are MMR deficient

Table 4 | Multivariable-adjusted HRs[a] and 95% CIs for associations between adiposity measures and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
BMI quartiles (kg/m²): range (median)									
Men – colon									
< 23.4 (22.2)	7993	61	1.00 (ref.)	27	1.00 (ref.)	25	1.00 (ref.)	19	1.00 (ref.)
23.4-24.9 (24.2)	8343	58	0.89 (0.60-1.32)	38	1.37 (0.81-2.30)	28	1.11 (0.62-1.96)	13	0.68 (0.33-1.42)
25.0-26.6 (25.7)	7683	69	1.12 (0.76-1.65)	42	1.66 (0.98-2.81)	29	1.22 (0.68-2.20)	21	1.21 (0.63-2.34)
> 26.6 (27.8)	7003	68	1.30 (0.88-1.92)	43	1.97 (1.17-3.32)	23	1.14 (0.61-2.14)	17	1.17 (0.59-2.31)
p-trend			0.112		0.007		0.603		0.378
per 5 kg/m²	31022	256	1.25 (1.00-1.57)	150	1.38 (1.05-1.82)	105	1.23 (0.87-1.72)	70	1.51 (1.01-2.26)
Men – rectum									
< 23.4 (22.2)	7993	18	1.00 (ref.)						
23.4-24.9 (24.2)	8343	13	0.67 (0.31-1.44)						
25.0-26.6 (25.7)	7683	19	1.06 (0.53-2.15)						
> 26.6 (27.8)	7003	18	1.21 (0.59-2.47)						
p-trend			0.415						
per 5 kg/m²	31022	68	1.17 (0.79-1.73)						
Women – colon									
<22.8 (21.5)	9014	52	1.00 (ref.)	30	1.00 (ref.)	53	1.00 (ref.)	43	1.00 (ref.)
22.8-24.7 (23.8)	8914	46	0.88 (0.58-1.35)	27	0.89 (0.52-1.53)	40	0.76 (0.49-1.19)	34	0.78 (0.49-1.26)
24.8-27.0 (25.7)	8141	65	1.48 (0.99-2.20)	30	1.08 (0.64-1.84)	39	0.82 (0.52-1.29)	24	0.61 (0.36-1.03)
>27.0 (29.2)	8158	59	1.37 (0.90-2.08)	29	1.01 (0.60-1.73)	41	0.90 (0.57-1.41)	30	0.78 (0.48-1.30)
p-trend			0.031		0.798		0.678		0.221
per 5 kg/m²	34228	222	1.31 (1.10-1.57)*	116	1.09 (0.84-1.42)	173	0.99 (0.81-1.22)	131	0.90 (0.70-1.15)
Women – rectum									
<22.8 (21.5)	9014	18	1.00 (ref.)						
22.8-24.7 (23.8)	8914	7	0.40 (0.17-0.98)						
24.8-27.0 (25.7)	8141	9	0.57 (0.24-1.33)						
>27.0 (29.2)	8158	21	1.46 (0.72-2.96)						
p-trend			0.312						
per 5 kg/m²	34228	55	1.21 (0.87-1.67)						
Clothing size: range (median)									
Men - colon									
≤50 (50)	10903	73	1.00 (ref.)	40	1.00 (ref.)	30	1.00 (ref.)	18	1.00 (ref.)
52 (52)	9750	84	1.26 (0.89-1.78)	52	1.45 (0.94-2.24)	29	1.04 (0.61-1.78)	23	1.37 (0.71-2.64)
54 (54)	5156	48	1.33 (0.88-2.02)	22	1.20 (0.69-2.09)	21	1.40 (0.77-2.55)	12	1.37 (0.64-2.95)
≥56 (56)	2619	29	1.63 (1.00-2.65)	17	1.80 (0.98-3.31)	9	1.19 (0.54-2.61)	7	1.55 (0.63-3.81)
p-trend			0.040		0.094		0.360		0.280
per 2 sizes	28428	234	1.24 (1.01-1.54)	131	1.31 (1.01-1.70)	89	1.18 (0.85-1.64)	60	1.33 (0.90-1.96)

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Clothing size: range (median)									
Men – rectum									
≤50 (50)	10903	21	1.00 (ref.)						
52 (52)	9750	17	0.90 (0.45-1.77)						
54 (54)	5156	17	1.71 (0.86-3.40)						
≥56 (56)	2619	9	1.78 (0.78-4.06)						
p-trend			0.073						
per 2 sizes	28428	64	1.17 (0.80-1.73)						
Women – colon									
≤40 (40)	6574	35	1.00 (ref.)	24	1.00 (ref.)	29	1.00 (ref.)	32	1.00 (ref.)
42 (42)	8582	54	1.13 (0.72-1.78)	25	0.75 (0.42-1.34)	47	1.23 (0.75-2.03)	35	0.82 (0.49-1.37)
44 (44)	9270	58	1.10 (0.70-1.72)	33	0.84 (0.48-1.45)	45	1.02 (0.62-1.68)	25	0.51 (0.29-0.88)
≥46 (46)	9454	70	1.33 (0.86-2.05)	30	0.74 (0.42-1.30)	49	1.11 (0.67-1.83)	37	0.76 (0.45-1.28)
p-trend			0.229		0.423		0.957		0.174
per 2 sizes	33880	217	1.26 (1.03-1.53)	112	1.05 (0.82-1.35)	170	1.04 (0.84-1.27)	129	0.90 (0.71-1.14)
Women – rectum									
≤40 (40)	6574	10	1.00 (ref.)						
42 (42)	8582	9	0.65 (0.26-1.62)						
44 (44)	9270	12	0.76 (0.32-1.84)						
≥46 (46)	9454	23	1.41 (0.63-3.13)						
p-trend			0.243						
per 2 sizes	33880	54	1.07 (0.77-1.51)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; BMI, body mass index; P-het, P-heterogeneity.

*Statistically significant p-heterogeneity, p=0.008 (reference group: wild-type for *KRAS*, *PIK3CA*, and *BRAF*, and pMMR). Note: other p-heterogeneity tests were not statistically significant.

[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate. BMI models were additionally adjusted for height (cm; continuous).

[b]Analyses for subgroups with <50 cases were not performed.

Table 5 | Multivariable-adjusted HRs[a] and 95% CIs for associations between physical activity measures and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Non-occupational physical activity (min/day): range (median)								
Men – colon								
≤ 30 (21.4)	4997	132	1.00 (ref.)	49	1.00 (ref.)	83	1.00 (ref.)	
31-60 (42.9)	10100	242	0.89 (0.69-1.16)	94	0.95 (0.64-1.39)	148	0.86 (0.63-1.17)	
61-90 (73.6)	6001	156	0.99 (0.75-1.32)	62	1.08 (0.71-1.63)	94	0.95 (0.67-1.33)	
> 90 (130.0)	9925	224	0.85 (0.65-1.11)	104	1.10 (0.76-1.60)	120	0.71 (0.51-0.97)	0.232
p-trend			0.356		0.402		0.050	
per 30 min/day	31022	754	0.99 (0.95-1.03)	309	1.02 (0.97-1.07)	445	0.97 (0.92-1.02)	0.204
Men – rectum								
≤ 30 (21.4)	4997	18	1.00 (ref.)	13	1.00 (ref.)	5	1.00 (ref.)	
31-60 (42.9)	10100	73	1.92 (1.12-3.30)	41	1.49 (0.78-2.85)	32	3.03 (1.16-7.89)	
61-90 (73.6)	6001	57	2.57 (1.47-4.47)	38	2.33 (1.21-4.47)	19	3.13 (1.15-8.52)	
> 90 (130.0)	9925	76	2.09 (1.22-3.59)	43	1.62 (0.85-3.08)	33	3.32 (1.28-8.60)	0.450
p-trend			0.012		0.104		0.033	
per 30 min/day	31022	224	1.04 (0.98-1.09)	135	1.04 (0.96-1.11)	89	1.03 (0.96-1.11)	0.850
Women – colon								
≤ 30 (19.3)	7756	169	1.00 (ref.)	52	1.00 (ref.)	117	1.00 (ref.)	
31-60 (42.9)	10923	198	0.83 (0.65-1.06)	44	0.58 (0.38-0.89)	154	0.94 (0.71-1.25)	
61-90 (75.0)	8000	148	0.84 (0.64-1.09)	47	0.85 (0.56-1.30)	101	0.83 (0.61-1.13)	
> 90 (115.7)	7550	115	0.70 (0.53-0.93)	36	0.69 (0.44-1.08)	79	0.71 (0.51-0.98)	0.145
p-trend			0.021		0.344		0.024	
per 30 min/day	34228	630	0.97 (0.91-1.03)	179	0.98 (0.88-1.10)	451	0.96 (0.89-1.03)	0.623
Women – rectum								
≤ 30 (19.3)	7756	31	1.00 (ref.)	14	1.00 (ref.)	17	1.00 (ref.)	
31-60 (42.9)	10923	44	1.03 (0.63-1.67)	26	1.30 (0.66-2.56)	18	0.78 (0.39-1.55)	
61-90 (75.0)	8000	34	1.06 (0.64-1.75)	12	0.79 (0.36-1.76)	22	1.27 (0.67-2.39)	
> 90 (115.7)	7550	22	0.72 (0.41-1.26)	12	0.83 (0.38-1.82)	10	0.61 (0.28-1.37)	0.214
p-trend			0.285		0.325		0.563	
per 30 min/day	34228	131	1.00 (0.88-1.14)	64	1.07 (0.89-1.28)	67	0.93 (0.79-1.08)	0.206
Occupational energy expenditure (kJ/min)								
Men - colon								
< 8	15144	365	1.00 (ref.)	152	1.00 (ref.)	213	1.00 (ref.)	
8-12	6368	133	0.83 (0.65-1.05)	54	0.80 (0.57-1.12)	79	0.86 (0.64-1.15)	
> 12	3561	66	0.71 (0.52-0.97)	20	0.51 (0.30-0.84)	46	0.85 (0.59-1.23)	0.201
p-trend	25073	564	0.017	226	0.006	338	0.274	

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Occupational energy expenditure (kJ/min)								
Men – rectum								
< 8	15144	107	1.00 (ref.)	65	1.00 (ref.)	42	1.00 (ref.)	
8-12	6368	57	1.35 (0.95-1.91)	35	1.38 (0.89-2.13)	22	1.30 (0.75-2.24)	
> 12	3561	21	0.84 (0.51-1.39)	14	0.91 (0.49-1.70)	7	0.73 (0.33-1.61)	
p-trend	25073	185	0.905	114	0.746	71	0.801	0.956
Occupational sitting time (hrs/day)								
Men – colon								
> 6	6511	187	1.00 (ref.)	85	1.00 (ref.)	102	1.00 (ref.)	
2-6	11617	244	0.70 (0.55-0.88)	87	0.55 (0.39-0.77)	157	0.82 (0.62-1.09)	
< 2	6944	133	0.63 (0.48-0.83)	54	0.56 (0.38-0.81)	79	0.70 (0.50-0.97)	
p-trend	25073	564	0.001	226	0.003	338	0.034	0.102
Men – rectum								
> 6	6511	60	1.00 (ref.)	39	1.00 (ref.)	21	1.00 (ref.)	
2-6	11617	69	0.62 (0.43-0.89)	40	0.55 (0.35-0.87)	29	0.75 (0.42-1.33)	
< 2	6944	56	0.88 (0.60-1.30)	35	0.84 (0.52-1.37)	21	0.96 (0.52-1.79)	
p-trend	25073	185	0.541	114	0.500	71	0.912	0.730

Abbreviations: HR, hazard ratio; CI, confdence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (defcient/profcient); NLCS, Netherlands Cohort Study; p-het, p-heterogeneity

[a]Hazard Ratios were adjusted for age (years; continuous), BMI (kg/m^2; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1–4; 5–14;>15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*), as well as MMR defcient cases

[c]This group includes cases with mutations in any of the genes (*KRAS*, *PIK3CA*, or *BRAF*) and/or cases that are MMR defcient

Table 6 | Multivariable-adjusted HRs[a] and 95% CIs for associations between physical activity measures and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Non-occupational physical activity (min/day): range (median)									
Men – colon									
≤ 30 (21.4)	4997	41	1.00 (ref.)	24	1.00 (ref.)	19	1.00 (ref.)	13	1.00 (ref.)
31-60 (42.9)	10100	83	0.97 (0.64-1.46)	63	1.30 (0.79-2.13)	34	0.87 (0.49-1.55)	18	0.66 (0.31-1.40)
61-90 (73.6)	6001	63	1.31 (0.85-2.02)	28	0.97 (0.55-1.72)	19	0.83 (0.43-1.61)	16	1.00 (0.47-2.15)
> 90 (130.0)	9925	69	0.84 (0.55-1.28)	35	0.73 (0.42-1.26)	33	0.82 (0.46-1.48)	23	0.80 (0.39-1.64)
p-trend			0.575		0.041		0.548		0.925
per 30 min/day	31022	256	0.95 (0.89-1.01)	150	0.96 (0.87-1.06)	105	1.02 (0.94-1.11)	70	1.02 (0.92-1.13)
Men – rectum									
≤ 30 (21.4)	4997	5	1.00 (ref.)						
31-60 (42.9)	10100	26	2.38 (0.90-6.31)						
61-90 (73.6)	6001	16	2.66 (0.96-7.40)						
> 90 (130.0)	9925	21	2.04 (0.76-5.45)						
p-trend			0.372						
per 30 min/day	31022	68	0.98 (0.89-1.08)						
Women – colon									
≤ 30 (19.3)	7756	55	1.00 (ref.)	36	1.00 (ref.)	47	1.00 (ref.)	32	1.00 (ref.)
31-60 (42.9)	10923	78	1.01 (0.69-1.47)	35	0.73 (0.45-1.20)	59	0.91 (0.60-1.37)	45	1.00 (0.63-1.61)
61-90 (75.0)	8000	48	0.84 (0.56-1.27)	28	0.79 (0.47-1.32)	35	0.72 (0.45-1.16)	34	1.00 (0.60-1.67)
> 90 (115.7)	7550	41	0.80 (0.52-1.23)	17	0.51 (0.28-0.93)	32	0.70 (0.44-1.14)	20	0.64 (0.36-1.14)
p-trend			0.197		0.042		0.095		0.150
per 30 min/day	34228	222	0.98 (0.90-1.08)	116	0.90 (0.77-1.04)	173	0.94 (0.84-1.05)	131	0.97 (0.85-1.09)
Women – rectum									
≤ 30 (19.3)	7756	14	1.00 (ref.)						
31-60 (42.9)	10923	15	0.79 (0.37-1.68)						
61-90 (75.0)	8000	20	1.40 (0.71-2.78)						
> 90 (115.7)	7550	6	0.45 (0.17-1.18)						
p-trend			0.366						
per 30 min/day	34228	55	0.86 (0.74-0.99)						
Occupational energy expenditure (kJ/min)									
Men - colon									
< 8	15144	115	1.00 (ref.)	72	1.00 (ref.)	59	1.00 (ref.)	32	1.00 (ref.)
8-12	6368	50	1.04 (0.72-1.49)	26	0.81 (0.50-1.32)	18	0.68 (0.38-1.20)	16	1.08 (0.56-2.08)
> 12	3561	25	0.93 (0.58-1.49)	16	0.84 (0.47-1.50)	10	0.65 (0.32-1.33)	10	1.02 (0.45-2.27)
p-trend	25073		0.855	114	0.425	87	0.139	58	0.919

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Occupational energy expenditure (kJ/min)									
Men – rectum									
< 8	15144	31	1.00 (ref.)						
8-12	6368	17	1.43 (0.77-2.64)						
> 12	3561	5	0.70 (0.27-1.82)						
p-trend	25073	53	0.892						
Occupational sitting time (hrs/day)									
Men – colon									
> 6	6511	57	1.00 (ref.)	34	1.00 (ref.)	22	1.00 (ref.)	14	1.00 (ref.)
2-6	11617	87	0.81 (0.56-1.17)	48	0.76 (0.48-1.20)	47	1.15 (0.67-1.96)	27	0.98 (0.49-1.94)
< 2	6944	46	0.75 (0.49-1.14)	32	0.84 (0.51-1.39)	18	0.73 (0.37-1.41)	17	0.99 (0.46-2.13)
p-trend	25073	190	0.181	114	0.508	87	0.325	58	0.992
Men – rectum									
> 6	6511	17	1.00 (ref.)						
2-6	11617	20	0.61 (0.32-1.19)						
< 2	6944	16	0.92 (0.46-1.86)						
p-trend	25073	53	0.826						

Abbreviations: HR, hazard ratio; CI, confdence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (defcient/profcient); NLCS, Netherlands Cohort Study; p-het, p-heterogeneity

p-heterogeneity tests (reference group for all tests: wild-type for *KRAS*, *PIK3CA*, and *BRAF*, and pMMR) were not statistically signifcant.

[a]Hazard ratios were adjusted for age (years; continuous), BMI (kg/m^2), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1–4; 5–14;>15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate

[b]Analyses for subgroups with < 50 cases were not performed

Several studies have focused on investigating associations between energy balance-related factors (i.e. BMI, waist circumference, physical activity) and risk of CRC in relation to specific (individual) mutations and/or MSI/MMR status, but results have been inconsistent[19-27]. To our knowledge, the current study is the first to combine cases into subgroups based on *KRAS*$_{mut}$, *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, and MMR status, and study potential etiological differences between these subgroups. Instead of comparing wild-type versus mutated tumors for individual genes and proficient versus deficient tumors for MMR, as done in previous studies, the all-wildtype+pMMR subgroup served as the reference group for all other subgroups in the current study. Combining mutation and MMR status into subgroups has some advantages. First, it has been suggested that mutations in *KRAS*, *PIK3CA*, and *BRAF* drive metabolic reprogramming toward the Warburg-effect[6, 12-14], and we have shown previously that MMR deficiency is associated with presence of the Warburg-effect[15]. Combining these molecular features, presumed to be involved in the same metabolic phenotype, thus results in a cleaner reference group compared to groups based on individual features (e.g. *KRAS* mutated versus wild-type). Our results show that co-occurrence of *KRAS*$_{mut}$ and *PIK3CA*$_{mut}$ is relatively common, as is co-occurrence of *BRAF*$_{mut}$ and dMMR. Using the all-wild-type+pMMR subgroup as the reference for all subgroups of individual mutations and MMR status, this reference group is less heterogeneous compared to, e.g., the *KRAS* wild-type (*KRAS*$_{wt}$) group, which still contains a large number of cases with a *PIK3CA* mutation. Second, differentiating subgroups on the basis of the combination of presence or absence of mutations and/or dMMR leads to increased statistical power, since most individual molecular features occurred in <20% of CRC cases (e.g., MMR deficiency: 10.7%).

Previous studies on adiposity and risk of CRC in relation to molecular features mainly focused on BMI[19-30], though some used additional adiposity measures like waist circumference[21, 23, 30]. Two cohort studies[20, 21] and two case–control studies[19, 22] investigated adiposity in relation to *KRAS*$_{mut}$ status in CRC. Our results are in line with those of Slattery et al[22], which showed positive associations of adiposity with *KRAS*$_{mut}$ but not *KRAS*$_{wt}$ colon cancer in women, whereas similar associations were observed for *KRAS*$_{mut}$ and *KRAS*$_{wt}$ in men. A study by Brändstedt et al[21] also reported positive associations between adiposity and *KRAS*$_{mut}$ but not *KRAS*$_{wt}$ CRC, but in men, not women. These and our results are in contrast with those of Carr et al[19] and Myte et al[20], who reported positive associations of adiposity with *KRAS*$_{wt}$ CRC (note: *KRAS*$_{wt}$+ *BRAF*$_{wt}$ in the study by Myte et al) but no or weak associations with *KRAS*$_{wt}$ CRC. Three cohort studies [20, 21, 23], including one study that used data from the NLCS with 7.3 years of follow-up[23], and two case–control studies[19, 28] studied adiposity in relation to *BRAF*^{mut} status in CRC. Our results are in line with all but one of these studies[20, 21, 23, 28], as these reported either a weaker positive association of adiposity with *BRAF*$_{mut}$ compared to *BRAF*$_{wt}$ CRC[21, 23], or no association with *BRAF*$_{mut}$ CRC[20, 28]. Even though Carr et al[19] observed this same difference in associations for men, associations between adiposity and CRC were

stronger for *BRAF*$_{mut}$ CRC than *BRAF*$_{wt}$ CRC in women. For MSI/MMR status, our results are in line with those of a recent meta-analysis by Carr et al[27], in which no difference in associations was observed between adiposity and MSI status in CRC. Our study is the first to investigate the association between adiposity and CRC risk in relation to *PIK3CA*$_{mut}$ status, and therefore cannot be compared to any previous data.

To our knowledge, associations between physical activity and colon cancer risk in relation to molecular features have only been investigated in a case–control study by Slattery et al for *KRAS*$_{mut}$[22], *BRAF*$_{mut}$[28], and MSI[26] status. Our results are partly in line with these studies, which showed stronger positive associations between physical inactivity and risk of *KRAS*$_{mut}$ colon cancer compared to *KRAS*$_{wt}$ colon cancer in men, whereas associations did not differ according to *KRAS*$_{mut}$ status in women[22]. For *BRAF*, they observed no association between physical activity and *BRAF*$_{mut}$ colon cancer[28]. Lastly, physical activity was associated with both MSS and MSI colon cancer in men, but only with MSS colon cancer in women[26]. Our results for *PIK3CA*$_{mut}$ CRC cannot be compared to any previous data, since studies investigating associations between physical activity and *PIK3CA*$_{mut}$ status in CRC are currently lacking.

The contradicting results across molecular pathological epidemiology (MPE) studies regarding associations of energy balance-related factors with risk of CRC according to *KRAS*$_{mut}$, *BRAF*$_{mut}$, and/or MSI/MMR status might be attributed to several factors. For example: use of different methods for assessing molecular features (e.g. assessment of different mutations or MSI versus MMR status); different timing and method of exposure measurements (i.e. BMI, waist circumference, physical activity); different study designs (i.e. cohort versus case–control); different approaches for (outcome) stratification (for example stratification on sex and tumor location); and/or chance findings due to multiple testing, caused by repeatedly splitting CRC into different molecular pathological subgroups. We therefore believe it is important that large prospective cohort studies replicate the current analyses, preferably stratified on tumor location and sex.

The current results suggest a role of *KRAS* mutations in the etiological pathway between adiposity and colon cancer risk in women (adiposity was only associated with *KRAS*$_{mut}$ colon cancers). In contrast, our results do not indicate a clear role of one of the molecular features in the etiological pathway between adiposity and colon cancer in men (adiposity was associated with all subgroups of molecular features in colon cancer). As mentioned above, the molecular features used in the current study have all been associated with the Warburg-effect[6, 12-15]. Associations with the all-wild-type+pMMR group indicate a low likelihood of Warburg-effect involvement, whereas associations with the any-mutation/dMMR subgroup or subgroups of individual molecular features

indicate a higher likelihood of Warburg-effect involvement. Therefore, the current results indicate a potential role of the Warburg-effect in the etiological pathway between adiposity and colon cancer in women through *KRAS* mutations, but not other molecular features. In men, a role of the Warburg-effect in the etiological pathway between adiposity and colon cancer is not indicated by the current results. In a previous study, we investigated associations between energy balancerelated factors and risk of Warburg-subtypes in CRC, based IHC expression of proteins involved in the Warburg-effect[10]. The results of this previous study indicated involvement of the Warburg-effect in associations between adiposity and colon cancer risk in both men and women, though additional mechanisms could be at play in women as well.

For physical activity, the current results indicate a role of molecular features ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and/ or MMR deficiency) in the etiological pathway between physical inactivity and colon cancer risk in women (physical activity was associated with any-mutation/dMMR colon cancer), and it seems that in particular *PIK3CA* mutations are involved in this association (strongest association observed with $PIK3CA_{mut}$ colon cancer). In men, the current results do not give a clear indication of involvement of molecular features in the association between physical activity and colon cancer. While non-occupational physical activity was inversely associated with the any-mutation/ dMMR subgroup, occupational physical activity was mainly associated with the all-wild-type+pMMR subgroup. It is assumed that occupational physical activity gives a better indication of physical activity for men than non-occupational physical activity. That is, while occupational physical activity represents long-term physical activity (median duration of longest held job: 29 years), non-occupational physical activity probably reflects the last few years before baseline. Therefore, the current results suggest that the molecular features studied here are not involved in the etiological pathway between physical inactivity and colon cancer risk in men. All in all, the current results indicate involvement of the Warburg-effect in associations between physical activity and colon cancer risk in women, but not men. Results of our previous study on Warburg-subtypes in CRC indicated that inverse associations between physical activity and colon cancer risk are explained by mechanisms other than the Warburg-effect[10].

Altogether, results from our previous study on Warburgsubtypes in CRC are only partly in line with the current results. Although the molecular features that were considered in the current study have been associated with the Warburg-effect[6, 12-15], they are additionally known for their involvement in numerous diverse (oncogenic) cellular pathways for cell growth, differentiation, proliferation, and survival[16-18]. Therefore, the molecular features used in the current study might not always be a good reflection of the Warburg-effect. Furthermore, tumors of cases in the all-wild-type+pMMR subgroup might express other molecular features, possibly also associated with the Warburg-effect, that were not assessed in the current study. This may have potentially influenced our results. Still, combining these molecular features into all-wild-type+pMMR and any-mutation/dMMR subgroups seemed to be a straightforward way of subgrouping CRC cases, especially for physical activity associations.

A major strength of the current study is the prospective cohort design with long follow-up (20.3 years) and availability of DNA from FFPE tumor material from a large number of incident CRC cases. Another strength was the detection of mutations using MassARRAY technology, which has been shown to be a suitable technique for mutation typing in (older) FFPE material[44]. The ColoCarta panel that was used includes assays for most of the *KRAS* (99%) and *BRAF* (98%) mutations, but it identifies only 78% of known *PIK3CA* mutations[45]. However, the most common *PIK3CA* mutations are included[46]. This makes it unlikely that additional detection of less common mutations would alter the current results, since the number of additional cases with a *PIK3CA* mutation would be rather small. As an indicator of MSI status, we used IHC expression of MLH1 and MSH2, which might have led to misclassification of some of the cases. However, it has been shown that loss of MLH1 or MSH2 expression was observed in ~ 90% of MSI cases[47].

In conclusion, results from this large prospective cohort study provide further insights in the associations between energy balance-related factors and CRC risk according to $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status. Associations between energy balance-related factors and risk of CRC varied by these molecular features, as well by sex and tumor location. Our results suggest a role of *KRAS* mutations in the etiological pathway between adiposity and colon cancer in women. For men, our results do not indicate a role of one of the molecular features in the etiological pathway of adiposity and colon cancer. Furthermore, the current results indicate a role of mutations in *KRAS*, *PIK3CA*, and/or *BRAF*, and/or MMR deficiency in the etiological pathway between physical inactivity and colon cancer risk in women, but not men, and it seems that in particular *PIK3CA* mutations are involved in this association. Our findings need to be replicated in additional large-scale MPE-studies.

REFERENCES

1. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiology and Prevention Biomarkers. 2007;16(12):2533-47.
2. Robsahm TE, Aagnes B, Hjartåker A, Langseth H, Bray FI, Larsen IK. Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. European Journal of Cancer Prevention. 2013;22(6):492-505.
3. Wolin KY, Yan Y, Colditz GA, Lee I. Physical activity and colon cancer prevention: a meta-analysis. British journal of cancer. 2009;100(4):611-6.
4. Samad A, Taylor R, Marshall T, Chapman MA. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Colorectal disease. 2005;7(3):204-13.
5. Huang XF, Chen JZ. Obesity, the PI3K/Akt signal pathway and colon cancer. Obesity reviews. 2009;10(6):610-6.
6. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330(6009):1340-4.
7. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiotherapy and oncology. 2009;92(3):329-33.
8. Schwartz L, T Supuran C, O Alfarouk K. The Warburg effect and the hallmarks of cancer. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents). 2017;17(2):164-70.
9. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. cell. 2011;144(5):646-74.
10. Jenniskens JCA, Offermans K, Simons CC, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors and risk of colorectal cancer expressing different levels of proteins involved in the Warburg-effect. Cancer Epidemiology Biomarkers & Prevention. 2021:cebp.EPI-21-0678-E.2021.
11. Warburg O. The metabolism of carcinoma cells. The Journal of Cancer Research. 1925;9(1):148-63.
12. Kimmelman AC. Metabolic dependencies in RAS-driven cancers. Clinical Cancer Research. 2015;21(8):1828-34.
13. Hutton JE, Wang X, Zimmerman LJ, Slebos RJ, Trenary IA, Young JD, Li M, Liebler DC. Oncogenic KRAS and BRAF drive metabolic reprogramming in colorectal cancer. Molecular & cellular proteomics. 2016;15(9):2924-38.
14. Jiang W, He T, Liu S, Zheng Y, Xiang L, Pei X, Wang Z, Yang H. The PIK3CA E542K and E545K mutations promote glycolysis and proliferation via induction of the β-catenin/SIRT3 signaling pathway in cervical cancer. Journal of hematology & oncology. 2018;11(1):1-15.
15. Offermans K, Jenniskens JCA, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Expression of proteins associated with the Warburg-effect and survival in colorectal cancer. J Pathol Clin Res. 2021.
16. Li W, Qiu T, Dong L, Zhang F, Guo L, Ying J. Prevalence and characteristics of PIK3CA mutation in mismatch repair-deficient colorectal cancer. Journal of Cancer. 2020;11(13):3827.
17. Haluska F, Pemberton T, Ibrahim N, Kalinsky K, editors. The RTK/RAS/BRAF/PI3K pathways in melanoma: biology, small molecule inhibitors, and potential applications. Seminars in oncology; 2007: Elsevier.
18. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology. 2010;138(6):2073-87. e3.

19. Carr PR, Amitay EL, Jansen L, Alwers E, Roth W, Herpel E, Kloor M, Schneider M, Bläker H, Chang-Claude J. Association of BMI and major molecular pathological markers of colorectal cancer in men and women. The American Journal of Clinical Nutrition. 2020.
20. Myte R, Gylling B, Häggström J, Häggström C, Zingmark C, Löfgren Burström A, Palmqvist R, Van Guelpen B. Metabolic factors and the risk of colorectal cancer by KRAS and BRAF mutation status. Int J Cancer. 2019;145(2):327-37.
21. Brändstedt J, Wangefjord S, Nodin B, Eberhard J, Sundström M, Manjer J, Jirström K. 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(6):e98964.
22. Slattery ML, Anderson K, Curtin K, Ma K-N, Schaffer D, Edwards S, Samowitz W. Lifestyle factors and Ki-ras mutations in colon cancer tumors. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2001;483(1-2):73-81.
23. Hughes LA, Williamson EJ, van Engeland M, Jenkins MA, Giles GG, Hopper JL, Southey MC, Young JP, Buchanan DD, Walsh MD. Body size and risk for colorectal cancers showing BRAF mutations or microsatellite instability: a pooled analysis. International journal of epidemiology. 2012;41(4):1060-72.
24. Campbell PT, Jacobs ET, Ulrich CM, Figueiredo JC, Poynter JN, McLaughlin JR, Haile RW, Jacobs EJ, Newcomb PA, Potter JD. Case–control study of overweight, obesity, and colorectal cancer risk, overall and by tumor microsatellite instability status. Journal of the National Cancer Institute. 2010;102(6):391-400.
25. Hoffmeister M, Bläker H, Kloor M, Roth W, Toth C, Herpel E, Frank B, Schirmacher P, Chang-Claude J, Brenner H. Body mass index and microsatellite instability in colorectal cancer: a population-based study. Cancer Epidemiology and Prevention Biomarkers. 2013;22(12):2303-11.
26. Slattery ML, Curtin K, Anderson K, Ma K-N, Ballard L, Edwards S, Schaffer D, Potter J, Leppert M, Samowitz WS. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. Journal of the National Cancer Institute. 2000;92(22):1831-6.
27. Carr P, Alwers E, Bienert S, Weberpals J, Kloor M, Brenner H, Hoffmeister M. Lifestyle factors and risk of sporadic colorectal cancer by microsatellite instability status: a systematic review and meta-analyses. Annals of Oncology. 2018;29(4):825-34.
28. Slattery ML, Curtin K, Sweeney C, Levin TR, Potter J, Wolff RK, Albertsen H, Samowitz WS. Diet and lifestyle factor associations with CpG island methylator phenotype and BRAF mutations in colon cancer. International journal of cancer. 2007;120(3):656-63.
29. Hanyuda A, Ogino S, Qian ZR, Nishihara R, Song M, Mima K, Inamura K, Masugi Y, Wu K, Meyerhardt JA. Body mass index and risk of colorectal cancer according to tumor lymphocytic infiltrate. International journal of cancer. 2016;139(4):854-68.
30. Brändstedt J, Wangefjord S, Borgquist S, Nodin B, Eberhard J, Manjer J, Jirström K. Influence of anthropometric factors on tumour biological characteristics of colorectal cancer in men and women: a cohort study. Journal of translational medicine. 2013;11(1):1-13.
31. van den Brandt PA, Goldbohm RA, Veer PVT, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. Journal of clinical epidemiology. 1990;43(3):285-95.
32. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika. 1986;73(1):1-11.
33. van den Brandt PA, Schouten LJ, Goldbohm

RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19(3):553-8.
34. Goldbohm RA, van den Brandt PA, Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. 1994.
35. Richman SD, Adams R, Quirke P, Butler R, Hemmings G, Chambers P, Roberts H, James MD, Wozniak S, Bathia R. Pre-trial inter-laboratory analytical validation of the FOCUS4 personalised therapy trial. Journal of clinical pathology. 2016;69(1):35-41.
36. Jenniskens JCA, Offermans K, Samarska I, Fazzi GE, Simons CCJM, Smits KM, Schouten LJ, Weijenberg MP, van den Brandt PA, Grabsch HI. Validity and reproducibility of immunohistochemical scoring by trained non-pathologists on Tissue MicroArrays. Cancer Epidemiology and Prevention Biomarkers. 2021.
37. Hughes LA, Schouten LJ, Goldbohm RA, van den Brandt PA, Weijenberg MP. Self-reported clothing size as a proxy measure for body size. Epidemiology. 2009;20(5):673-6.
38. Simons CCJM, Hughes LA, Van Engeland M, Goldbohm RA, Van Den Brandt PA, Weijenberg MP. Physical activity, occupational sitting time, and colorectal cancer risk in the Netherlands cohort study. American journal of epidemiology. 2013;177(6):514-30.
39. Lin DY, Wei L-J. The robust inference for the Cox proportional hazards model. Journal of the American statistical Association. 1989;84(408):1074-8.
40. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982;69(1):239-41.
41. Janssen I, Heymsfield SB, Allison DB, Kotler DP, Ross R. Body mass index and waist circumference independently contribute to the prediction of nonabdominal, abdominal subcutaneous, and visceral fat. The American journal of clinical nutrition. 2002;75(4):683-8.
42. De Vogel S, Bongaerts BW, Wouters KA, Kester AD, Schouten LJ, de Goeij AF, de Bruïne AP, Goldbohm RA, van den Brandt PA, van Engeland M. Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. Carcinogenesis. 2008;29(9):1765-73.
43. Wacholder S, Gail MH, Pee D, Brookmeyer R. Alternative variance and efficiency calculations for the case-cohort design. Biometrika. 1989;76(1):117-23.
44. Fleitas T, Ibarrola-Villava M, Ribas G, Cervantes A. MassARRAY determination of somatic oncogenic mutations in solid tumors: Moving forward to personalized medicine. Cancer treatment reviews. 2016;49:57-64.
45. Fumagalli D, Gavin PG, Taniyama Y, Kim S-I, Choi H-J, Paik S, Pogue-Geile KL. A rapid, sensitive, reproducible and cost-effective method for mutation profiling of colon cancer and metastatic lymph nodes. BMC cancer. 2010;10(1):101.
46. Gray RT, Cantwell MM, Coleman HG, Loughrey MB, Bankhead P, McQuaid S, O'neill RF, Arthur K, Bingham V, McGready C. Evaluation of PTGS2 expression, PIK3CA mutation, aspirin use and colon cancer survival in a population-based cohort study. Clinical and translational gastroenterology. 2017;8(4):e91.
47. Lanza G, Gafà R, Maestri I, Santini A, Matteuzzi M, Cavazzini L. Immunohistochemical pattern of MLH1/MSH2 expression is related to clinical and pathological features in colorectal adenocarcinomas with microsatellite instability. Modern Pathology. 2002;15(7):741.

SUPPLEMENTARY METHODS

DEPARAFFINIZATION OF FFPE SECTIONS

FFPE tissue sections containing tumor primary tumor were deparaffinized using an adapted version of the protocol for Purification of genomic DNA from FFPE tissue using the QIAamp® DNA FFPE Tissue Kit and Deparaffinization Solution (Qiagen, Hilden, Germany). The adapted protocol included the following steps: I) Add 320 µl Deparaffinization Solution to 2 x 20 µm sections and vortex vigorously for 10 s. Centrifuge briefly to collect the sample in the bottom of the tube; II) Incubate at 56°C for 3 min, and then allow to cool at room temperature (15–25°C); III) Add 200 µl Buffer ATL, and mix by vortexing; IV) Centrifuge for 1 min at 11,000 x g (10,000 rpm); V) Add 20 µl proteinase K to the lower clear phase and mix gently by pipetting up and down; VI) Incubate at 56°C for 1 h; VII) Mix clear phase by pipetting up and down; VIII) Incubate at 90°C for 1 h; IX) Briefly centrifuge the 1.5 ml tube to remove drops from inside the lid; X) Transfer the 2 ml microcentrifuge tubes to QiaSymphony.

Supplementary Table S1 | ColoCarta panel genes and mutations.

Gene	Assay	Mutation
BRAF	15/16	*V600E/K/L/M/R*
	9	*D594G/V*
KRAS	1	*G12A/D/V*
	2	*G12C/R/S*
	4	*G13D/V*
	5	*A59T*
	7	*Q61L/P/R*
	8	*Q61H_A/H_G*
PIK3CA	1	*R88Q*
	3	*C420R*
	5	*E542K*
	6	*E545K*
	7	*Q546K*
	8	*H701P*
	9	*H1047L/R*
NRAS	1	*G12A/D/V*
	2	*G12C/R/S*
	3	*G13A/D/V*
	4	*G13C/R/S*
	7	*Q61H*
	8	*Q61E/K*
HRAS	6	*Q61L/P/R*
MET	1	*R970C*
	2	*T992I*

Supplementary Table S2 | Baseline characteristics [mean (SD) or %] of CRC cases by availability of mutation and MMR status; NLCS, 1986-2006.

	Information on mutation and MMR status	
	Not available	**Available**
N	2248	2349
Men (%)	55.7	56.1
Age	62.1 (4.2)	61.8 (4.1)
Overweight/obesity[a] (%)	49.0	50.0
Clothing size[b]	48.1 (5.0)	48.2 (5.1)
Non-occupational physical activity >60 min/day	47.5	46.6
Occupational energy expenditure >12 kJ/min[c]	11.2	10.7
Occupational sitting time <2 hours/day[c]	33.1	32.7
Height (cm)	171.8 (8.4)	172.1 (8.5)
Total energy intake (kcal/day)	1936 (504)	1944 (492)
Family history of colorectal cancer (%)	8.4	9.5
Alcohol consumption (g/day)	12.2 (16.2)	11.4 (14.8)
Processed meat intake (g/day)	14.0 (15.7)	13.2 (14.2)
Red meat intake (g/day)	87.7 (41.2)	87.5 (39.6)
Never cigarette smokers (%)	32.9	32.1
University or higher vocational education (%)	14.3	15.1

Abbreviations: SD, standard deviation; CRC, colorectal cancer; MMR, mismatch repair; NLCS, Netherlands Cohort Study.
[a]Body mass index ≥25.
[b]Lower body clothing size. Based on fewer participants due to extra missings.
[c]Based on fewer participants due to shorter follow-up (17.3 years), only available for men.

Supplementary Table S3 | Age-adjusted HRs[a] and 95% CIs for associations between adiposity measures and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
BMI quartiles (kg/m²): range (median)							
Men – colon							
< 23.4 (22.2)	7993	167	1.00 (ref.)	66	1.00 (ref.)	101	1.00 (ref.)
23.4-24.9 (24.2)	8343	188	1.06 (0.83-1.36)	79	1.13 (0.79-1.60)	109	1.02 (0.75-1.38)
25.0-26.6 (25.7)	7683	200	1.23 (0.96-1.57)	77	1.20 (0.84-1.72)	123	1.24 (0.92-1.67)
> 26.6 (27.8)	7003	199	1.35 (1.06-1.74)	87	1.51 (1.06-2.14)	112	1.25 (0.92-1.70)
p-trend			0.008		0.021		0.069
per kg/m	31022	754	1.28 (1.11-1.48)	309	1.35 (1.09-1.67)	445	1.23 (1.04-1.46)
Men – rectum							
< 23.4 (22.2)	7993	58	1.00 (ref.)	34	1.00 (ref.)	24	1.00 (ref.)
23.4-24.9 (24.2)	8343	54	0.88 (0.59-1.30)	35	0.97 (0.60-1.59)	19	0.75 (0.40-1.38)
25.0-26.6 (25.7)	7683	65	1.16 (0.80-1.70)	42	1.29 (0.80-2.07)	23	0.99 (0.55-1.77)
> 26.6 (27.8)	7003	47	0.93 (0.62-1.40)	24	0.81 (0.47-1.40)	23	1.08 (0.60-1.95)
p-trend			0.881		0.824		0.622
per kg/m	31022	224	1.02 (0.82-1.27)	135	0.95 (0.72-1.25)	89	1.13 (0.81-1.58)
Women – colon							
<22.8 (21.5)	9014	181	1.00 (ref.)	56	1.00 (ref.)	125	1.00 (ref.)
22.8-24.7 (23.8)	8914	146	0.80 (0.62-1.03)	43	0.77 (0.51-1.17)	103	0.81 (0.61-1.09)
24.8-27.0 (25.7)	8141	147	0.89 (0.69-1.15)	36	0.71 (0.46-1.10)	111	0.97 (0.73-1.30)
>27.0 (29.2)	8158	156	0.94 (0.73-1.20)	44	0.87 (0.57-1.31)	112	0.97 (0.73-1.29)
p-trend			0.770		0.432		0.890
per 5 kg/m²	34228	630	1.02 (0.90-1.15)	179	0.87 (0.69-1.09)	451	1.08 (0.95-1.24)
Women – rectum							
<22.8 (21.5)	9014	37	1.00 (ref.)	16	1.00 (ref.)	21	1.00 (ref.)
22.8-24.7 (23.8)	8914	26	0.71 (0.42-1.19)	15	0.96 (0.47-1.95)	11	0.52 (0.25-1.10)
24.8-27.0 (25.7)	8141	32	0.96 (0.59-1.57)	18	1.26 (0.64-2.49)	14	0.73 (0.37-1.46)
>27.0 (29.2)	8158	36	1.07 (0.67-1.73)	15	1.05 (0.52-2.14)	21	1.09 (0.59-2.03)
p-trend			0.565		0.696		0.661
per 5 kg/m²	34228	131	1.10 (0.89-1.36)	64	1.08 (0.80-1.46)	67	1.12 (0.84-1.49)
Clothing size: range (median)							
Men - colon							
≤50 (50)	10903	211	1.00 (ref.)	90	1.00 (ref.)	121	1.00 (ref.)
52 (52)	9750	247	1.27 (1.02-1.59)	104	1.27 (0.94-1.73)	143	1.28 (0.97-1.67)
54 (54)	5156	136	1.33 (1.02-1.72)	53	1.23 (0.85-1.78)	83	1.40 (1.02-1.92)
≥56 (56)	2619	90	1.76 (1.29-2.39)	39	1.82 (1.20-2.77)	51	1.71 (1.18-2.50)
p-trend			<0.001		0.011		0.002
per 2 sizes	28428	684	1.31 (1.15-1.49)	286	1.32 (1.11-1.58)	398	1.30 (1.11-1.53)

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Clothing size: range (median)							
Men – rectum							
≤50 (50)	10903	78	1.00 (ref.)	46	1.00 (ref.)	32	1.00 (ref.)
52 (52)	9750	69	0.99 (0.70-1.40)	46	1.14 (0.74-1.74)	23	0.79 (0.46-1.36)
54 (54)	5156	43	1.18 (0.79-1.76)	25	1.18 (0.71-1.97)	18	1.17 (0.65-2.11)
≥56 (56)	2619	16	0.88 (0.50-1.55)	7	0.67 (0.30-1.51)	9	1.16 (0.54-2.47)
p-trend			0.881		0.782		0.580
per 2 sizes	28428	206	0.97 (0.80-1.19)	124	0.96 (0.75-1.22)	82	1.00 (0.73-1.37)
Women – colon							
≤40 (40)	6574	128	1.00 (ref.)	46	1.00 (ref.)	82	1.00 (ref.)
42 (42)	8582	150	0.88 (0.67-1.16)	34	0.57 (0.36-0.91)	116	1.06 (0.77-1.45)
44 (44)	9270	159	0.84 (0.64-1.11)	48	0.74 (0.48-1.13)	111	0.90 (0.65-1.24)
≥46 (46)	9454	182	0.93 (0.71-1.22)	50	0.76 (0.50-1.16)	132	1.04 (0.76-1.42)
p-trend			0.664		0.481		0.938
per 2 sizes	33880	619	1.08 (0.94-1.23)	178	1.06 (0.82-1.39)	441	1.08 (0.94-1.24)
Women – rectum							
≤40 (40)	6574	23	1.00 (ref.)	11	1.00 (ref.)	12	1.00 (ref.)
42 (42)	8582	30	0.99 (0.57-1.74)	17	1.20 (0.55-2.59)	13	0.81 (0.37-1.80)
44 (44)	9270	35	1.06 (0.62-1.84)	20	1.31 (0.62-2.76)	15	0.85 (0.39-1.86)
≥46 (46)	9454	42	1.25 (0.74-2.13)	16	1.04 (0.48-2.28)	26	1.43 (0.70-2.89)
p-trend			0.348		0.920		0.239
per 2 sizes	33880	130	1.02 (0.83-1.26)	64	0.98 (0.72-1.32)	66	1.07 (0.81-1.43)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; BMI, body mass index.

[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR deficient.

Supplementary Table S4 | Age-adjusted HRs[a] and 95% CIs for associations between adiposity measures and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
BMI quartiles (kg/m²): range (median)									
Men – colon									
< 23.4 (22.2)	7993	61	1.00 (ref.)	27	1.00 (ref.)	25	1.00 (ref.)	19	1.00 (ref.)
23.4-24.9 (24.2)	8343	58	0.90 (0.61-1.31)	38	1.33 (0.80-2.22)	28	1.05 (0.60-1.84)	13	0.65 (0.32-1.33)
25.0-26.6 (25.7)	7683	69	1.15 (0.80-1.67)	42	1.60 (0.97-2.64)	29	1.18 (0.68-2.05)	21	1.11 (0.59-2.10)
> 26.6 (27.8)	7003	68	1.26 (0.87-1.83)	43	1.81 (1.09-2.99)	23	1.03 (0.57-1.84)	17	1.00 (0.51-1.95)
p-trend			0.122		0.013		0.808		0.659
per kg/m	31022	256	1.24 (1.00-1.53)	150	1.33 (1.03-1.72)	105	1.18 (0.86-1.63)	70	1.38 (0.92-2.07)
Men – rectum									
< 23.4 (22.2)	7993	18	1.00 (ref.)						
23.4-24.9 (24.2)	8343	13	0.68 (0.33-1.40)						
25.0-26.6 (25.7)	7683	19	1.08 (0.56-2.08)						
> 26.6 (27.8)	7003	18	1.12 (0.58-2.19)						
p-trend			0.494						
per kg/m	31022	68	1.14 (0.79-1.64)						
Women – colon									
<22.8 (21.5)	9014	52	1.00 (ref.)	30	1.00 (ref.)	53	1.00 (ref.)	43	1.00 (ref.)
22.8-24.7 (23.8)	8914	46	0.88 (0.58-1.33)	27	0.89 (0.52-1.53)	40	0.74 (0.48-1.13)	34	0.77 (0.48-1.24)
24.8-27.0 (25.7)	8141	65	1.37 (0.93-2.02)	30	1.10 (0.65-1.85)	39	0.80 (0.52-1.24)	24	0.61 (0.36-1.02)
>27.0 (29.2)	8158	59	1.23 (0.83-1.83)	29	1.05 (0.62-1.78)	41	0.83 (0.54-1.28)	30	0.75 (0.46-1.22)
p-trend			0.094		0.688		0.469		0.168
per 5 kg/m²	34228	222	1.27 (1.07-1.51)	116	1.13 (0.88-1.47)	173	0.98 (0.80-1.19)	131	0.89 (0.70-1.14)
Women – rectum									
<22.8 (21.5)	9014	18	1.00 (ref.)						
22.8-24.7 (23.8)	8914	7	0.39 (0.16-0.94)						
24.8-27.0 (25.7)	8141	9	0.55 (0.24-1.24)						
>27.0 (29.2)	8158	21	1.28 (0.67-2.43)						
p-trend			0.410						
per 5 kg/m²	34228	55	1.20 (0.89-1.63)						
Clothing size: range (median)									
Men - colon									
≤50 (50)	10903	73	1.00 (ref.)	40	1.00 (ref.)	30	1.00 (ref.)	18	1.00 (ref.)
52 (52)	9750	84	1.23 (0.88-1.73)	52	1.43 (0.93-2.20)	29	1.04 (0.61-1.76)	23	1.35 (0.71-2.55)
54 (54)	5156	48	1.33 (0.89-1.97)	22	1.15 (0.67-1.98)	21	1.40 (0.78-2.49)	12	1.30 (0.61-2.75)
≥56 (56)	2619	29	1.60 (1.00-2.57)	17	1.77 (0.97-3.23)	9	1.20 (0.56-2.60)	7	1.51 (0.62-3.70)
p-trend			0.036		0.117		0.336		0.330
per 2 sizes	28428	234	1.23 (1.01-1.51)	131	1.29 (1.00-1.67)	89	1.18 (0.87-1.62)	60	1.31 (0.89-1.93)

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Clothing size: range (median)									
Men – rectum									
≤50 (50)	10903	21	1.00 (ref.)						
52 (52)	9750	17	0.87 (0.45-1.68)						
54 (54)	5156	17	1.65 (0.86-3.18)						
≥56 (56)	2619	9	1.74 (0.78-3.86)						
p-trend			0.077						
per 2 sizes	28428	64	1.17 (0.80-1.69)						
Women – colon									
≤40 (40)	6574	35	1.00 (ref.)	24	1.00 (ref.)	29	1.00 (ref.)	32	1.00 (ref.)
42 (42)	8582	54	1.16 (0.74-1.81)	25	0.78 (0.44-1.39)	47	1.20 (0.74-1.96)	35	0.82 (0.49-1.35)
44 (44)	9270	58	1.11 (0.71-1.74)	33	0.93 (0.54-1.59)	45	1.02 (0.62-1.66)	25	0.52 (0.30-0.89)
≥46 (46)	9454	70	1.30 (0.84-2.01)	30	0.81 (0.47-1.42)	49	1.07 (0.66-1.73)	37	0.74 (0.45-1.22)
p-trend			0.275		0.650		0.935		0.146
per 2 sizes	33880	217	1.24 (1.02-1.50)	112	1.10 (0.86-1.41)	170	1.03 (0.85-1.26)	129	0.90 (0.71-1.13)
Women – rectum									
≤40 (40)	6574	10	1.00 (ref.)						
42 (42)	8582	9	0.68 (0.28-1.70)						
44 (44)	9270	12	0.83 (0.35-1.97)						
≥46 (46)	9454	23	1.55 (0.72-3.35)						
p-trend			0.153						
per 2 sizes	33880	54	1.14 (0.83-1.58)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; dMMR, mismatch repair deficient; NLCS, Netherlands Cohort Study; mut, mutated; BMI, body mass index.

[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

[b]Analyses for subgroups with <50 cases were not performed.

Supplementary Table S5 | Age-adjusted HRs[a] and 95% CIs for associations between physical activity measures and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Non-occupational physical activity (min/day): range (median)							
Men – colon							
≤ 30 (21.4)	4997	132	1.00 (ref.)	49	1.00 (ref.)	83	1.00 (ref.)
31-60 (42.9)	10100	242	0.87 (0.67-1.13)	94	0.91 (0.63-1.32)	148	0.85 (0.63-1.16)
61-90 (73.6)	6001	156	0.96 (0.73-1.28)	62	1.02 (0.68-1.53)	94	0.93 (0.66-1.30)
> 90 (130.0)	9925	224	0.80 (0.62-1.04)	104	1.01 (0.70-1.47)	120	0.68 (0.49-0.93)
p-trend			0.169		0.676		0.025
per 30 min/day	31022	754	0.98 (0.94-1.02)	309	1.01 (0.96-1.06)	445	0.96 (0.91-1.01)
Men – rectum							
≤ 30 (21.4)	4997	18	1.00 (ref.)	13	1.00 (ref.)	5	1.00 (ref.)
31-60 (42.9)	10100	73	1.95 (1.14-3.33)	41	1.51 (0.80-2.86)	32	3.09 (1.19-8.04)
61-90 (73.6)	6001	57	2.56 (1.47-4.46)	38	2.34 (1.22-4.49)	19	3.12 (1.15-8.48)
> 90 (130.0)	9925	76	2.08 (1.22-3.55)	43	1.65 (0.87-3.13)	33	3.20 (1.23-8.32)
p-trend			0.015		0.097		0.051
per 30 min/day	31022	224	1.03 (0.98-1.09)	135	1.04 (0.97-1.12)	89	1.02 (0.95-1.10)
Women – colon							
≤ 30 (19.3)	7756	169	1.00 (ref.)	52	1.00 (ref.)	117	1.00 (ref.)
31-60 (42.9)	10923	198	0.84 (0.66-1.07)	44	0.59 (0.39-0.91)	154	0.95 (0.72-1.25)
61-90 (75.0)	8000	148	0.86 (0.66-1.12)	47	0.86 (0.57-1.31)	101	0.86 (0.64-1.17)
> 90 (115.7)	7550	115	0.71 (0.54-0.94)	36	0.70 (0.45-1.09)	79	0.72 (0.52-0.99)
p-trend			0.028		0.352		0.033
per 30 min/day	34228	630	0.97 (0.91-1.03)	179	0.98 (0.88-1.09)	451	0.96 (0.90-1.03)
Women – rectum							
≤ 30 (19.3)	7756	31	1.00 (ref.)	14	1.00 (ref.)	17	1.00 (ref.)
31-60 (42.9)	10923	44	1.01 (0.63-1.62)	26	1.29 (0.67-2.51)	18	0.76 (0.39-1.49)
61-90 (75.0)	8000	34	1.06 (0.64-1.76)	12	0.81 (0.36-1.78)	22	1.29 (0.68-2.45)
> 90 (115.7)	7550	22	0.73 (0.42-1.28)	12	0.85 (0.39-1.88)	10	0.63 (0.28-1.39)
p-trend			0.341		0.379		0.625
per 30 min/day	34228	131	1.01 (0.89-1.15)	64	1.08 (0.91-1.28)	67	0.93 (0.80-1.09)
Occupational energy expenditure (kJ/min)							
Men - colon							
< 8	15144	365	1.00 (ref.)	152	1.00 (ref.)	213	1.00 (ref.)
8-12	6368	133	0.86 (0.68-1.09)	54	0.85 (0.61-1.18)	79	0.87 (0.65-1.16)
> 12	3561	66	0.75 (0.55-1.01)	20	0.55 (0.34-0.90)	46	0.89 (0.62-1.26)
p-trend	25073	564	0.038	226	0.014	338	0.361

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Occupational energy expenditure (kJ/min)							
Men – rectum							
< 8	15144	107	1.00 (ref.)	65	1.00 (ref.)	42	1.00 (ref.)
8-12	6368	57	1.28 (0.91-1.80)	35	1.31 (0.85-2.00)	22	1.25 (0.74-2.11)
> 12	3561	21	0.83 (0.51-1.36)	14	0.92 (0.51-1.68)	7	0.69 (0.31-1.55)
p-trend	25073	185	0.962	114	0.788	71	0.659
Occupational sitting time (hrs/day)							
Men – colon							
> 6	6511	187	1.00 (ref.)	85	1.00 (ref.)	102	1.00 (ref.)
2-6	11617	244	0.71 (0.57-0.89)	87	0.56 (0.41-0.77)	157	0.83 (0.63-1.10)
< 2	6944	133	0.66 (0.51-0.85)	54	0.59 (0.41-0.85)	79	0.71 (0.51-0.99)
p-trend	25073	564	0.002	226	0.005	338	0.041
Men – rectum							
> 6	6511	60	1.00 (ref.)	39	1.00 (ref.)	21	1.00 (ref.)
2-6	11617	69	0.64 (0.44-0.92)	40	0.57 (0.36-0.90)	29	0.75 (0.42-1.33)
< 2	6944	56	0.88 (0.59-1.29)	35	0.85 (0.53-1.37)	21	0.93 (0.50-1.71)
p-trend	25073	185	0.526	114	0.518	71	0.825

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study.

[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR deficient.

Supplementary Table S6 | Age-adjusted HRs[a] and 95% CIs for associations between physical activity measures and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Non-occupational physical activity (min/day): range (median)									
Men – colon									
≤ 30 (21.4)	4997	41	1.00 (ref.)	24	1.00 (ref.)	19	1.00 (ref.)	13	1.00 (ref.)
31-60 (42.9)	10100	83	0.97 (0.65-1.45)	63	1.26 (0.77-2.07)	34	0.85 (0.47-1.52)	18	0.65 (0.31-1.36)
61-90 (73.6)	6001	63	1.27 (0.83-1.94)	28	0.96 (0.55-1.70)	19	0.82 (0.43-1.58)	16	1.02 (0.48-2.15)
> 90 (130.0)	9925	69	0.79 (0.52-1.20)	35	0.70 (0.41-1.21)	33	0.80 (0.44-1.45)	23	0.79 (0.39-1.61)
p-trend			0.324		0.029		0.515		0.934
per 30 min/day	31022	256	0.94 (0.88-1.00)	150	0.96 (0.87-1.06)	105	1.02 (0.93-1.11)	70	1.02 (0.92-1.14)
Men – rectum									
≤ 30 (21.4)	4997	5	1.00 (ref.)						
31-60 (42.9)	10100	26	2.50 (0.95-6.61)						
61-90 (73.6)	6001	16	2.64 (0.95-7.32)						
> 90 (130.0)	9925	21	2.00 (0.74-5.38)						
p-trend			0.464						
per 30 min/day	31022	68	0.98 (0.88-1.08)						
Women – colon									
≤ 30 (19.3)	7756	55	1.00 (ref.)	36	1.00 (ref.)	47	1.00 (ref.)	32	1.00 (ref.)
31-60 (42.9)	10923	78	1.02 (0.70-1.48)	35	0.70 (0.43-1.14)	59	0.91 (0.61-1.36)	45	1.02 (0.63-1.63)
61-90 (75.0)	8000	48	0.87 (0.57-1.31)	28	0.78 (0.46-1.30)	35	0.75 (0.47-1.19)	34	1.05 (0.64-1.74)
> 90 (115.7)	7550	41	0.79 (0.51-1.21)	17	0.50 (0.28-0.91)	32	0.73 (0.45-1.18)	20	0.66 (0.37-1.18)
p-trend			0.192		0.040		0.131		0.207
per 30 min/day	34228	222	0.98 (0.89-1.08)	116	0.89 (0.76-1.05)	173	0.95 (0.85-1.05)	131	0.98 (0.87-1.10)
Women – rectum									
≤ 30 (19.3)	7756	14	1.00 (ref.)						
31-60 (42.9)	10923	15	0.77 (0.37-1.60)						
61-90 (75.0)	8000	20	1.40 (0.71-2.79)						
> 90 (115.7)	7550	6	0.45 (0.17-1.18)						
p-trend			0.388						
per 30 min/day	34228	55	0.86 (0.74-1.00)						
Occupational energy expenditure (kJ/min)									
Men - colon									
< 8	15144	115	1.00 (ref.)	72	1.00 (ref.)	59	1.00 (ref.)	32	1.00 (ref.)
8-12	6368	50	1.02 (0.72-1.45)	26	0.85 (0.54-1.35)	18	0.71 (0.41-1.23)	16	1.14 (0.61-2.11)
> 12	3561	25	0.89 (0.56-1.41)	16	0.92 (0.52-1.62)	10	0.69 (0.35-1.37)	10	1.26 (0.60-2.63)
p-trend	25073	190	0.709	114	0.625	87	0.177	58	0.509

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Occupational energy expenditure (kJ/min)									
Men – rectum									
< 8	15144	31	1.00 (ref.)						
8-12	6368	17	1.30 (0.71-2.36)						
> 12	3561	5	0.66 (0.26-1.70)						
p-trend	25073	53	0.686						
Occupational sitting time (hrs/day)									
Men – colon									
> 6	6511	57	1.00 (ref.)	34	1.00 (ref.)	22	1.00 (ref.)	14	1.00 (ref.)
2-6	11617	87	0.83 (0.58-1.18)	48	0.77 (0.49-1.21)	47	1.15 (0.68-1.94)	27	1.02 (0.52-1.98)
< 2	6944	46	0.74 (0.49-1.12)	32	0.87 (0.53-1.44)	18	0.75 (0.40-1.42)	17	1.09 (0.52-2.25)
p-trend	25073	190	0.156	114	0.602	87	0.352	58	0.819
Men – rectum									
> 6	6511	17	1.00 (ref.)						
2-6	11617	20	0.63 (0.33-1.22)						
< 2	6944	16	0.86 (0.43-1.72)						
p-trend	25073	53	0.700						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; dMMR, mismatch repair deficient; NLCS, Netherlands Cohort Study; mut, mutated.

[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

[b]Analyses for subgroups with <50 cases were not performed.

Supplementary Table S7 | Multivariable-adjusted HRs[a] and 95% CIs for associations between adiposity measures and CRC subgroups based on individual mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	*KRAS*$_{wt}$		*PIK3CA*$_{wt}$[b]		*BRAF*$_{wt}$[b]		pMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
BMI quartiles (kg/m²): range (median)									
Men – colon									
< 23.4 (22.2)	7993	103	1.00 (ref.)	140	1.00 (ref.)	142	1.00 (ref.)	148	1.00 (ref.)
23.4-24.9 (24.2)	8343	130	1.16 (0.87-1.57)	150	1.00 (0.76-1.32)	160	1.06 (0.81-1.38)	175	1.11 (0.85-1.44)
25.0-26.6 (25.7)	7683	131	1.26 (0.93-1.71)	158	1.12 (0.85-1.48)	171	1.21 (0.92-1.59)	179	1.21 (0.93-1.59)
> 26.6 (27.8)	7003	131	1.47 (1.09-1.99)	156	1.30 (0.99-1.72)	176	1.46 (1.11-1.91)	182	1.44 (1.10-1.89)
p-trend			0.011		0.044		0.004		0.006
per kg/m	31022	498	1.33 (1.11-1.58)	604	1.28 (1.09-1.51)	649	1.32 (1.12-1.54)	684	1.28 (1.10-1.50)
Men – rectum									
< 23.4 (22.2)	7993	40	1.00 (ref.)						
23.4-24.9 (24.2)	8343	41	0.94 (0.59-1.49)						
25.0-26.6 (25.7)	7683	46	1.18 (0.75-1.85)						
> 26.6 (27.8)	7003	29	0.81 (0.49-1.36)						
p-trend			0.706						
per kg/m	31022	156	0.96 (0.73-1.26)						
Women – colon									
<22.8 (21.5)	9014	129	1.00 (ref.)	151	1.00 (ref.)	128	1.00 (ref.)	138	1.00 (ref.)
22.8-24.7 (23.8)	8914	100	0.78 (0.58-1.05)	119	0.80 (0.60-1.05)	106	0.83 (0.62-1.11)	112	0.82 (0.62-1.09)
24.8-27.0 (25.7)	8141	82	0.71 (0.52-0.98)	117	0.89 (0.67-1.19)	108	0.97 (0.72-1.30)	123	1.03 (0.77-1.37)
>27.0 (29.2)	8158	97	0.86 (0.63-1.18)	127	1.01 (0.76-1.34)	115	1.05 (0.78-1.42)	126	1.08 (0.81-1.44)
p-trend			0.255		0.856		0.583		0.393
per 5 kg/m²	34228	408	0.90 (0.77-1.05)	514	1.03 (0.90-1.18)	457	1.06 (0.91-1.22)	499	1.08 (0.94-1.24)
Women – rectum									
<22.8 (21.5)	9014	19	1.00 (ref.)						
22.8-24.7 (23.8)	8914	19	0.96 (0.49-1.87)						
24.8-27.0 (25.7)	8141	23	1.25 (0.66-2.37)						
>27.0 (29.2)	8158	15	0.78 (0.39-1.58)						
p-trend			0.714						
per 5 kg/m²	34228	76	0.99 (0.74-1.33)						
Clothing size: range (median)									
Men - colon									
≤50 (50)	10903	138	1.00 (ref.)	171	1.00 (ref.)	181	1.00 (ref.)	193	1.00 (ref.)
52 (52)	9750	163	1.31 (1.01-1.71)	195	1.26 (0.99-1.61)	218	1.34 (1.06-1.69)	224	1.29 (1.02-1.63)
54 (54)	5156	88	1.36 (1.00-1.86)	114	1.39 (1.04-1.85)	115	1.35 (1.01-1.79)	124	1.36 (1.03-1.79)
≥56 (56)	2619	61	1.89 (1.32-2.70)	73	1.79 (1.28-2.52)	81	1.90 (1.37-2.64)	83	1.83 (1.32-2.53)
p-trend			0.001		<0.001		<0.001		<0.001
per 2 sizes	28428	450	1.37 (1.18-1.60)	553	1.33 (1.15-1.54)	595	1.35 (1.17-1.55)	624	1.33 (1.16-1.52)

(continued)	Person years at risk	*KRAS*$_{wt}$		*PIK3CA*$_{wt}$[b]		*BRAF*$_{wt}$[b]		pMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Clothing size: range (median)									
Men – rectum									
≤50 (50)	10903	57	1.00 (ref.)						
52 (52)	9750	52	1.03 (0.69-1.53)						
54 (54)	5156	26	0.99 (0.61-1.62)						
≥56 (56)	2619	7	0.55 (0.25-1.23)						
p-trend			0.293						
per 2 sizes	28428	142	0.91 (0.73-1.13)						
Women – colon									
≤40 (40)	6574	93	1.00 (ref.)	104	1.00 (ref.)	99	1.00 (ref.)	96	1.00 (ref.)
42 (42)	8582	96	0.79 (0.57-1.10)	125	0.92 (0.68-1.24)	103	0.78 (0.57-1.07)	115	0.90 (0.66-1.23)
44 (44)	9270	101	0.74 (0.53-1.01)	126	0.83 (0.62-1.12)	114	0.78 (0.57-1.06)	134	0.94 (0.70-1.28)
≥46 (46)	9454	112	0.81 (0.59-1.12)	152	1.01 (0.75-1.36)	133	0.91 (0.67-1.23)	145	1.02 (0.76-1.38)
p-trend			0.230		0.994		0.701		0.747
per 2 sizes	33880	402	1.00 (0.85-1.18)	507	1.09 (0.94-1.27)	449	1.10 (0.94-1.29)	490	1.14 (0.98-1.32)
Women – rectum									
≤40 (40)	6574	13	1.00 (ref.)						
42 (42)	8582	21	1.16 (0.56-2.40)						
44 (44)	9270	23	1.23 (0.61-2.45)						
≥46 (46)	9454	19	0.94 (0.45-1.95)						
p-trend			0.835						
per 2 sizes	33880	76	0.94 (0.72-1.22)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; wt, wild-type.

[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate. BMI models were additionally adjusted for height (cm; continuous).

[b]Analyses were not performed when <50 cases showed a mutation or dMMR.

Supplementary Table S8 | Multivariable-adjusted HRs[a] and 95% CIs for associations between physical activity measures and CRC subgroups based on individual mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	*KRAS*$_{wt}$		*PIK3CA*$_{wt}$[b]		*BRAF*$_{wt}$[b]		pMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Non-occupational physical activity (min/day): range (median)									
Men – colon									
≤ 30 (21.4)	4997	91	1.00 (ref.)	108	1.00 (ref.)	113	1.00 (ref.)	119	1.00 (ref.)
31-60 (42.9)	10100	159	0.86 (0.63-1.16)	179	0.80 (0.60-1.07)	208	0.90 (0.68-1.18)	224	0.92 (0.70-1.20)
61-90 (73.6)	6001	93	0.85 (0.61-1.19)	128	1.00 (0.74-1.36)	137	1.02 (0.76-1.38)	140	1.00 (0.74-1.34)
> 90 (130.0)	9925	155	0.85 (0.63-1.15)	189	0.87 (0.66-1.16)	191	0.85 (0.64-1.13)	201	0.86 (0.65-1.13)
p-trend			0.414		0.838		0.422		0.346
per 30 min/day	31022	498	1.01 (0.96-1.06)	604	1.00 (0.95-1.04)	649	0.98 (0.94-1.03)	684	0.99 (0.94-1.03)
Men – rectum									
≤ 30 (21.4)	4997	13	1.00 (ref.)						
31-60 (42.9)	10100	47	1.72 (0.91-3.26)						
61-90 (73.6)	6001	41	2.52 (1.32-4.81)						
> 90 (130.0)	9925	55	2.12 (1.13-3.97)						
p-trend			0.012						
per 30 min/day	31022	156	1.06 (0.99-1.12)						
Women – colon									
≤ 30 (19.3)	7756	114	1.00 (ref.)	133	1.00 (ref.)	122	1.00 (ref.)	137	1.00 (ref.)
31-60 (42.9)	10923	120	0.74 (0.56-1.00)	163	0.86 (0.66-1.12)	139	0.80 (0.61-1.06)	153	0.79 (0.60-1.04)
61-90 (75.0)	8000	100	0.84 (0.62-1.14)	120	0.85 (0.64-1.14)	113	0.88 (0.66-1.18)	114	0.80 (0.60-1.07)
> 90 (115.7)	7550	74	0.66 (0.47-0.91)	98	0.75 (0.56-1.01)	83	0.70 (0.51-0.96)	95	0.72 (0.53-0.97)
p-trend			0.036		0.078		0.063		0.044
per 30 min/day	34228	408	0.96 (0.89-1.03)	514	0.98 (0.92-1.05)	457	0.98 (0.91-1.05)	499	0.97 (0.90-1.03)
Women – rectum									
≤ 30 (19.3)	7756	17	1.00 (ref.)						
31-60 (42.9)	10923	29	1.20 (0.65-2.24)						
61-90 (75.0)	8000	14	0.78 (0.38-1.60)						
> 90 (115.7)	7550	16	0.92 (0.46-1.86)						
p-trend			0.495						
per 30 min/day	34228	76	1.08 (0.93-1.26)						
Occupational energy expenditure (kJ/min)									
Men - colon									
< 8	15144	250	1.00 (ref.)	293	1.00 (ref.)	306	1.00 (ref.)	333	1.00 (ref.)
8-12	6368	83	0.74 (0.56-0.98)	107	0.83 (0.64-1.08)	115	0.86 (0.67-1.11)	117	0.80 (0.63-1.03)
> 12	3561	41	0.61 (0.42-0.89)	50	0.67 (0.47-0.95)	56	0.72 (0.51-1.01)	56	0.67 (0.48-0.94)
p-trend	25073		0.003	450	0.016	477	0.038	506	0.008

(continued)	Person years at risk	*KRAS*$_{wt}$		*PIK3CA*$_{wt}$[b]		*BRAF*$_{wt}$[b]		pMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Occupational energy expenditure (kJ/min)									
Men – rectum									
< 8	15144	76	1.00 (ref.)						
8-12	6368	40	1.31 (0.87-1.98)						
> 12	3561	16	0.89 (0.50-1.58)						
p-trend	25073	132	0.849						
Occupational sitting time (hrs/day)									
Men – colon									
> 6	6511	130	1.00 (ref.)	153	1.00 (ref.)	165	1.00 (ref.)	173	1.00 (ref.)
2-6	11617	157	0.65 (0.50-0.85)	196	0.68 (0.53-0.88)	197	0.64 (0.50-0.82)	217	0.67 (0.53-0.86)
< 2	6944	87	0.58 (0.43-0.80)	101	0.59 (0.44-0.79)	115	0.62 (0.47-0.82)	116	0.60 (0.45-0.80)
p-trend	25073	374	0.001	450	<0.001	477	0.001	506	<0.001
Men – rectum									
> 6	6511	43	1.00 (ref.)						
2-6	11617	49	0.62 (0.40-0.96)						
< 2	6944	40	0.87 (0.55-1.38)						
p-trend	25073	132	0.574						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; wt, wild-type.

[a]Hazard Ratios were adjusted for age (years; continuous), BMI (kg/m2), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Analyses were not performed when <50 cases showed a mutation or dMMR.

Supplementary Table S9 | Multivariable-adjusted HRs[a] and 95% CIs for associations between (mutually adjusted) adiposity measures and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986–2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
BMI quartiles (kg/m²): range (median)							
Men – colon							
< 23.4 (22.2)	7356	153	1.00 (ref.)	61	1.00 (ref.)	92	1.00 (ref.)
23.4-24.9 (24.2)	7756	176	0.97 (0.73-1.27)	73	0.95 (0.64-1.42)	103	0.98 (0.70-1.37)
25.0-26.6 (25.7)	7056	184	1.01 (0.75-1.37)	72	0.92 (0.61-1.41)	112	1.08 (0.75-1.56)
> 26.6 (27.8)	6260	171	1.02 (0.72-1.43)	80	1.06 (0.67-1.68)	91	0.98 (0.65-1.50)
p-trend			0.836		0.842		0.904
per kg/m	28428	684	1.16 (0.95-1.40)	286	1.20 (0.91-1.58)	398	1.12 (0.89-1.42)
Men – rectum							
< 23.4 (22.2)	7356	54	1.00 (ref.)	30	1.00 (ref.)	24	1.00 (ref.)
23.4-24.9 (24.2)	7756	51	0.82 (0.54-1.25)	33	0.98 (0.58-1.66)	18	0.64 (0.34-1.21)
25.0-26.6 (25.7)	7056	62	1.11 (0.72-1.70)	40	1.35 (0.78-2.34)	22	0.82 (0.43-1.55)
> 26.6 (27.8)	6260	39	0.78 (0.46-1.31)	21	0.79 (0.38-1.64)	18	0.74 (0.38-1.45)
p-trend			0.693		0.932		0.570
per kg/m	28428	206	0.97 (0.73-1.30)	124	0.95 (0.65-1.39)	82	1.01 (0.67-1.52)
Women – colon							
<22.8 (21.5)	8964	179	1.00 (ref.)	56	1.00 (ref.)	123	1.00 (ref.)
22.8-24.7 (23.8)	8874	145	0.78 (0.58-1.05)	43	0.73 (0.43-1.23)	102	0.80 (0.57-1.12)
24.8-27.0 (25.7)	8101	143	0.84 (0.58-1.22)	35	0.63 (0.29-1.39)	108	0.94 (0.63-1.39)
>27.0 (29.2)	7941	152	0.89 (0.52-1.51)	44	0.75 (0.22-2.54)	108	0.95 (0.57-1.59)
p-trend			0.710		0.563		0.987
per 5 kg/m²	33880	619	1.00 (0.78-1.28)	178	0.70 (0.40-1.23)	441	1.15 (0.90-1.46)
Women – rectum							
<22.8 (21.5)	8964	37	1.00 (ref.)	16	1.00 (ref.)	21	1.00 (ref.)
22.8-24.7 (23.8)	8874	26	0.80 (0.44-1.45)	15	1.00 (0.41-2.40)	11	0.65 (0.29-1.41)
24.8-27.0 (25.7)	8101	32	1.18 (0.62-2.24)	18	1.37 (0.55-3.46)	14	1.03 (0.43-2.49)
>27.0 (29.2)	7941	35	1.58 (0.67-3.74)	15	1.25 (0.34-4.55)	20	2.01 (0.66-6.16)
p-trend			0.230		0.562		0.243
per 5 kg/m²	33880	130	1.40 (0.95-2.05)	64	1.31 (0.75-2.28)	66	1.49 (0.89-2.48)
Clothing size: range (median)							
Men - colon							
≤50 (50)	10903	211	1.00 (ref.)	90	1.00 (ref.)	121	1.00 (ref.)
52 (52)	9750	247	1.10 (0.86-1.41)	104	1.17 (0.82-1.65)	143	1.06 (0.79-1.43)
54 (54)	5156	136	1.05 (0.76-1.44)	53	1.07 (0.68-1.67)	83	1.03 (0.71-1.51)
≥56 (56)	2619	90	1.14 (0.74-1.76)	39	1.37 (0.75-2.52)	51	1.00 (0.60-1.67)
p-trend			0.648		0.453		0.985
per 2 sizes	28428	684	1.13 (0.95-1.34)	286	1.22 (0.96-1.54)	398	1.06 (0.86-1.31)

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Clothing size: range (median)							
Men – rectum							
≤50 (50)	10903	78	1.00 (ref.)	46	1.00 (ref.)	32	1.00 (ref.)
52 (52)	9750	69	0.99 (0.68-1.45)	46	1.07 (0.66-1.74)	23	0.86 (0.49-1.51)
54 (54)	5156	43	1.20 (0.74-1.95)	25	1.08 (0.59-2.00)	18	1.39 (0.69-2.82)
≥56 (56)	2619	16	0.90 (0.45-1.82)	7	0.59 (0.22-1.61)	9	1.47 (0.60-3.59)
p-trend			0.782		0.635		0.274
per 2 sizes	28428	206	0.95 (0.74-1.22)	124	0.91 (0.66-1.26)	82	1.01 (0.73-1.41)
Women – colon							
≤40 (40)	6574	128	1.00 (ref.)	46	1.00 (ref.)	82	1.00 (ref.)
42 (42)	8582	150	0.82 (0.60-1.11)	34	0.62 (0.37-1.03)	116	0.92 (0.65-1.31)
44 (44)	9270	159	0.74 (0.52-1.05)	48	0.85 (0.48-1.52)	111	0.71 (0.47-1.07)
≥46 (46)	9454	182	0.79 (0.50-1.25)	50	1.02 (0.47-2.18)	132	0.74 (0.44-1.24)
p-trend			0.270		0.778		0.136
per 2 sizes	33880	619	1.19 (0.91-1.57)	178	1.62 (0.96-2.73)	441	1.05 (0.82-1.35)
Women – rectum							
≤40 (40)	6574	23	1.00 (ref.)	11	1.00 (ref.)	12	1.00 (ref.)
42 (42)	8582	30	0.89 (0.49-1.61)	17	1.12 (0.49-2.52)	13	0.69 (0.30-1.61)
44 (44)	9270	35	0.94 (0.50-1.77)	20	1.23 (0.52-2.91)	15	0.69 (0.28-1.71)
≥46 (46)	9454	42	1.06 (0.46-2.47)	16	0.89 (0.24-3.33)	26	1.19 (0.40-3.52)
p-trend			0.831		0.943		0.685
per 2 sizes	33880	130	0.80 (0.54-1.17)	64	0.89 (0.52-1.53)	66	0.72 (0.42-1.21)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; BMI, body mass index.

[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate. BMI models were additionally adjusted for height (cm; continuous). Mutual adjustment: clothing size (size; continuous) was included in BMI models, and BMI (kg/m2; continuous) was included in clothing size models.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR deficient.

Supplementary Table S10 | Multivariable-adjusted HRs[a] and 95% CIs for associations between (mutually adjusted) adiposity measures and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986–2006.

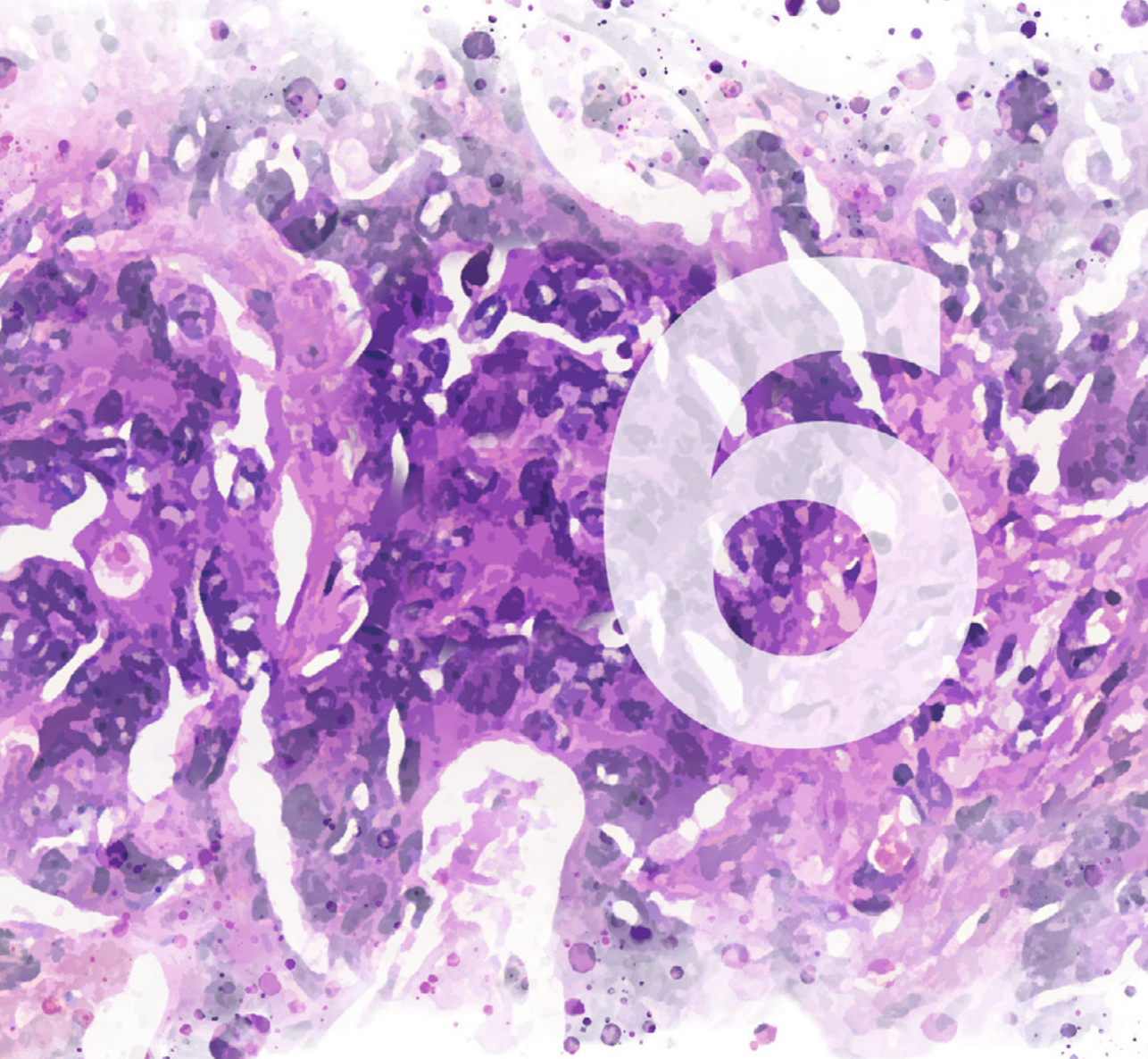
	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
BMI quartiles (kg/m²): range (median)									
Men – colon									
< 23.4 (22.2)	7993	56	1.00 (ref.)	25	1.00 (ref.)	22	1.00 (ref.)	18	1.00 (ref.)
23.4-24.9 (24.2)	8343	56	0.89 (0.58-1.36)	35	1.30 (0.74-2.30)	27	1.12 (0.60-2.09)	13	0.62 (0.29-1.31)
25.0-26.6 (25.7)	7683	62	1.00 (0.64-1.58)	38	1.52 (0.84-2.78)	26	1.09 (0.52-2.29)	17	0.80 (0.36-1.80)
> 26.6 (27.8)	7003	60	1.13 (0.68-1.87)	33	1.49 (0.77-2.92)	14	0.68 (0.27-1.72)	12	0.61 (0.24-1.55)
p-trend			0.539		0.200		0.505		0.457
per kg/m	31022	234	1.19 (0.91-1.57)	131	1.26 (0.88-1.82)	89	0.99 (0.60-1.63)	60	1.05 (0.61-1.83)
Men – rectum									
< 23.4 (22.2)	7993	18	1.00 (ref.)						
23.4-24.9 (24.2)	8343	13	0.57 (0.27-1.20)						
25.0-26.6 (25.7)	7683	18	0.78 (0.38-1.60)						
> 26.6 (27.8)	7003	15	0.69 (0.32-1.49)						
p-trend			0.563						
per kg/m	31022	64	0.91 (0.58-1.43)						
Women – colon									
<22.8 (21.5)	9014	50	1.00 (ref.)	29	1.00 (ref.)	53	1.00 (ref.)	43	1.00 (ref.)
22.8-24.7 (23.8)	8914	46	0.86 (0.53-1.40)	27	0.99 (0.54-1.79)	39	0.68 (0.42-1.10)	33	0.76 (0.45-1.30)
24.8-27.0 (25.7)	8141	64	1.35 (0.78-2.32)	30	1.24 (0.62-2.45)	37	0.66 (0.36-1.18)	23	0.59 (0.30-1.14)
>27.0 (29.2)	8158	57	1.17 (0.57-2.41)	26	1.16 (0.50-2.72)	41	0.69 (0.32-1.48)	30	0.83 (0.36-1.88)
p-trend			0.351		0.588		0.330		0.437
per 5 kg/m²	34228	217	1.43 (1.04-1.96)	112	1.22 (0.79-1.89)	170	0.93 (0.65-1.35)	129	0.99 (0.64-1.53)
Women – rectum									
<22.8 (21.5)	9014	18	1.00 (ref.)						
22.8-24.7 (23.8)	8914	7	0.51 (0.20-1.29)						
24.8-27.0 (25.7)	8141	9	0.85 (0.32-2.30)						
>27.0 (29.2)	8158	20	2.81 (0.87-9.07)						
p-trend			0.131						
per 5 kg/m²	34228	54	1.63 (0.95-2.77)						
Clothing size: range (median)									
Men - colon									
≤50 (50)	10903	73	1.00 (ref.)	40	1.00 (ref.)	30	1.00 (ref.)	18	1.00 (ref.)
52 (52)	9750	84	1.03 (0.71-1.48)	52	1.21 (0.76-1.91)	29	0.78 (0.42-1.42)	23	1.18 (0.56-2.46)
54 (54)	5156	48	0.96 (0.61-1.53)	22	0.89 (0.48-1.63)	21	0.87 (0.42-1.77)	12	1.06 (0.43-2.60)
≥56 (56)	2619	29	0.90 (0.48-1.70)	17	1.08 (0.51-2.28)	9	0.54 (0.20-1.46)	7	1.03 (0.31-3.42)
p-trend			0.739		0.848		0.364		0.976
per 2 sizes	28428	234	0.99 (0.77-1.26)	131	1.09 (0.79-1.51)	89	0.90 (0.59-1.36)	60	1.22 (0.72-2.06)

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Clothing size: range (median)									
Men – rectum									
≤50 (50)	10903	21	1.00 (ref.)						
52 (52)	9750	17	0.91 (0.46-1.79)						
54 (54)	5156	17	1.76 (0.80-3.91)						
≥56 (56)	2619	9	1.83 (0.71-4.75)						
p-trend			0.105						
per 2 sizes	28428	64	1.01 (0.70-1.47)						
Women – colon									
≤40 (40)	6574	35	1.00 (ref.)	24	1.00 (ref.)	29	1.00 (ref.)	32	1.00 (ref.)
42 (42)	8582	54	0.89 (0.55-1.44)	25	0.62 (0.33-1.18)	47	1.27 (0.73-2.21)	35	0.73 (0.40-1.32)
44 (44)	9270	58	0.73 (0.43-1.24)	33	0.62 (0.31-1.23)	45	1.06 (0.55-2.05)	25	0.42 (0.20-0.89)
≥46 (46)	9454	70	0.65 (0.33-1.28)	30	0.44 (0.19-1.06)	49	1.24 (0.55-2.79)	37	0.57 (0.23-1.42)
p-trend			0.165		0.111		0.838		0.109
per 2 sizes	33880	217	1.07 (0.76-1.49)	112	1.12 (0.72-1.75)	170	1.15 (0.78-1.68)	129	0.85 (0.54-1.34)
Women – rectum									
≤40 (40)	6574	10	1.00 (ref.)						
42 (42)	8582	9	0.55 (0.21-1.43)						
44 (44)	9270	12	0.61 (0.22-1.63)						
≥46 (46)	9454	23	1.08 (0.35-3.36)						
p-trend			0.722						
per 2 sizes	33880	54	0.66 (0.37-1.17)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; mut, mutated; BMI, body mass index.

[a]Hazard Ratios were adjusted for age (years; continuous), non-occupational physical activity (minutes/day; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate. BMI models were additionally adjusted for height (cm; continuous). Mutual adjustment: clothing size (size; continuous) was included in BMI models, and BMI (kg/m2; continuous) was included in clothing size models.

[b]Analyses for subgroups with <50 cases were not performed.



6

ENERGY BALANCE-RELATED FACTORS IN CHILDHOOD AND ADOLESCENCE AND RISK OF COLORECTAL CANCER BASED ON *KRAS*, *PIK3CA* AND *BRAF* MUTATIONS AND MMR STATUS

Josien C. A. Jenniskens, Kelly Offermans, Colinda C. J. M. Simons, Iryna Samarska, Gregorio E. Fazzi, Jaleesa R. M. van der Meer, Kim M. Smits, Leo J. Schouten, Matty P. Weijenberg, Heike I. Grabsch, Piet A. van den Brandt

Molecular Carcinogenesis, 2022.

ABSTRACT

INTRODUCTION

KRAS mutations ($KRAS_{mut}$), $PIK3CA_{mut}$, $BRAF_{mut}$, and deficient DNA mismatch repair (dMMR) have been associated with the Warburg-effect. We previously reported differential associations between early-life energy balance-related factors (height, energy restriction, BMI) and colorectal cancer (CRC) subtypes based on the Warburg-effect. We now investigated associations of early-life energy balance-related factors and risk of CRC subgroups based on mutation and MMR status.

METHODS

Data from the Netherlands Cohort Study was used. $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status were available for 2,349 CRC cases, and complete covariate data for 1,934 cases and 3,911 subcohort members. Multivariable-adjusted Cox-regression was used to estimate associations of height, energy restriction proxies (exposure to Dutch Hunger Winter, Second World War, Economic Depression), and early adult BMI (age 20 years) with risk of CRC based on individual molecular features and combinations thereof (all-wild-type+MMR-proficient [pMMR]; any-mutation/dMMR).

RESULTS

Height was positively associated with any-mutation/dMMR CRC but not all-wild-type+pMMR CRC, with the exception of rectal cancer in men, and with heterogeneity in associations observed for colon cancer in men (P-heterogeneity=0.049) and rectal cancer in women (P-heterogeneity=0.014). Results on early-life energy restriction proxies in relation to risk of CRC subgroups did not show clear patterns. Early adult BMI was positively, but not significantly, associated with $KRAS_{mut}$ colon cancer in men and with $BRAF_{mut}$ and dMMR colon cancer in women.

CONCLUSIONS

Our results suggest a role of $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR in the etiological pathway between height and CRC risk. $KRAS_{mut}$ might potentially play a role in associations of early adult BMI with colon cancer risk in men, and $BRAF_{mut}$ and dMMR in women.

INTRODUCTION

Energy balance-related factors in early-life, i.e. from childhood to adolescence, have been associated with colorectal cancer (CRC) risk later in life[1-9]. Whilst there are reports that early-life energy restriction decreases CRC risk[1-3], the opposite has been reported as well[8, 9]. Increased adult-attained height, a surrogate measure of fetal and early-life (nutritional) exposures, and increased childhood or early adult body mass index (BMI) have been associated with increased CRC risk[4, 6, 7]. Mechanisms underlying these long-term effects of early-life energy balance-related factors are currently unknown. We have previously reported differential associations of early-life energy balance-related factors with CRC risk based on expression levels of proteins involved in the Warburg-effect[10].

It has been suggested that the Warburg-effect, which is characterized by increased aerobic glycolysis[11-13], may play an important role in carcinogenesis [14, 15]. Mutations in oncogenes such as *KRAS*, *PIK3CA*, and *BRAF* as well as the presence of DNA mismatch repair (MMR) deficiency have been associated with the presence of the Warburg-effect[11, 16-19]. Whilst mutations in *KRAS*, *PIK3CA*, and *BRAF* ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, respectively) and MMR deficiency (dMMR), a surrogate of microsatellite instability (MSI), are common molecular characteristics of CRC[20-22], there are only few studies investigating associations between early-life energy balance-related factors and risk of CRC in relation to these molecular features[23-26].

We hypothesized that associations of early-life energy balance-related factors with CRC risk differ between subgroups based on molecular features ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, dMMR). We investigated whether early-life energy balance-related factors were associated with CRC risk in relation to these molecular features individually as well as combined into an all-wild-type+pMMR subgroup (i.e. cases wild-type for all genes and MMR proficient [pMMR]) and any-mutation/dMMR subgroup (i.e. cases with a mutation in any of the genes and/or dMMR). We believe that combining these molecular features into subgroups has some advantages. First, mutations in *KRAS*, *PIK3CA*, and *BRAF* have all been associated with metabolic reprogramming towards the Warburg-effect[11, 16-18], and MMR deficiency has previously been associated with estimated presence of the Warburg-effect[19]. Therefore, combining these molecular features, presumed to be involved in the same metabolic phenotype, results in a cleaner reference group compared to groups based on individual features (e.g. *KRAS*-mutated versus –wild-type). Second, using combination subgroups based on multiple molecular features results in increased statistical power, since most individual molecular features occur in <20% of CRC cases. The all-wild-type+pMMR subgroup was used as reference group for all other subgroups. Associations with subgroups of individual molecular features and/or with the any-mutation/dMMR subgroup, but not with the all-wild-type+pMMR subgroup, were considered to indicate potential Warburg-effect involvement in the etiological pathway between the energy balance-related factor and CRC.

MATERIALS AND METHODS

DESIGN AND STUDY POPULATION

We used data from the large prospective Netherlands Cohort Study (NLCS), which included 120,852 subjects aged 55-69 years at baseline in 1986. All participants completed a mailed, self-administered baseline questionnaire on cancer risk factors[27]. By completing and returning the questionnaire, participants agreed to participate in the study. The NLCS was approved by institutional review boards from Maastricht University and the Netherlands Organization for Applied Scientific Research. Ethical approval was obtained from the Medical Ethical Committee of Maastricht University Medical Center+. The NLCS uses a case-cohort approach for data processing and analysis[28]. Immediately after baseline, 5,000 participants were randomly sampled from the full cohort, and accumulated person-years were estimated from this subcohort. Information on vital status of subcohort members was obtained by biennial active follow-up and by linkage with municipal population registries. Through annual record linkage with the Netherlands Cancer Registry and PALGA, the nationwide Dutch Pathology Registry[29], incident cancer cases from the full cohort were identified. Completeness of cancer follow-up by the Netherlands Cancer Registry and PALGA was estimated to be over 96%[30]. For the current study, follow-up covered 20.3 years (September 17, 1986 until January 1, 2007). After excluding cases and subcohort members who reported a history of cancer (except skin cancer) at baseline, a total of 4,597 incident CRC cases and 4,774 subcohort members were available (Supplementary Figure S1). As described previously[31], formalin-fixed paraffin-embedded (FFPE) tissue blocks from primary tumor and matched normal colon tissue from 3,872 CRC cases were requested from participating laboratories as part of the Rainbow-TMA project during 2012-2017[27]. Tissue blocks from 3,021 CRC cases were collected from 43 Dutch pathology laboratories (78% retrieval rate) (Supplementary Figure S1) and used to extract tumor DNA.

TISSUE MICROARRAY CONSTRUCTION AND IMMUNOHISTOCHEMISTRY

Three 0.6mm cores were sampled from FFPE blocks of 2,694 CRC cases and combined into 78 tissue microarray (TMA) blocks (Supplementary Figure S1), as described previously[19]. From these TMA blocks, 5µm thick sections were cut and stained with Hematoxylin & Eosin (H&E) according to a standard protocol or subjected to immunohistochemistry (IHC) using an automated immunostainer (DAKO Autostainer Link 48, Glostrup, Denmark). MMR status, as an indicator for the presence of absence of MSI, was assessed using IHC staining for MLH1 and MSH2 as described previously[19]. All TMA sections were scanned using an Aperio scanner (Leica Microsystems, Milton Keynes, UK) at 40x magnification at the University of Leeds (UK) Scanning Facility or at the Department of Pathology, Aachen University Hospital (Germany).

For quality control, H&E-stained TMA sections combined with pan-cytokeratin stained sections, if necessary, were reviewed to confirm presence of adenocarcinoma for each core. Requiring at least one core per case, 2,497 cases passed quality control (Supplementary Figure S1). IHC-stained TMAs for MLH1 and MSH2 were scored

according to the protocol published by Richman et al[32] by three non-pathologists (GF: senior histology technician; KO: PhD-student; JJ: PhD-student) after appropriate training[33], as well as by an experienced histopathologist (HG). Tumors with complete loss of either MLH1 or MSH2 expression were classified as MMR deficient (dMMR), and those expressing both MLH1 and MSH2 were classified as MMR proficient (pMMR). MMR status information was available for 2,455 CRC cases (Supplementary Figure S1).

DNA ISOLATION AND MUTATION DETECTION

Two 20 µm thick sections were cut from primary tumor containing FFPE blocks for DNA extraction. Sections were deparaffinized manually with Buffer ATL (Cat. No. 939011, Qiagen, Hilden, Germany), Proteinase K (Cat. No. 19131, Qiagen), and Deparaffinization Solution (Cat. No. 19093, Qiagen), using an adapted version of the manufacturer's protocol (Supplementary Methods). For DNA isolation, the QIAsymphony® DSP DNA Mini Kit (Cat. No. 937236, Qiagen) and the QIAsymphony® (Qiagen) instrument were used according to the manufacturer's protocol (Tissue_HC_200 protocol). The Quantus™ Fluorometer (Promega, Madison, WI, USA) with a QuantiFluor® dsDNA system (Promega) was used to determine double-stranded DNA concentrations. Mutations in tumor DNA were analyzed at the Institut für Immunologie und Genetik (Kaiserslautern, Germany). The ColoCarta panel (Agena Bioscience, Hamburg) was used to screen for more than 32 mutations in 6 genes (*BRAF*, *HRAS*, *KRAS*, *MET*, *NRAS*, *PIK3CA*; see Supplementary Table S1 for specific mutations), using Matrix Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) mass spectrometry (cut-offs: Z-score ≥ 4.00; spectrum quality ≥ 0.750; typer peak probability ≥ 0.850; primer extension rate cut-off ≥ 0.200). A mutation frequency of ≥7.5% for any of the alleles was considered evidence of a mutation in the corresponding gene. A failed reaction at a single nucleotide position resulted in missing data for the corresponding gene status only if the reactions at all other positions were wild-type. Information on *KRAS*, *PIK3CA*, and *BRAF* mutation status was complete for 2,349 CRC cases (Supplementary Figure S1).

The following subgroups were used for statistical analyses: I) all-wild-type+pMMR – cases wild-type for all genes (*KRAS*, *BRAF*, and *PIK3CA*) and MMR-proficient; II) any-mutation/dMMR – cases with a mutation in any of the genes (*KRAS*, *BRAF*, and *PIK3CA*) and/or MMR-deficient; III) $KRAS_{mut}$ – cases with a (non-exclusive) *KRAS* mutation; IV) $BRAF_{mut}$; V) $PIK3CA_{mut}$; VI) dMMR. Note: subgroups based on individual mutations and MMR status can overlap since multiple mutations and/or MMR deficiency can occur within the same tumor. Frequencies of molecular features and of co-mutations within this cohort have been published previously[34].

ENERGY BALANCE-RELATED FACTORS

Three proxy variables were used to assess exposure to energy restriction during childhood, as described in more detail previously[35, 36]: I) place of residence during the Dutch Hunger Winter (1944-45): living in a city in the western part of the Netherlands during the Hunger Winter indicated severe energy restriction, with caloric intake of 400-800 kcal per day at the height of the famine[37, 38]; II) place of residence in 1942, reflecting World War II (WWII; 1940-44): living in a Dutch city in 1942 with more than

40,000 inhabitants was used as an indicator for energy restriction; III) employment status of the father during the Dutch Economic Depression (1932–40): unemployment of the father during the Economic Depression indicated a lack of variation in the food pattern, though sufficient calories were available. Participants of the NLCS were 12-28 years old during the Hunger Winter, 8-28 years old during WWII, and 0-23 years old during the Economic Depression. Height was self-reported at baseline (cm), early adult BMI (age 20 years) was calculated by using self-reported weight at age 20 years and height at baseline (kg/m^2).

COX REGRESSION MODELS

After excluding participants with incomplete or inconsistent data on exposure variables or confounders, 3,911 subcohort members and 1,934 CRC cases were available for analyses (Supplementary Figure S1). Associations between early-life energy balance-related factors and CRC subgroups based on molecular features were investigated stratified on sex and tumor location. Cox proportional hazard models were used to estimate Hazard Ratios (HRs) and 95% confidence intervals (CIs) for the associations between subgroups of CRC and early-life energy restriction proxies (place of residence during the Hunger Winter; place of residence during WWII; employment status of the father during the Economic Depression), height (according to sex-specific quartiles, and per 5 cm increase), and early adult BMI (according to sex-specific quartiles, and per 5 kg/m^2 increase). Standard errors of the HRs were estimated using the Huber-White sandwich estimator to account for additional variance introduced by sampling from the cohort[39]. The proportional hazard assumption was tested using the scaled Schoenfeld residuals[40] and by introducing time-covariate interactions into the models. Cases with rectosigmoid cancer were excluded from analyses.

All multivariable models were adjusted for age, family history of CRC (yes/no), alcohol intake (0; 0.1-4; 5-14; >15 g/day), energy intake at baseline (kcal/day), non-occupational physical activity (min/day), red meat consumption (g/day), and processed meat consumption (g/day). Models on energy restriction and adult-attained height were additionally adjusted for BMI at baseline (kg/m^2). Models on adolescent BMI were additionally adjusted for height (cm). These confounders were based on previous literature in the field[5-7].

Heterogeneity in associations between early-life energy balance-related factors and CRC subgroups based on molecular features were evaluated using an adapted version of the competing risks procedure in Stata developed for the case-cohort design, as described previously[41, 42]. All subgroups were compared pair-wise with the all-wild-type+pMMR subgroup, which was the reference group for heterogeneity tests of all subgroups.

All analyses were conducted in Stata Statistical Software: Release 15 (StataCorp., 2017, College Station, TX).

RESULTS

COHORT CHARACTERISTICS IN SUBGROUPS BASED ON MOLECULAR FEATURES

Information on early-life energy balance-related factors of CRC cases, overall and according to subgroups based on molecular features, are shown in Table 1. Both men and women in the any-mutation/dMMR subgroup were taller compared to those in the all-wild-type+pMMR subgroup, with the exception of men with rectal cancer. Amongst male colon cancer cases, those with a $BRAF_{mut}$ or dMMR tumor were tallest, whereas no clear difference was observed amongst female colon cancer cases. In general, cases in the any-mutation/dMMR subgroup were more often exposed to energy restriction early in life, with the exception of men with colon cancer, where cases in the all-wild-type+pMMR subgroup were more often exposed to energy restriction. Cases in the any-mutation/dMMR subgroup were more often overweight at age 20 years, with the exception of men with rectal cancer. Amongst male colon cancer cases, those with a $KRAS_{mut}$ or $PIK3CA_{mut}$ tumor were most frequently overweight at age 20 years, whereas amongst female colon cancer cases this seemed to be the case for $PIK3CA_{mut}$ and dMMR tumors.

COX REGRESSION ANALYSES

Multivariable-adjusted Cox regression models on early-life energy balance-related factors and risk of CRC in subgroups based on mutation and MMR status are shown in Tables 2-5. Age-adjusted Cox regression models are shown in Supplementary Tables S2-5. Age was included as a time-varying covariate in all models, because of violation of the proportional hazards assumption.

Adult-attained height

Results on associations between adult-attained height and risk of CRC in subgroups based on molecular features, by tumor location and sex, are shown in Table 2. In men, height was positively associated with risk of overall colon cancer, and especially with the any-mutation/dMMR subgroup [HR_{5cm} (95% CI): 1.17 (1.08-1.27), p-$trend_{quartiles}$: 0.001]. Although in men positive associations were observed between height and all subgroups of individual molecular features in colon cancer, associations were strongest for $BRAF_{mut}$ [HR_{5cm} (95% CI): 1.23 (1.06-1.43), p-$trend_{quartiles}$: 0.025] and dMMR colon cancer [HR_{5cm} (95% CI): 1.24 (1.03-1.48), p-$trend_{quartiles}$: 0.057]. Compared to the all-wild-type+pMMR subgroup, statistically significant heterogeneity was observed for any-mutation/dMMR (P-heterogeneity=0.049) and $BRAF_{mut}$ (P-heterogeneity=0.049) colon cancer in men. No associations were observed between height and overall rectal cancer in men, and stratification in subgroups of combinations of molecular features did not lead to clear associations. It seems, however, that height was positively associated with $KRAS_{mut}$ rectal cancer in men, but this association did not reach statistical significance. In women, borderline significant and significant positive associations were observed for height and risk of overall colon and rectal cancer, respectively. These positive associations were observed for the any-mutation/dMMR subgroups [colon: HR_{5cm} (95% CI): 1.09 (0.99-1.19), p-$trend_{quartiles}$: 0.069; rectum: 1.38 (1.12-1.70), p-$trend_{quartiles}$: 0.010], but not for the all-wild-type+pMMR subgroups. For rectal cancer, this positive

Table 1 | Characteristics [mean (SD) or %] of CRC cases in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Total	Wild-type+ pMMR[c]	Any-mutation/ dMMR[d]	*KRAS*$_{mut}$	*PIK3CA*$_{mut}$[e]	*BRAF*$_{mut}$[e]	dMMR[e]
Men - colon							
N[a]	754	309	445	256	150	105	70
Age (years)	61.6 (4.2)	61.2 (4.2)	61.9 (4.2)	62.0 (4.2)	61.5 (4.3)	62.2 (4.1)	62.7 (4.1)
Height (cm)	177.4 (6.8)	176.8 (6.4)	177.8 (7.1)	177.3 (6.9)	177.3 (7.3)	178.3 (7.1)	178.2 (7.3)
Hunger Winter (living in Western city %)	16.6	18.7	15.1	15.9	16.2	11.7	9.0
WWII (living in city %)	46.1	50.7	42.9	42.9	40.5	45.1	34.8
Economic depression (father unemployed %)	10.9	12.8	9.6	10.9	8.3	7.1	7.6
Overweight[b] at age 20 years (%)	8.5	7.7	9.0	10.3	9.7	6.0	6.8
Weight change since age 20 years (kg)	11.4 (9.7)	11.4 (10.1)	11.4 (9.4)	11.0 (9.5)	12.0 (9.8)	11.3 (9.6)	12.5 (10.6)
Men - rectum							
N[a]	224	135	89	68	30	8	1
Age (years)	60.8 (3.9)	60.4 (4.0)	61.4 (3.9)	61.7 (3.9)			
Height (cm)	177.0 (6.7)	177.0 (6.5)	177.0 (6.9)	178.0 (6.4)			
Hunger Winter (living in Western city %)	16.1	14.8	18.0	14.7			
WWII (living in city %)	46.7	45.4	48.7	45.6			
Economic depression (father unemployed %)	10.2	11.7	8.0	7.4			
Overweight[b] at age 20 years (%)	5.6	5.9	5.2	5.1			
Weight change since age 20 years (kg)	9.7 (9.3)	9.3 (8.9)	10.5 (9.9)	10.8 (10.0)			
Women - colon							
N[a]	630	179	451	222	116	173	131
Age (years)	62.0 (4.1)	61.1 (3.9)	62.3 (4.1)	62.2 (4.1)	62.2 (4.2)	62.6 (4.1)	62.3 (4.0)
Height (cm)	165.9 (6.2)	165.7 (5.9)	166.0 (6.3)	166.0 (6.8)	165.7 (7.1)	165.9 (5.9)	165.8 (5.8)

(continued)	Total	Wild-type+ pMMR[c]	Any-mutation/ dMMR[d]	*KRAS*$_{mut}$	*PIK3CA*$_{mut}$[e]	*BRAF*$_{mut}$[e]	dMMR[e]
Women - colon							
Hunger Winter (living in Western city %)	24.4	20.7	25.8	24.4	24.4	28.2	28.7
WWII (living in city %)	48.2	47.6	48.5	46.3	53.3	48.6	52.8
Economic depression (father unemployed %)	11.8	12.1	11.6	14.2	17.3	10.2	11.1
Overweight[b] at age 20 years (%)	7.7	4.4	9.0	10.8	11.4	9.0	11.5
Weight change since age 20 years (kg)	10.1 (10.3)	10.5 (10.4)	9.9 (10.3)	11.0 (10.5)	9.4 (10.5)	9.1 (10.0)	8.3 (11.2)
Women - rectum							
N[a]	131	64	67	55	15	6	2
Age (years)	61.5 (4.2)	60.9 (4.3)	62.0 (4.0)	61.7 (4.0)			
Height (cm)	166.5 (6.5)	165.3 (6.9)	167.6 (5.9)	167.8 (6.2)			
Hunger Winter (living in Western city %)	25.0	23.4	26.6	22.6			
WWII (living in city %)	50.0	42.9	57.5	54.1			
Economic depression (father unemployed %)	9.5	4.8	13.9	14.8			
Overweight[b] at age 20 years (%)	7.6	6.8	8.3	8.0			
Weight change since age 20 years (kg)	9.5 (9.9)	9.5 (10.7)	9.5 (9.1)	10.1 (9.1)			

Abbreviations: SD, standard deviation; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; mut, mutated; WWII, World War II

[a]Total number based on the most complete variable (height). Numbers of other variables might not add up to the same total because of missing values.
[b]Body mass index ≥25.
[c]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR-deficient cases.
[d]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR-deficient.
[e]Analyses for subgroups with <50 cases were not performed.

Table 2 | Multivariable-adjusted HRs[a] and 95% CIs for associations between height and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Height quartiles (cm): range (median)								
Men – colon								
<173 (170)	8935	191	1.00 (ref.)	77	1.00 (ref.)	114	1.00 (ref.)	
173-176 (175)	7680	182	1.17 (0.91-1.50)	84	1.32 (0.93-1.88)	98	1.06 (0.78-1.44)	
177-181 (179)	7097	175	1.22 (0.95-1.57)	78	1.35 (0.94-1.92)	97	1.13 (0.83-1.54)	
>181 (185)	7310	206	1.49 (1.16-1.91)	70	1.26 (0.88-1.81)	136	1.64 (1.23-2.21)	0.049
p-trend			0.002		0.178		0.001	
per 5 cm	31022	754	1.13 (1.05-1.20)	309	1.06 (0.97-1.16)	445	1.17 (1.08-1.27)	0.049
Men – rectum								
<173 (170)	8935	61	1.00 (ref.)	39	1.00 (ref.)	22	1.00 (ref.)	
173-176 (175)	7680	46	0.84 (0.56-1.27)	26	0.73 (0.44-1.23)	20	1.04 (0.55-1.96)	
177-181 (179)	7097	67	1.36 (0.93-2.00)	41	1.27 (0.79-2.05)	26	1.52 (0.83-2.78)	
>181 (185)	7310	50	0.99 (0.66-1.50)	29	0.85 (0.51-1.42)	21	1.27 (0.67-2.40)	0.855
p-trend			0.464		0.943		0.264	
per 5 cm	31022	224	1.04 (0.93-1.15)	135	1.03 (0.91-1.17)	89	1.05 (0.88-1.24)	0.836
Women – colon								
<162 (158)	8764	140	1.00 (ref.)	39	1.00 (ref.)	101	1.00 (ref.)	
162-165 (164)	9216	173	1.17 (0.90-1.52)	54	1.31 (0.85-2.02)	119	1.12 (0.83-1.52)	
166-169 (168)	7771	147	1.17 (0.89-1.54)	45	1.31 (0.83-2.07)	102	1.12 (0.82-1.54)	
>169 (172)	8477	170	1.27 (0.97-1.66)	41	1.08 (0.67-1.73)	129	1.34 (0.99-1.82)	0.459
p-trend			0.107		0.794		0.069	
per 5 cm	34228	630	1.07 (0.99-1.15)	179	1.02 (0.90-1.15)	451	1.09 (0.99-1.19)	0.413
Women – rectum								
<162 (158)	8764	29	1.00 (ref.)	18	1.00 (ref.)	11	1.00 (ref.)	
162-165 (164)	9216	24	0.79 (0.45-1.37)	11	0.59 (0.28-1.23)	13	1.12 (0.49-2.56)	
166-169 (168)	7771	38	1.46 (0.88-2.44)	20	1.25 (0.65-2.42)	18	1.81 (0.83-3.94)	
>169 (172)	8477	40	1.44 (0.86-2.41)	15	0.88 (0.43-1.80)	25	2.37 (1.11-5.06)	0.285
p-trend			0.047		0.788		0.010	
per 5 cm	34228	131	1.15 (0.98-1.35)	64	0.95 (0.75-1.21)	67	1.38 (1.12-1.70)	0.014

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[d]		*BRAF*$_{mut}$[d]		dMMR[d]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Height quartiles (cm): range (median)									
Men – colon									
<173 (170)	8935	74	1.00 (ref.)	40	1.00 (ref.)	24	1.00 (ref.)	16	1.00 (ref.)
173-176 (175)	7680	56	0.63-1.36)	32	1.00 (0.61-1.63)	24	1.21 (0.67-2.18)	16	1.26 (0.62-2.59)
177-181 (179)	7097	55	0.98 (0.67-1.44)	33	1.11 (0.68-1.80)	22	1.22 (0.67-2.24)	16	1.39 (0.68-2.82)
>181 (185)	7310	71	1.34 (0.93-1.94)	45	1.56 (0.98-2.47)	35	1.95 (1.12-3.39)	22	1.92 (0.98-3.73)
p-trend			0.140		0.063		0.025		0.057
per 5 cm	31022	256	1.14 (1.03-1.26)	150	1.11 (0.97-1.27)	105	1.23 (1.06-1.43)*	70	1.24 (1.03-1.48)
Men – rectum									
<173 (170)	8935	13	1.00 (ref.)						
173-176 (175)	7680	16	1.42 (0.67-3.03)						
177-181 (179)	7097	21	2.01 (0.97-4.17)						
>181 (185)	7310	18	1.81 (0.85-3.86)						
p-trend			0.063						
per 5 cm	31022	68	1.14 (0.95-1.36)						
Women – colon									
<162 (158)	8764	53	1.00 (ref.)	29	1.00 (ref.)	37	1.00 (ref.)	27	1.00 (ref.)
162-165 (164)	9216	56	1.03 (0.69-1.54)	25	0.83 (0.48-1.45)	44	1.10 (0.70-1.76)	41	1.42 (0.85-2.36)
166-169 (168)	7771	49	1.07 (0.70-1.63)	30	1.17 (0.68-2.01)	43	1.24 (0.77-2.00)	28	1.10 (0.62-1.93)
>169 (172)	8477	64	1.33 (0.89-1.97)	32	1.17 (0.69-1.97)	49	1.34 (0.84-2.14)	35	1.31 (0.75-2.27)
p-trend			0.164		0.355		0.191		0.567
per 5 cm	34228	222	1.09 (0.95-1.23)	116	1.05 (0.88-1.25)	173	1.08 (0.95-1.23)	131	1.07 (0.92-1.24)
Women – rectum									
<162 (158)	8764	9	1.00 (ref.)						
162-165 (164)	9216	10	1.07 (0.43-2.69)						
166-169 (168)	7771	15	1.87 (0.80-4.38)						
>169 (172)	8477	21	2.43 (1.07-5.53)						
p-trend			0.012						
per 5 cm	34228	55	1.40 (1.11-1.77)*						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; (d/p)MMR, mismatch repair (deficient/proficient); p-het, p-heterogeneity.

*Statistically significant p-heterogeneity; *BRAF*$_{mut}$ men colon: p=0.049; *KRAS*$_{mut}$ women rectum: p=0.017 (reference group: wild-type for *KRAS*, *PIK3CA*, and *BRAF*, and pMMR)

[a]Hazard Ratios were adjusted for age (years; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), BMI at baseline (kg/m^2; continuous), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR-deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR-deficient.

[d]Analyses for subgroups with <50 cases were not performed.

Table 3 | Multivariable-adjusted HRs[a] and 95% CIs for associations between early-life energy restriction and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Place of residence during the Dutch Hunger Winter (1944-45)								
Men – colon	30174	736		300		436		
Non-west	15188	442	1.00 (ref.)	182	1.00 (ref.)	260	1.00 (ref.)	
Western rural	4129	75	0.67 (0.50-0.90)	29	0.63 (0.41-0.97)	46	0.70 (0.49-0.99)	
Western city	6524	122	0.64 (0.50-0.81)	56	0.71 (0.51-1.00)	66	0.59 (0.43-0.80)	0.546
Men – rectum	30174	224		135		89		
Non-west	15188	138	1.00 (ref.)	84	1.00 (ref.)	54	1.00 (ref.)	
Western rural	4129	28	0.76 (0.49-1.17)	19	0.83 (0.49-1.40)	9	0.64 (0.31-1.32)	
Western city	6524	36	0.61 (0.41-0.90)	20	0.55 (0.33-0.92)	16	0.70 (0.39-1.25)	0.781
Women – colon	33722	620		174		446		
Non-west	18083	360	1.00 (ref.)	107	1.00 (ref.)	253	1.00 (ref.)	
Western rural	4851	76	0.79 (0.59-1.06)	20	0.68 (0.41-1.12)	56	0.84 (0.60-1.16)	
Western city	9234	151	0.82 (0.65-1.02)	36	0.64 (0.43-0.95)	115	0.90 (0.70-1.15)	0.369
Women – rectum	33722	128		64		64		
Non-west	18083	76	1.00 (ref.)	37	1.00 (ref.)	39	1.00 (ref.)	
Western rural	4851	12	0.63 (0.34-1.18)	6	0.66 (0.27-1.58)	6	0.61 (0.25-1.48)	
Western city	9234	32	0.87 (0.56-1.34)	15	0.83 (0.45-1.54)	17	0.90 (0.49-1.65)	0.986
Place of residence during World War II (1942)								
Men - colon	23793	560		231		329		
Rural area	11327	283	1.00 (ref.)	105	1.00 (ref.)	178	1.00 (ref.)	
Urban area	11713	258	0.89 (0.72-1.09)	117	1.08 (0.80-1.45)	141	0.77 (0.60-1.00)	0.041
Men - rectum	23973	184		108		76		
Rural area	11327	95	1.00 (ref.)	57	1.00 (ref.)	38	1.00 (ref.)	
Urban area	11713	86	0.89 (0.64-1.23)	49	0.85 (0.56-1.29)	37	0.94 (0.58-1.53)	0.689
Women - colon	26164	483		126		357		
Rural area	11882	235	1.00 (ref.)	62	1.00 (ref.)	173	1.00 (ref.)	
Urban area	13562	233	0.88 (0.71-1.09)	60	0.87 (0.59-1.27)	173	0.88 (0.69-1.13)	0.829
Women - rectum	26164	96		49		47		
Rural area	11882	44	1.00 (ref.)	25	1.00 (ref.)	19	1.00 (ref.)	
Urban area	13562	48	0.98 (0.63-1.51)	21	0.75 (0.40-1.38)	27	1.29 (0.71-2.35)	0.216
Employment of the father during the Dutch Economic Depression (1932-40)								
Men - colon	29841	724		296		428		
Employed	26697	645	1.00 (ref.)	258	1.00 (ref.)	387	1.00 (ref.)	
Unemployed	3145	79	0.97 (0.73-1.30)	38	1.16 (0.79-1.70)	41	0.84 (0.58-1.22)	0.180
Men - rectum	29841	216		128		88		
Employed	26697	194	1.00 (ref.)	113	1.00 (ref.)	81	1.00 (ref.)	
Unemployed	3145	22	0.96 (0.60-1.53)	15	1.16 (0.66-2.05)	7	0.69 (0.31-1.54)	0.757

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Employment of the father during the Dutch Economic Depression (1932-40)								
Women – colon	32597	604		174		430		
Employed	29046	533	1.00 (ref.)	153	1.00 (ref.)	380	1.00 (ref.)	
Unemployed	3552	71	1.09 (0.81-1.47)	21	1.13 (0.69-1.86)	50	1.07 (0.76-1.51)	0.904
Women – rectum	32597	127		62		65		
Employed	29046	115	1.00 (ref.)	59	1.00 (ref.)	56	1.00 (ref.)	
Unemployed	3552	12	0.83 (0.45-1.53)	3	0.40 (0.12-1.30)	9	1.29 (0.63-2.62)	0.892

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study; P-het, P-heterogeneity.

[a]Hazard Ratios were adjusted for age (years; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), BMI at baseline (kg/m^2; continuous), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR deficient.

Table 4 | Multivariable-adjusted HRs[a] and 95% CIs for associations between early-life energy restriction and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Place of residence during the Dutch Hunger Winter (1944-45)									
Men – colon	30174	251		148		103		67	
Non-west	15188	149	1.00 (ref.)	92	1.00 (ref.)	62	1.00 (ref.)	42	1.00 (ref.)
Western rural	4129	29	0.78 (0.50-1.20)	16	0.69 (0.39-1.21)	7	0.44 (0.20-0.98)	6	0.57 (0.24-1.35)
Western city	6524	40	0.61 (0.42-0.90)	24	0.62 (0.39-0.99)	12	0.44 (0.23-0.84)	6	0.34 (0.14-0.83)
Men – rectum	30174	68							
Non-west	15188	45	1.00 (ref.)						
Western rural	4129	8	0.70 (0.32-1.53)						
Western city	6524	10	0.52 (0.26-1.06)						
Women – colon	33722	221		115		170		129	
Non-west	18083	130	1.00 (ref.)	65	1.00 (ref.)	90	1.00 (ref.)	72	1.00 (ref.)
Western rural	4851	24	0.68 (0.43-1.09)	15	0.89 (0.49-1.62)	25	1.05 (0.65-1.69)	14	0.76 (0.41-1.38)
Western city	9234	54	0.81 (0.57-1.14)	28	0.86 (0.54-1.37)	48	1.05 (0.72-1.52)	37	1.03 (0.68-1.57)
Women – rectum	33722	53							
Non-west	18083	33	1.00 (ref.)						
Western rural	4851	6	0.70 (0.28-1.72)						
Western city	9234	12	0.76 (0.37-1.54)						
Place of residence during World War II (1942)									
Men - colon	23793	196		116		71		46	
Rural area	11327	110	1.00 (ref.)	66	1.00 (ref.)	35	1.00 (ref.)	26	1.00 (ref.)
Urban area	11713	84	0.72 (0.53-0.99)*	47	0.73 (0.49-1.09)*	32	0.89 (0.53-1.48)	16	0.66 (0.34-1.29)
Men - rectum	23973	57							
Rural area	11327	31	1.00 (ref.)						
Urban area	11713	26	0.81 (0.46-1.42)						
Women - colon	26164	175		90		138		106	
Rural area	11882	87	1.00 (ref.)	41	1.00 (ref.)	67	1.00 (ref.)	49	1.00 (ref.)
Urban area	13562	81	0.83 (0.60-1.15)	48	1.08 (0.69-1.69)	67	0.88 (0.61-1.27)	56	0.99 (0.66-1.50)
Women - rectum	26164	37							
Rural area	11882	16	1.00 (ref.)						
Urban area	13562	20	1.18 (0.61-2.28)						
Employment of the father during the Dutch Economic Depression (1932-40)									
Men - colon	29841	247		144		99		66	
Employed	26697	220	1.00 (ref.)	132	1.00 (ref.)	92	1.00 (ref.)	61	1.00 (ref.)
Unemployed	3145	27	0.95 (0.61-1.48)	12	0.73 (0.39-1.35)	7	0.62 (0.28-1.36)	5	0.67 (0.26-1.71)
Men - rectum	29841	68							
Employed	26697	63	1.00 (ref.)						
Unemployed	3145	5	0.62 (0.24-1.56)						

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Employment of the father during the Dutch Economic Depression (1932-40)									
Women – colon	32597	212		110		167		126	
Employed	29046	182	1.00 (ref.)	91	1.00 (ref.)	150	1.00 (ref.)	112	1.00 (ref.)
Unemployed	3552	30	1.33 (0.87-2.04)	19	1.62 (0.95-2.78)	17	0.91 (0.53-1.56)	14	1.05 (0.57-1.90)
Women – rectum	32597	54							
Employed	29046	46	1.00 (ref.)						
Unemployed	3552	8	1.43 (0.67-3.04)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; mut, mutated; dMMR, mismatch repair deficiency.

*Statistically significant p-heterogeneity; WWII *KRAS*$_{mut}$ men colon: p=0.049; WWII *PIK3CA*$_{mut}$ men colon: p=0.045 (reference group: wild-type for *KRAS*, *PIK3CA*, and *BRAF*, and dMMR)

[a]Hazard Ratios were adjusted for age (years; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), BMI at baseline (kg/m^2; continuous), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]Analyses for subgroups with <50 cases were not performed.

Table 5 | Multivariable-adjusted HRs[a] and 95% CIs for associations between adolescent BMI and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]		p-het
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	
Quartiles of BMI at age 20 years (kg/m²): range (median)								
Men – colon								
<20.2 (19.2)	6103	155	1.00 (ref.)	64	1.00 (ref.)	91	1.00 (ref.)	
20.2-21.6 (21.0)	6458	148	0.93 (0.70-1.22)	51	0.76 (0.51-1.14)	97	1.05 (0.76-1.46)	
21.7-23.3 (22.4)	6308	151	1.00 (0.75-1.32)	67	1.04 (0.70-1.53)	84	0.97 (0.69-1.37)	
>23.3 (24.3)	6012	173	1.24 (0.94-1.63)	77	1.27 (0.87-1.86)	96	1.21 (0.86-1.69)	0.253
p-trend			0.112		0.099		0.369	
per 5 kg/m²	24881	627	1.04 (0.85-1.27)	259	0.94 (0.71-1.25)	368	1.12 (0.87-1.43)	0.497
Men – rectum								
<20.2 (19.2)	6103	41	1.00 (ref.)	25	1.00 (ref.)	16	1.00 (ref.)	
20.2-21.6 (21.0)	6458	62	1.44 (0.94-2.20)	32	1.18 (0.68-2.05)	30	1.83 (0.98-3.44)	
21.7-23.3 (22.4)	6308	45	1.09 (0.68-1.74)	31	1.19 (0.66-2.12)	14	0.91 (0.42-1.93)	
>23.3 (24.3)	6012	47	1.19 (0.75-1.89)	30	1.19 (0.66-2.15)	17	1.15 (0.57-2.32)	0.293
p-trend			0.808		0.583		0.713	
per 5 kg/m²	24881	195	1.12 (0.85-1.48)	118	1.13 (0.79-1.60)	77	1.09 (0.72-1.66)	0.703
Women – colon								
<19.6 (18.4)	7795	142	1.00 (ref.)	44	1.00 (ref.)	98	1.00 (ref.)	
19.6-21.2 (20.5)	7731	143	1.03 (0.78-1.36)	41	0.94 (0.59-1.49)	102	1.08 (0.79-1.48)	
21.3-23.0 (22.0)	7964	147	1.04 (0.79-1.38)	44	0.98 (0.63-1.54)	103	1.08 (0.78-1.48)	
>23.0 (24.2)	7683	137	1.07 (0.81-1.42)	29	0.67 (0.40-1.11)	108	1.27 (0.92-1.75)	0.284
p-trend			0.645		0.168		0.171	
per 5 kg/m²	31173	569	1.07 (0.91-1.26)	158	0.89 (0.69-1.16)	411	1.15 (0.96-1.39)	0.149
Women – rectum								
<19.6 (18.4)	7795	23	1.00 (ref.)	10	1.00 (ref.)	13	1.00 (ref.)	
19.6-21.2 (20.5)	7731	37	1.71 (0.98-2.98)	20	2.03 (0.92-4.47)	17	1.47 (0.68-3.20)	
21.3-23.0 (22.0)	7964	29	1.30 (0.73-2.31)	15	1.43 (0.62-3.31)	14	1.22 (0.55-2.67)	
>23.0 (24.2)	7683	30	1.44 (0.80-2.60)	14	1.37 (0.59-3.22)	16	1.57 (0.71-2.48)	0.875
p-trend			0.408		0.757		0.359	
per 5 kg/m²	31173	119	1.15 (0.87-1.52)	59	1.03 (0.70-1.52)	60	1.27 (0.86-1.90)	0.847

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[d]		*BRAF*$_{mut}$[d]		dMMR[d]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Quartiles of BMI at age 20 years (kg/m²): range (median)									
Men – colon									
<20.2 (19.2)	6103	50	1.00 (ref.)	32	1.00 (ref.)	19	1.00 (ref.)	19	1.00 (ref.)
20.2-21.6 (21.0)	6458	50	0.96 (0.63-1.47)	33	0.98 (0.59-1.62)	25	1.39 (0.75-2.58)	16	0.89 (0.45-1.76)
21.7-23.3 (22.4)	6308	57	1.16 (0.76-1.77)	23	0.72 (0.41-1.25)	19	1.13 (0.57-2.23)	9	0.52 (0.23-1.17)
>23.3 (24.3)	6012	57	1.26 (0.82-1.94)	35	1.18 (0.71-1.97)	21	1.36 (0.70-2.66)	15	0.94 (0.46-1.93)
p-trend			0.198		0.774		0.515		0.581
per 5 kg/m²	24881	214	1.30 (0.94-1.80)	123	1.08 (0.73-1.60)	84	1.00 (0.66-1.53)	59	0.74 (0.43-1.27)
Men – rectum									
<20.2 (19.2)	6103	11	1.00 (ref.)						
20.2-21.6 (21.0)	6458	24	2.15 (1.04-4.47)						
21.7-23.3 (22.4)	6308	12	1.19 (0.50-2.83)						
>23.3 (24.3)	6012	12	1.32 (0.57-3.05)						
p-trend			0.965						
per 5 kg/m²	24881	59	1.21 (0.75-1.95)						
Women – colon									
<19.6 (18.4)	7795	46	1.00 (ref.)	25	1.00 (ref.)	37	1.00 (ref.)	25	1.00 (ref.)
19.6-21.2 (20.5)	7731	52	1.19 (0.78-1.83)	29	1.24 (0.71-2.17)	35	0.97 (0.59-1.60)	28	1.16 (0.66-2.04)
21.3-23.0 (22.0)	7964	49	1.10 (0.71-1.69)	20	0.85 (0.46-1.56)	43	1.20 (0.74-1.93)	30	1.24 (0.71-2.17)
>23.0 (24.2)	7683	56	1.39 (0.91-2.14)	31	1.45 (0.82-2.54)	41	1.32 (0.81-2.16)	30	1.42 (0.79-2.53)
p-trend			0.192		0.403		0.189		0.231
per 5 kg/m²	31173	203	1.17 (0.91-1.50)	105	1.25 (0.91-1.71)	156	1.29 (0.97-1.70)	113	1.36 (0.98-1.90)
Women – rectum									
<19.6 (18.4)	7795	11	1.00 (ref.)						
19.6-21.2 (20.5)	7731	13	1.32 (0.56-3.10)						
21.3-23.0 (22.0)	7964	12	1.25 (0.53-2.92)						
>23.0 (24.2)	7683	14	1.64 (0.69-3.90)						
p-trend			0.299						
per 5 kg/m²	31173	50	1.22 (0.78-1.90)						

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; (d/p)MMR, mismatch repair (deficient/proficient); p-het, p-heterogeneity.

Note: p-heterogeneity tests for individual molecular features (*KRAS*$_{mut}$, *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, and dMMR) were not statistically significant (reference group for all tests: wild-type for *KRAS*, *PIK3CA*, and *BRAF*, and pMMR).

[a]Hazard Ratios were adjusted for age (years; continuous), height (cm; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR-deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR-deficient.

[d]Analyses for subgroups with <50 cases were not performed.

association was observed for the *KRAS*$_{mut}$ subgroup as well [HR$_{5cm}$ (95% CI): 1.40 (1.11-1.77), p-trend$_{quartiles}$: 0.012]. For colon cancer, positive associations were observed for all subgroups of individual molecular features, but none reached statistical significance. For rectal cancer in women, statistically significant heterogeneity was observed for any-mutation/dMMR (P-heterogeneity=0.014) and *KRAS*$_{mut}$ (P-heterogeneity=0.017) subgroups compared to the all-wild-type+pMMR subgroup (Table 2).

Proxies for early-life energy restriction

Results on associations between energy restriction proxies and risk of CRC in subgroups based on molecular features, by tumor location and sex, are shown in Tables 3 and 4. Living in a western part of the Netherlands during the Dutch Hunger Winter (1944-45) was associated with a decreased risk of overall colon and rectal cancer in men (Table 3). For colon cancer, inverse associations were observed for both the any-mutation/dMMR subgroup [HR$_{western\ rural}$ (95% CI): 0.70 (0.49-0.99); HR$_{western\ city}$: 0.59 (0.43-0.80)] and the all-wild-type+pMMR subgroup [HR$_{western\ rural}$ (95% CI): 0.63 (0.41-0.97); HR$_{western\ city}$: 0.71 (0.51-1.00)]. Although statistically significant inverse associations were observed with all subgroups of individual molecular features in colon cancer, associations were strongest for *BRAF*$_{mut}$ [HR$_{western\ rural}$ (95% CI): 0.44 (0.20-0.98); HR$_{western\ city}$: 0.44 (0.23-0.84)] and dMMR [HR$_{western\ rural}$ (95% CI): 0.57 (0.24-1.35); HR$_{western\ city}$: 0.34 (0.14-0.83)] tumors (Table 4). For rectal cancer, inverse associations were observed for all subgroups of (combinations of) molecular features (Tables 3-4), but only the association with the all-wild-type+pMMR subgroup was statistically significant [HR$_{western\ rural}$ (95% CI): 0.83 (0.49-1.40); HR$_{western\ city}$: 0.55 (0.33-0.92)]. In women, the inverse association between exposure to the Hunger Winter and overall colon cancer risk was borderline significant (Table 3). This association was statistically significant for the all-wild-type+pMMR subgroup [HR$_{western\ rural}$ (95% CI): 0.68 (0.41-1.12); HR$_{western\ city}$: 0.64 (0.43-0.95)], whereas no association was observed for the any-mutation/dMMR subgroup (Table 3). No clear associations were observed for subgroups of individual molecular features in colon cancer in women (Table 4). Furthermore, no clear associations were observed between exposure to the Hunger Winter and risk of overall rectal cancer in women, and stratification in subgroups based on individual molecular features or combinations thereof did not lead to clear associations (Tables 3-4).

Place of residence during WWII (1942) was not associated with overall colon cancer in men (Table 3). However, living in an urban area during WWII showed a borderline significant inverse association with any-mutation/dMMR colon cancer [HR$_{urban\ area}$ (95% CI): 0.77 (0.60-1.00)], but not with all-wild-type+pMMR colon cancer (P-heterogeneity=0.041). Even though inverse associations were observed for all subgroups of individual molecular features, only the association with *KRAS*$_{mut}$ colon cancer reached statistical significance in men [HR$_{urban\ area}$ (95% CI): 0.72 (0.53-0.99)] (Table 4). Compared to the all-wild-type+pMMR subgroup, associations with *KRAS*$_{mut}$ and *PIK3CA*$_{mut}$ were statistically different (P-heterogeneity=0.049 and 0.045, respectively). Place of residence during WWII was not associated with risk of rectal cancer in men, nor with colon or rectal cancer in women, and stratification in subgroups did not lead to clear associations (Tables 3-4).

Lastly, employment status of the father during the Dutch Economic Depression (1932-40) was not associated with overall colon or rectal cancer risk, neither in men nor in women (Table 4). Stratification in subgroups of (combinations of) molecular features did not lead to clear associations (Tables 4-5).

Early adult BMI

Results on associations between early adult BMI (age 20 years) and risk of CRC in subgroups based on molecular features, by tumor location and sex, are shown in Table 5. BMI at age 20 years was not associated with risk of overall colon or rectal cancer in men, and stratification in subgroups based on (combinations of) molecular features only led to a non-significant positive association with $KRAS_{mut}$ colon cancer [$HR_{5kg/m2}$ (95% CI): 1.30 (0.94-1.80); p-$trend_{quartiles}$: 0.198]. No clear associations were observed for overall colon or rectal cancer in women either. However, stratification in subgroups led to borderline significant positive associations for $BRAF_{mut}$ [$HR_{5kg/m2}$ (95% CI): 1.29 (0.97-1.70), p-$trend_{quartiles}$: 0.189] and dMMR [$HR_{5kg/m2}$ (95% CI): 1.36 (0.98-1.90), p-$trend_{quartiles}$: 0.231] colon cancer in women, but not for other subgroups of individual molecular features or combinations thereof. For rectal cancer, stratification in subgroups based on (combinations of) molecular features did not lead to clear associations (Table 5).

DISCUSSION

$KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR have all been associated with the Warburg-effect[11,16-19]. We previously reported differential associations between early-life energy balance-related factors and CRC subtypes based on expression of proteins involved in the Warburg-effect[10]. Using data from a large prospective cohort study, we now investigated associations between early-life energy balance-related factors and risk of CRC subgroups based on $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status. Associations between early-life energy balance-related factors and risk of CRC varied by abovementioned molecular features, as well as by sex and tumor location. Height was positively associated with any-mutation/dMMR CRC but not all-wild-type+pMMR CRC, with the exception of men with rectal cancer, and this difference reached statistically significant heterogeneity in men with colon cancer and women with rectal cancer. Results on early-life energy restriction proxies, reflecting exposure to energy restriction during the Dutch Hunger Winter, WWII, and the Dutch Economic Depression, in relation to risk of CRC subgroups did not show clear patterns. A high early adult BMI (age 20 years) was (non-significantly) associated with an increased risk of $KRAS_{mut}$ colon cancer in men and of $BRAF_{mut}$ and dMMR colon cancer in women.

In the current study, we combined CRC cases into subgroups based on common molecular features ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, dMMR) and studied potential etiological differences between these subgroups. We believe that combining these molecular features into subgroups has some advantages. First, mutations in *KRAS*, *PIK3CA*, and

BRAF have all been associated with metabolic reprogramming towards the Warburg-effect[11, 16-18], and MMR deficiency has previously been associated with estimated presence of the Warburg-effect [19]. Therefore, combining these molecular features, presumed to be involved in the same metabolic phenotype, results in a cleaner reference group compared to groups based on individual features (e.g. *KRAS*-mutated versus –wild-type). We observed that co-occurrence of $KRAS_{mut}$ and $PIK3CA_{mut}$ and of $BRAF_{mut}$ and dMMR was relatively common in the current cohort. Thus, by using the all-wild-type+pMMR subgroup as the reference for all subgroups of individual molecular features, this reference group is less heterogeneous regarding the Warburg-effect compared to, for example, the *KRAS*-wild-type group, which still contains a large number of cases with a *PIK3CA* mutation. Second, using combination subgroups based on multiple molecular features results in increased statistical power, since most individual molecular features occurred in <20% of CRC cases (e.g., MMR deficiency: 10.7%).

Up to now, only a very limited number of studies have investigated associations between early-life energy balance-related factors and risk of CRC in relation to specific (individual) molecular features[23-26]. Our results are concordant with those of a pooled analysis of the NLCS, using 7.3 years of follow-up, and the Melbourne Collaborative Cohort study, which showed stronger associations between height and $BRAF_{mut}$ compared to $BRAF_{wt}$ CRC in both cohorts, and stronger associations for MSI compared to microsatellite stable (MSS) CRC, again in both cohorts[25]. However, our results on height are not in line with those of Brändstedt et al[23, 24], who did not find clear (differences between) associations based on $KRAS_{mut}$, $BRAF_{mut}$, or MMR status, which could have been related to limited statistical power. With respect to associations between early adult BMI and CRC in relation to MSI status, results of the current study are not in line with a previous case-control study[26], which showed the strongest association with MSI low CRC (men and women combined), whereas we observed stronger associations with dMMR colon cancer in women. This difference could be related to the difference in study design (cohort versus case-control) or by the difference in stratification on tumor location and sex. To the best of our knowledge, studies on early-life energy restriction and risk of CRC subgroups based on molecular features are currently lacking.

The current results suggest a role of the molecular features ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR) in the etiological pathway of height with CRC. Regarding colon cancer, in women all individual molecular features seemed equally involved in this etiological pathway, whereas in men it seems that especially *BRAF* mutations and MMR deficiency are involved in this etiological pathway (note: *BRAF* mutations and MMR deficiency often co-occur). Regarding rectal cancer, *KRAS* mutations seem to be involved in this etiological pathway, both in men and women. However, it should be noted that the majority of mutations in rectal cancer were observed in *KRAS*, the other molecular features considered here could thus not be investigated for rectal cancer in the current study due to limited power. The molecular features investigated in the current study have all been associated with the Warburg-effect[11, 16-19]. Since we hypothesized that associations with the any-mutation/dMMR subgroup or subgroups of individual molecular features indicate a higher likelihood of Warburg-effect involvement, the

current results would indicate a potential role of the Warburg-effect in the etiological pathway between height and CRC in both men and women. Previously, we investigated associations between early-life energy balance-related factors and risk of Warburg-subtypes in CRC, based on immunohistochemical expression of proteins involved in the Warburg-effect[10]. In contrast to the current results, those of our previous study did not indicate a role of the Warburg-effect in the etiological pathway between height and CRC risk.

For energy restriction, the current results do not give a clear indication on whether or not the studied molecular features (*KRAS*$_{mut}$, *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, and dMMR) are involved in the etiological pathway with CRC risk, since results varied across the three proxy measures. Neither our results in men nor in women suggest a role of molecular features in the etiological pathway between exposure to energy restriction during the Hunger Winter and CRC risk. In contrast, our results suggest a potential role of these molecular features, in particular *KRAS*$_{mut}$ and/or *PIK3CA*$_{mut}$, in the etiological pathway between exposure to energy restriction during WWII and colon cancer risk in men. It should be noted that *KRAS* and *PIK3CA* were often co-mutated. Other analyses of the WWII and Economic Depression proxies with CRC risk did not show clear associations. Therefore, the current results on energy restriction do not indicate a role of the Warburg-effect in the etiological pathway between energy restriction during the Hunger Winter or Economic Depression and risk of CRC in men or women, but do indicate a possible role of the Warburg-effect in the etiological pathway between energy restriction experienced during WWII and risk of colon cancer in men. This is largely in line with results of our previous study on Warburg-subtypes in CRC[10]. In our previous study, we observed inverse associations between energy restriction during the Hunger Winter and CRC regardless of Warburg-subtypes in men, and with Warburg-low colon cancer in women, which does not indicate a role of the Warburg-effect. However, we previously observed a (non-significant) inverse association between energy restriction during WWII and specifically Warburg-high colon cancer in men, which suggests a role of the Warburg-effect.

As previously described[10], the variation of early adult BMI across participants from the NLCS is limited, possibly explaining the lack of statistically significant associations between early adult BMI and CRC risk. While this, to a certain extent, may have prevented us from detecting involvement of molecular features (*KRAS*$_{mut}$, *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, and dMMR) in the association between early adult BMI and CRC risk in men and women, *BRAF* mutations and/or MMR deficiency seemed to play a role in the risk enhancement observed for colon cancer in women with a high early adult BMI, whereas *KRAS* mutations may be involved in the risk enhancement for colon cancer in men. The Warburg-effect might thus potentially play a role in the etiological pathway between early adult BMI and colon cancer. These results are partly in line with our previous study on Warburg-subtypes in CRC[10], where we observed a positive association between early adult BMI and specifically Warburg-high colon cancer in men, and with Warburg-moderate rectal cancer in women.

All in all, results of the current study are only partly in line with those of our previous study on Warburg-subtypes in CRC assessed by immunohistochemistry. As mentioned, the molecular features that were considered in the current study have all been associated with the Warburg-effect[11, 16-19]. However, these molecular features are additionally known for their involvement in numerous diverse (oncogenic) cellular pathways for cell growth, differentiation, proliferation and survival[20-22]. These molecular features might thus not always be a good reflection of the Warburg-effect. In addition, whilst being wild-type for the genes currently studied as well as MMR proficient, tumors of cases in the all-wild-type+pMMR subgroup might still be characterized by other molecular features that were not assessed in the current study. These molecular features may possibly also be associated with the Warburg-effect, potentially reducing any contrast between the all-wild-type+pMMR group and other groups. Nevertheless, the combination of molecular features into the all-wild-type+pMMR and any-mutation/dMMR subgroups seems to be a straightforward way of subtyping CRC cases.

Strengths of the current study are the prospective cohort design with long follow-up (20.3 years) and availability of DNA from FFPE tumor material from a large number of incident CRC cases. A limitation of the current study is that we did not have a validation cohort available to confirm our results. Therefore, replication of the current results in additional large prospective cohorts is needed. In addition, despite the large sample size, the number of cases in final statistical analyses was limited for some groups (especially rectal cancer) due to heterogeneity in sex and tumor location. Furthermore, the use of the MassARRAY technology in the current study to detect mutations has been shown to be suitable for mutation typing in (older) FFPE material[43]. However, even though the ColoCarta panel includes most known mutations in *KRAS* (99%) and *BRAF* (98%), only 78% of known *PIK3CA* mutations are included[44]. As the most common *PIK3CA* mutations are included[45], it appears to be unlikely that additional detection of less common mutations would have altered the current results. As for MSI status, the usage of MLH1 and MSH2 immunohistochemical expression as an indicator of MSI status might have led to misclassification of some of the cases since not all MMR genes were included. However, it has been shown that loss of MLH1 or MSH2 expression was observed in ~90% of MSI cases[46].

Another strength of the current study is that the NLCS provides the unique opportunity to investigate associations between exposure to (severe) energy restriction early in life and risk of CRC in relation to common molecular features. Nevertheless, the proxy measures for energy restriction during the Dutch Hunger Winter, WWII, and the Dutch Economic Depression might entail exposure misclassification, since individual data on dietary exposures during those times were not available. However, it has previously been shown that the proxy measure for energy restriction during the Hunger Winter was reasonably adequate among female subcohort members[47].

In conclusion, results from this large prospective cohort study provide further insights into the associations between early-life energy balance-related factors and CRC risk according to $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and MMR status. Our results indicate a role of these molecular features in the etiological pathway between height and CRC risk. *BRAF* mutations and/or MMR deficiency seemed to be mainly involved in the association of height with colon cancer in men, whereas *KRAS* mutations seem to be important for rectal cancer in both men and women. Furthermore, *KRAS* mutations might potentially be involved in the etiological pathway between early adult BMI and colon cancer risk in men, whereas *BRAF* mutations and/or MMR deficiency potentially play a role in the etiological pathway between early adult BMI and colon cancer risk in women.

REFERENCES

1. Hughes LA, van den Brandt PA, Goldbohm RA, De Goeij AF, De Bruine AP, Van Engeland M, Weijenberg MP. Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from the Netherlands Cohort Study. International journal of epidemiology. 2010;39(5):1333-44.
2. Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age–period–cohort analysis. European journal of cancer prevention. 2002;11(5):489-95.
3. Svensson E, Møller B, Tretli S, Barlow L, Engholm G, Pukkala E, Rahu M, Tryggvadottir L. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. Cancer Causes & Control. 2005;16(3):215-23.
4. Garcia H, Song M. Early-life obesity and adulthood colorectal cancer risk: a meta-analysis. Revista Panamericana de Salud Pública. 2019;43:e3.
5. Karahalios A, English DR, Simpson JA. Weight change and risk of colorectal cancer: a systematic review and meta-analysis. American journal of epidemiology. 2015;181(11):832-45.
6. Abar L, Vieira AR, Aune D, Sobiecki JG, Vingeliene S, Polemiti E, Stevens C, Greenwood DC, Chan DS, Schlesinger S. Height and body fatness and colorectal cancer risk: an update of the WCRF–AICR systematic review of published prospective studies. European journal of nutrition. 2018;57(5):1701-20.
7. Gunnell D, Okasha M, Davey Smith G, Oliver S, Sandhu J, Holly J. Height, leg length, and cancer risk: a systematic review. Epidemiologic reviews. 2001;23(2):313-42.
8. Xie SH, Lagergren J. A possible link between famine exposure in early life and future risk of gastrointestinal cancers: Implications from age-period-cohort analysis. International journal of cancer. 2017;140(3):636-45.
9. Keinan-Boker L, Vin-Raviv N, Liphshitz I, Linn S, Barchana M. Cancer incidence in Israeli jewish survivors of World War II. Journal of the National Cancer Institute. 2009;101(21):1489-500.
10. Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors in childhood and adolescence and risk of colorectal cancer expressing different levels of proteins involved in the Warburg-effect. International Journal of Cancer.
11. Levine AJ, Puzio-Kuter AM. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. Science. 2010;330(6009):1340-4.
12. Feron O. Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. Radiotherapy and oncology. 2009;92(3):329-33.
13. Warburg O. The metabolism of carcinoma cells. The Journal of Cancer Research. 1925;9(1):148-63.
14. Schwartz L, T Supuran C, O Alfarouk K. The Warburg effect and the hallmarks of cancer. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents). 2017;17(2):164-70.
15. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. cell. 2011;144(5):646-74.
16. Kimmelman AC. Metabolic dependencies in RAS-driven cancers. Clinical Cancer Research. 2015;21(8):1828-34.
17. Hutton JE, Wang X, Zimmerman LJ, Slebos RJ, Trenary IA, Young JD, Li M, Liebler DC. Oncogenic KRAS and BRAF drive metabolic reprogramming in colorectal cancer. Molecular & cellular proteomics.

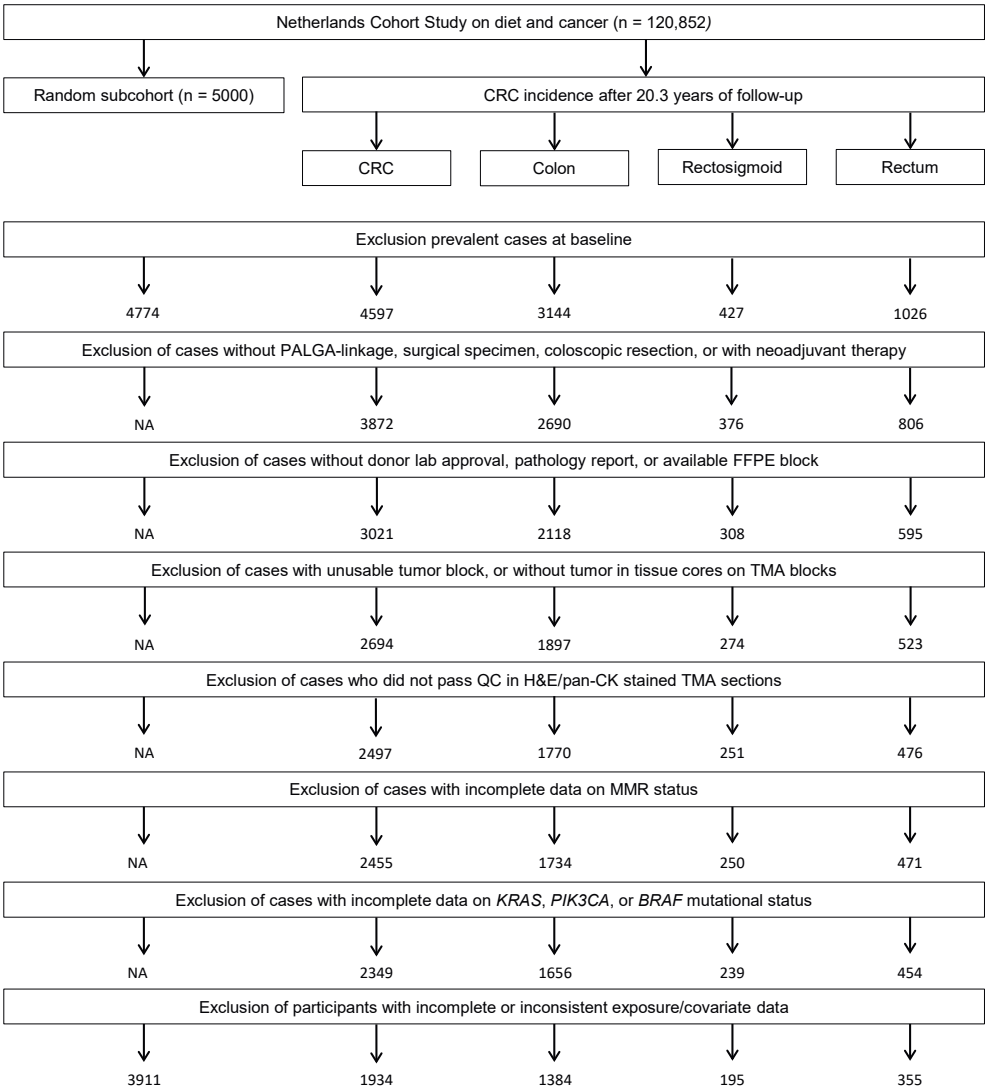
2016;15(9):2924-38.
18. Jiang W, He T, Liu S, Zheng Y, Xiang L, Pei X, Wang Z, Yang H. The PIK3CA E542K and E545K mutations promote glycolysis and proliferation via induction of the β-catenin/SIRT3 signaling pathway in cervical cancer. Journal of hematology & oncology. 2018;11(1):1-15.
19. Offermans K, Jenniskens JCA, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Expression of proteins associated with the Warburg-effect and survival in colorectal cancer. J Pathol Clin Res. 2021.
20. Li W, Qiu T, Dong L, Zhang F, Guo L, Ying J. Prevalence and characteristics of PIK3CA mutation in mismatch repair-deficient colorectal cancer. Journal of Cancer. 2020;11(13):3827.
21. Haluska F, Pemberton T, Ibrahim N, Kalinsky K, editors. The RTK/RAS/BRAF/PI3K pathways in melanoma: biology, small molecule inhibitors, and potential applications. Seminars in oncology; 2007: Elsevier.
22. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology. 2010;138(6):2073-87. e3.
23. Brändstedt J, Wangefjord S, Borgquist S, Nodin B, Eberhard J, Manjer J, Jirström K. Influence of anthropometric factors on tumour biological characteristics of colorectal cancer in men and women: a cohort study. Journal of translational medicine. 2013;11(1):1-13.
24. Brändstedt J, Wangefjord S, Nodin B, Eberhard J, Sundström M, Manjer J, Jirström K. 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(6):e98964.
25. Hughes LA, Williamson EJ, van Engeland M, Jenkins MA, Giles GG, Hopper JL, Southey MC, Young JP, Buchanan DD, Walsh MD. Body size and risk for colorectal cancers showing BRAF mutations or microsatellite instability: a pooled analysis. International journal of epidemiology. 2012;41(4):1060-72.
26. Campbell PT, Jacobs ET, Ulrich CM, Figueiredo JC, Poynter JN, McLaughlin JR, Haile RW, Jacobs EJ, Newcomb PA, Potter JD. Case–control study of overweight, obesity, and colorectal cancer risk, overall and by tumor microsatellite instability status. Journal of the National Cancer Institute. 2010;102(6):391-400.
27. van den Brandt PA, Goldbohm RA, Veer PVT, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. Journal of clinical epidemiology. 1990;43(3):285-95.
28. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika. 1986;73(1):1-11.
29. van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch Cancer Registry for Epidemiological Research. Int J Epidemiol. 1990;19(3):553-8.
30. Goldbohm RA, van den Brandt PA, Dorant E. Estimation of the coverage of Dutch municipalities by cancer registries and PALGA based on hospital discharge data. 1994.
31. Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors and risk of colorectal cancer expressing different levels of proteins involved in the Warburg-effect. Cancer Epidemiology Biomarkers & Prevention. 2021:cebp.EPI-21-0678-E.2021.
32. Richman SD, Adams R, Quirke P, Butler R, Hemmings G, Chambers P, Roberts H, James MD, Wozniak S, Bathia R. Pre-trial inter-laboratory analytical validation of the FOCUS4 personalised therapy trial. Journal

of clinical pathology. 2016;69(1):35-41.
33. Jenniskens JCA, Offermans K, Samarska I, Fazzi GE, Simons CCJM, Smits KM, Schouten LJ, Weijenberg MP, van den Brandt PA, Grabsch HI. Validity and reproducibility of immunohistochemical scoring by trained non-pathologists on Tissue MicroArrays. Cancer Epidemiology and Prevention Biomarkers. 2021;30(10):1867-74.
34. Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, van der Meer JR, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors and risk of colorectal cancer based on KRAS, PIK3CA, and BRAF mutations and MMR status. Journal of Cancer Research and Clinical Oncology. 2022:1-20.
35. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey L. Energy restriction early in life and colon carcinoma risk. Cancer. 2003;97(1):46-55.
36. Simons CCJM, Schouten LJ, Godschalk R, van Engeland M, van den Brandt PA, van Schooten FJ, Weijenberg MP. Body size, physical activity, genetic variants in the insulin-like growth factor pathway and colorectal cancer risk. Carcinogenesis. 2015;36(9):971-81.
37. Trienekens GM. Tussen ons volk en de honger: de voedselvoorziening, 1940-19451985.
38. Burger G, Drummond J, Sandstead H. Malnutrition and starvation in Western Netherlands. Part I The Hague, Netherlands: General State Printing Office. 1948.
39. Lin DY, Wei L-J. The robust inference for the Cox proportional hazards model. Journal of the American statistical Association. 1989;84(408):1074-8.
40. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika. 1982;69(1):239-41.
41. De Vogel S, Bongaerts BW, Wouters KA, Kester AD, Schouten LJ, de Goeij AF, de Bruïne AP, Goldbohm RA, van den Brandt PA, van Engeland M. Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. Carcinogenesis. 2008;29(9):1765-73.
42. Wacholder S, Gail MH, Pee D, Brookmeyer R. Alternative variance and efficiency calculations for the case-cohort design. Biometrika. 1989;76(1):117-23.
43. Fleitas T, Ibarrola-Villava M, Ribas G, Cervantes A. MassARRAY determination of somatic oncogenic mutations in solid tumors: Moving forward to personalized medicine. Cancer treatment reviews. 2016;49:57-64.
44. Fumagalli D, Gavin PG, Taniyama Y, Kim S-I, Choi H-J, Paik S, Pogue-Geile KL. A rapid, sensitive, reproducible and cost-effective method for mutation profiling of colon cancer and metastatic lymph nodes. BMC cancer. 2010;10(1):101.
45. Gray RT, Cantwell MM, Coleman HG, Loughrey MB, Bankhead P, McQuaid S, O'neill RF, Arthur K, Bingham V, McGready C. Evaluation of PTGS2 expression, PIK3CA mutation, aspirin use and colon cancer survival in a population-based cohort study. Clinical and translational gastroenterology. 2017;8(4):e91.
46. Lanza G, Gafà R, Maestri I, Santini A, Matteuzzi M, Cavazzini L. Immunohistochemical pattern of MLH1/MSH2 expression is related to clinical and pathological features in colorectal adenocarcinomas with microsatellite instability. Modern Pathology. 2002;15(7):741.
47. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey L. Diet in adolescence and the risk of breast cancer: results of the Netherlands Cohort Study. Cancer Causes & Control. 1999;10(3):189-99.

SUPPLEMENTARY METHODS

DEPARAFFINIZATION OF FFPE SECTIONS

FFPE tissue sections containing tumor primary tumor were deparaffinized using an adapted version of the protocol for Purification of genomic DNA from FFPE tissue using the QIAamp® DNA FFPE Tissue Kit and Deparaffinization Solution (Qiagen, Hilden, Germany). The adapted protocol included the following steps: I) Add 320 µl Deparaffinization Solution to 2 x 20 µm sections and vortex vigorously for 10 s. Centrifuge briefly to collect the sample in the bottom of the tube; II) Incubate at 56°C for 3 min, and then allow to cool at room temperature (15–25°C); III) Add 200 µl Buffer ATL, and mix by vortexing; IV) Centrifuge for 1 min at 11,000 x g (10,000 rpm); V) Add 20 µl proteinase K to the lower clear phase and mix gently by pipetting up and down; VI) Incubate at 56°C for 1 h; VII) Mix clear phase by pipetting up and down; VIII) Incubate at 90°C for 1 h; IX) Briefly centrifuge the 1.5 ml tube to remove drops from inside the lid; X) Transfer the 2 ml microcentrifuge tubes to QiaSymphony.



Supplementary Figure S1 | Flow diagram of the number of CRC cases and subcohort members; NLCS, 1986-2006. *Abbreviations: CRC, colorectal cancer; NA, not applicable; PALGA, Dutch Pathology Registry; FFPE, formalin-fixed paraffin-embedded; TMA, tissue microarray; QC, quality control; H&E, Hematoxylin & Eosin; pan-CK, pan-cytokeratin; MMR, mismatch repair.*

Supplementary Table S1 | ColoCarta panel genes and mutations.

Gene	Assay	Mutation
BRAF	15/16	*V600E/K/L/M/R*
	9	*D594G/V*
KRAS	1	*G12A/D/V*
	2	*G12C/R/S*
	4	*G13D/V*
	5	*A59T*
	7	*Q61L/P/R*
	8	*Q61H_A/H_G*
PIK3CA	1	*R88Q*
	3	*C420R*
	5	*E542K*
	6	*E545K*
	7	*Q546K*
	8	*H701P*
	9	*H1047L/R*
NRAS	1	*G12A/D/V*
	2	*G12C/R/S*
	3	*G13A/D/V*
	4	*G13C/R/S*
	7	*Q61H*
	8	*Q61E/K*
HRAS	6	*Q61L/P/R*
MET	1	*R970C*
	2	*T992I*

Supplementary Table S2 | Age-adjusted HRs[a] and 95% CIs for associations between height and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Height quartiles (cm): range (median)							
Men – colon							
<173 (170)	8935	191	1.00 (ref.)	77	1.00 (ref.)	114	1.00 (ref.)
173-176 (175)	7680	182	1.13 (0.89-1.44)	84	1.29 (0.92-1.80)	98	1.03 (0.76-1.39)
177-181 (179)	7097	175	1.17 (0.91-1.49)	78	1.28 (0.91-1.81)	97	1.09 (0.80-1.48)
>181 (185)	7310	206	1.35 (1.06-1.71)	70	1.12 (0.79-1.59)	136	1.51 (1.13-2.00)
p-trend			0.016		0.500		0.006
per 5 cm	31022	754	1.10 (1.03-1.17)	309	1.03 (0.94-1.12)	445	1.15 (1.06-1.24)
Men – rectum							
<173 (170)	8935	61	1.00 (ref.)	39	1.00 (ref.)	22	1.00 (ref.)
173-176 (175)	7680	46	0.88 (0.59-1.32)	26	0.77 (0.46-1.29)	20	1.07 (0.58-1.99)
177-181 (179)	7097	67	1.39 (0.96-2.01)	41	1.32 (0.83-2.08)	26	1.51 (0.84-2.70)
>181 (185)	7310	50	1.00 (0.67-1.48)	29	0.89 (0.54-1.47)	21	1.19 (0.64-2.20)
p-trend			0.451		0.801		0.351
per 5 cm	31022	224	1.04 (0.94-1.15)	135	1.04 (0.92-1.18)	89	1.03 (0.88-1.21)
Women – colon							
<162 (158)	8764	140	1.00 (ref.)	39	1.00 (ref.)	101	1.00 (ref.)
162-165 (164)	9216	173	1.17 (0.90-1.51)	54	1.31 (0.85-2.02)	119	1.11 (0.83-1.50)
166-169 (168)	7771	147	1.18 (0.90-1.54)	45	1.29 (0.82-2.03)	102	1.14 (0.84-1.55)
>169 (172)	8477	170	1.27 (0.98-1.65)	41	1.08 (0.68-1.70)	129	1.35 (1.01-1.81)
p-trend			0.083		0.798		0.049
per 5 cm	34228	630	1.06 (0.99-1.15)	179	1.02 (0.90-1.14)	451	1.09 (0.99-1.18)
Women – rectum							
<162 (158)	8764	29	1.00 (ref.)	18	1.00 (ref.)	11	1.00 (ref.)
162-165 (164)	9216	24	0.79 (0.45-1.37)	11	0.58 (0.27-1.25)	13	1.12 (0.50-2.52)
166-169 (168)	7771	38	1.48 (0.90-2.45)	20	1.26 (0.65-2.40)	18	1.86 (0.87-3.98)
>169 (172)	8477	40	1.44 (0.87-2.36)	15	0.86 (0.43-1.72)	25	2.40 (1.16-4.95)
p-trend			0.036		0.834		0.006
per 5 cm	34228	131	1.15 (0.98-1.34)	64	0.95 (0.76-1.19)	67	1.38 (1.14-1.68)

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[d]		*BRAF*$_{mut}$[d]		dMMR[d]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Height quartiles (cm): range (median)									
Men – colon									
<173 (170)	8935	74	1.00 (ref.)	40	1.00 (ref.)	24	1.00 (ref.)	16	1.00 (ref.)
173-176 (175)	7680	56	0.90 (0.62-1.31)	32	0.95 (0.59-1.53)	24	1.20 (0.67-2.15)	16	1.22 (0.60-2.47)
177-181 (179)	7097	55	0.95 (0.65-1.39)	33	1.05 (0.65-1.70)	22	1.19 (0.66-2.15)	16	1.30 (0.64-2.63)
>181 (185)	7310	71	1.21 (0.85-1.73)	45	1.40 (0.90-2.19)	35	1.87 (1.09-3.18)	22	1.78 (0.92-3.43)
p-trend			0.303		0.134		0.031		0.090
per 5 cm	31022	256	1.11 (1.01-1.22)	150	1.09 (0.95-1.23)	105	1.21 (1.05-1.39)	70	1.21 (1.01-1.43)
Men – rectum									
<173 (170)	8935	13	1.00 (ref.)						
173-176 (175)	7680	16	1.46 (0.70-3.07)						
177-181 (179)	7097	21	2.08 (1.03-4.20)						
>181 (185)	7310	18	1.75 (0.84-3.62)						
p-trend			0.064						
per 5 cm	31022	68	1.13 (0.96-1.34)						
Women – colon									
<162 (158)	8764	53	1.00 (ref.)	29	1.00 (ref.)	37	1.00 (ref.)	27	1.00 (ref.)
162-165 (164)	9216	56	1.00 (0.67-1.49)	25	0.82 (0.47-1.42)	44	1.12 (0.71-1.77)	41	1.43 (0.87-2.37)
166-169 (168)	7771	49	1.05 (0.69-1.58)	30	1.17 (0.69-1.98)	43	1.31 (0.82-2.07)	28	1.16 (0.67-2.01)
>169 (172)	8477	64	1.27 (0.86-1.88)	32	1.17 (0.70-1.96)	49	1.40 (0.90-2.20)	35	1.36 (0.81-2.30)
p-trend			0.214		0.340		0.105		0.404
per 5 cm	34228	222	1.07 (0.95-1.21)	116	1.05 (0.88-1.24)	173	1.09 (0.96-1.23)	131	1.08 (0.94-1.24)
Women – rectum									
<162 (158)	8764	9	1.00 (ref.)						
162-165 (164)	9216	10	1.05 (0.43-2.61)						
166-169 (168)	7771	15	1.88 (0.82-4.35)						
>169 (172)	8477	21	2.44 (1.10-5.41)						
p-trend			0.009						
per 5 cm	34228	55	1.40 (1.12-1.75)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; (d/p)MMR, mismatch repair (deficient/proficient).

[a]Hazard Ratios were adjusted for age (years; continuous) age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR-deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR-deficient.

[d]Analyses for subgroups with <50 cases were not performed.

Supplementary Table S3 | Age-adjusted HRs[a] and 95% CIs for associations between early-life energy restriction and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Place of residence during the Dutch Hunger Winter (1944-45)							
Men – colon	30174	736		300		436	
Non-west	15188	442	1.00 (ref.)	182	1.00 (ref.)	260	1.00 (ref.)
Western rural	4129	75	0.65 (0.48-0.86)	29	0.60 (0.39-0.91)	46	0.68 (0.48-0.96)
Western city	6524	122	0.71 (0.54-0.81)	56	0.71 (0.51-0.98)	66	0.59 (0.44-0.79)
Men – rectum	30174	224		135		89	
Non-west	15188	138	1.00 (ref.)	84	1.00 (ref.)	54	1.00 (ref.)
Western rural	4129	28	0.75 (0.49-1.16)	19	0.83 (0.49-1.39)	9	0.63 (0.30-1.29)
Western city	6524	36	0.60 (0.41-0.89)	20	0.55 (0.33-0.91)	16	0.69 (0.39-1.22)
Women – colon	33722	620		174		446	
Non-west	18083	360	1.00 (ref.)	107	1.00 (ref.)	253	1.00 (ref.)
Western rural	4851	76	0.80 (0.60-1.07)	20	0.70 (0.43-1.15)	56	0.85 (0.61-1.17)
Western city	9234	151	0.82 (0.66-1.03)	36	0.66 (0.45-0.99)	115	0.89 (0.70-1.14)
Women – rectum	33722	128		64		64	
Non-west	18083	76	1.00 (ref.)	37	1.00 (ref.)	39	1.00 (ref.)
Western rural	4851	12	0.59 (0.32-1.11)	6	0.61 (0.25-1.45)	6	0.58 (0.24-1.39)
Western city	9234	32	0.83 (0.54-1.27)	15	0.81 (0.44-1.48)	17	0.85 (0.48-1.53)
Place of residence during World War II (1942)							
Men - colon	23793	560		231		329	
Rural area	11327	283	1.00 (ref.)	105	1.00 (ref.)	178	1.00 (ref.)
Urban area	11713	258	0.88 (0.72-1.07)	117	1.08 (0.81-1.44)	141	0.76 (0.59-0.97)
Men - rectum	23973	184		108		76	
Rural area	11327	95	1.00 (ref.)	57	1.00 (ref.)	38	1.00 (ref.)
Urban area	11713	86	0.88 (0.65-1.21)	49	0.84 (0.57-1.25)	37	0.94 (0.59-1.50)
Women - colon	26164	483		126		357	
Rural area	11882	235	1.00 (ref.)	62	1.00 (ref.)	173	1.00 (ref.)
Urban area	13562	233	0.87 (0.70-1.07)	60	0.86 (0.59-1.25)	173	0.87 (0.68-1.10)
Women - rectum	26164	96		49		47	
Rural area	11882	44	1.00 (ref.)	25	1.00 (ref.)	19	1.00 (ref.)
Urban area	13562	48	0.96 (0.62-1.47)	21	0.75 (0.41-1.37)	27	1.22 (0.67-2.23)
Employment of the father during the Dutch Economic Depression (1932-40)							
Men - colon	29841	724		296		428	
Employed	26697	645	1.00 (ref.)	258	1.00 (ref.)	387	1.00 (ref.)
Unemployed	3145	79	1.01 (0.76-1.35)	38	1.23 (0.85-1.79)	41	0.87 (0.61-1.25)
Men - rectum	29841	216		128		88	
Employed	26697	194	1.00 (ref.)	113	1.00 (ref.)	81	1.00 (ref.)
Unemployed	3145	22	0.96 (0.60-1.53)	15	1.14 (0.65-1.99)	7	0.72 (0.33-1.59)

(continued)	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Employment of the father during the Dutch Economic Depression (1932-40)							
Women – colon	32597	604		174		430	
Employed	29046	533	1.00 (ref.)	153	1.00 (ref.)	380	1.00 (ref.)
Unemployed	3552	71	1.08 (0.80-1.44)	21	1.14 (0.70-1.85)	50	1.05 (0.75-1.47)
Women – rectum	32597	127		62		65	
Employed	29046	115	1.00 (ref.)	59	1.00 (ref.)	56	1.00 (ref.)
Unemployed	3552	12	0.86 (0.46-1.59)	3	0.43 (0.13-1.38)	9	1.29 (0.63-2.66)

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; (d/p)MMR, mismatch repair (deficient/proficient); NLCS, Netherlands Cohort Study.

[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR deficient.

Supplementary Table S4 | Age-adjusted HRs[a] and 95% CIs for associations between early-life energy restriction and CRC for individual mutations and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Place of residence during the Dutch Hunger Winter (1944-45)									
Men – colon	30174	251		148		103		67	
Non-west	15188	149	1.00 (ref.)	92	1.00 (ref.)	62	1.00 (ref.)	42	1.00 (ref.)
Western rural	4129	29	0.75 (0.49-1.15)	16	0.66 (0.38-1.15)	7	0.43 (0.20-0.96)	6	0.56 (0.24-1.34)
Western city	6524	40	0.62 (0.43-0.90)	24	0.60 (0.38-0.96)	12	0.45 (0.24-0.85)	6	0.34 (0.14-0.80)
Men – rectum	30174	68							
Non-west	15188	45	1.00 (ref.)						
Western rural	4129	8	0.68 (0.32-1.47)						
Western city	6524	10	0.52 (0.26-1.04)						
Women – colon	33722	221		115		170		129	
Non-west	18083	130	1.00 (ref.)	65	1.00 (ref.)	90	1.00 (ref.)	72	1.00 (ref.)
Western rural	4851	24	0.70 (0.45-1.11)	15	0.88 (0.49-1.57)	25	1.07 (0.67-1.71)	14	0.75 (0.41-1.35)
Western city	9234	54	0.81 (0.58-1.14)	28	0.84 (0.53-1.33)	48	1.04 (0.72-1.51)	37	1.01 (0.67-1.53)
Women – rectum	33722	53							
Non-west	18083	33	1.00 (ref.)						
Western rural	4851	6	0.68 (0.28-1.65)						
Western city	9234	12	0.71 (0.36-1.40)						
Place of residence during World War II (1942)									
Men - colon	23793	196		116		71		46	
Rural area	11327	110	1.00 (ref.)	66	1.00 (ref.)	35	1.00 (ref.)	26	1.00 (ref.)
Urban area	11713	84	0.73 (0.54-0.99)	47	0.69 (0.46-1.01)	32	0.87 (0.53-1.42)	16	0.58 (0.31-1.09)
Men - rectum	23973	57							
Rural area	11327	31	1.00 (ref.)						
Urban area	11713	26	0.81 (0.47-1.38)						
Women - colon	26164	175		90		138		106	
Rural area	11882	87	1.00 (ref.)	41	1.00 (ref.)	67	1.00 (ref.)	49	1.00 (ref.)
Urban area	13562	81	0.81 (0.58-1.11)	48	1.01 (0.65-1.56)	67	0.87 (0.61-1.24)	56	1.01 (0.67-1.50)
Women - rectum	26164	37							
Rural area	11882	16	1.00 (ref.)						
Urban area	13562	20	1.08 (0.55-2.12)						
Employment of the father during the Dutch Economic Depression (1932-40)									
Men - colon	29841	247		144		99		66	
Employed	26697	220	1.00 (ref.)	132	1.00 (ref.)	92	1.00 (ref.)	61	1.00 (ref.)
Unemployed	3145	27	1.00 (0.65-1.54)	12	0.76 (0.41-1.40)	7	0.62 (0.28-1.36)	5	0.66 (0.27-1.67)
Men - rectum	29841	68							
Employed	26697	63	1.00 (ref.)						
Unemployed	3145	5	0.66 (0.26-1.65)						

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[b]		*BRAF*$_{mut}$[b]		dMMR[b]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Employment of the father during the Dutch Economic Depression (1932-40)									
Women – colon	32597	212		110		167		126	
Employed	29046	182	1.00 (ref.)	91	1.00 (ref.)	150	1.00 (ref.)	112	1.00 (ref.)
Unemployed	3552	30	1.32 (0.87-2.01)	19	1.67 (0.99-2.82)	17	0.90 (0.53-1.53)	14	1.00 (0.56-1.80)
Women – rectum	32597	54							
Employed	29046	46	1.00 (ref.)						
Unemployed	3552	8	1.41 (0.66-3.03)						

Abbreviations: HR, hazard ratio; CI, confidence interval; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; mut, mutated; dMMR, mismatch repair deficiency.
[a]Hazard Ratios were adjusted for age (years; continuous) and age was included as a time-varying covariate.
[b]Analyses for subgroups with <50 cases were not performed.

Supplementary Table S5 | Age-adjusted HRs[a] and 95% CIs for associations between adolescent BMI and CRC in subgroups based on mutation and MMR status, by sex and tumor location; NLCS, 1986-2006.

	Person years at risk	Total		Wild-type+pMMR[b]		Any-mutation/dMMR[c]	
		n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)	n_{cases}	HR (95% CI)
Quartiles of BMI at age 20 years (kg/m²): range (median)							
Men – colon							
<20.2 (19.2)	6103	155	1.00 (ref.)	64	1.00 (ref.)	91	1.00 (ref.)
20.2-21.6 (21.0)	6458	148	0.91 (0.69-1.19)	51	0.74 (0.50-1.11)	97	1.02 (0.74-1.41)
21.7-23.3 (22.4)	6308	151	0.94 (0.71-1.23)	67	1.00 (0.69-1.46)	84	0.89 (0.64-1.25)
>23.3 (24.3)	6012	173	1.16 (0.89-1.51)	77	1.24 (0.86-1.78)	96	1.10 (0.80-1.53)
p-trend			0.263		0.125		0.762
per 5 kg/m²	24881	627	1.01 (0.83-1.23)	259	0.94 (0.72-1.24)	368	1.07 (0.84-1.35)
Men – rectum							
<20.2 (19.2)	6103	41	1.00 (ref.)	25	1.00 (ref.)	16	1.00 (ref.)
20.2-21.6 (21.0)	6458	62	1.41 (0.92-2.14)	32	1.17 (0.68-2.02)	30	1.78 (0.96-3.32)
21.7-23.3 (22.4)	6308	45	1.05 (0.67-1.65)	31	1.18 (0.68-2.04)	14	0.85 (0.41-1.76)
>23.3 (24.3)	6012	47	1.17 (0.75-1.82)	30	1.21 (0.69-2.10)	17	1.10 (0.55-2.22)
p-trend			0.878		0.528		0.574
per 5 kg/m²	24881	195	1.09 (0.84-1.43)	118	1.13 (0.81-1.57)	77	1.05 (0.70-1.56)
Women – colon							
<19.6 (18.4)	7795	142	1.00 (ref.)	44	1.00 (ref.)	98	1.00 (ref.)
19.6-21.2 (20.5)	7731	143	1.02 (0.78-1.34)	41	0.94 (0.60-1.47)	102	1.06 (0.78-1.45)
21.3-23.0 (22.0)	7964	147	1.02 (0.78-1.34)	44	0.98 (0.63-1.52)	103	1.04 (0.76-1.42)
>23.0 (24.2)	7683	137	1.01 (0.77-1.33)	29	0.67 (0.41-1.09)	108	1.17 (0.86-1.59)
p-trend			0.940		0.153		0.371
per 5 kg/m²	31173	569	1.04 (0.89-1.21)	158	0.89 (0.70-1.14)	411	1.10 (0.92-1.31)
Women – rectum							
<19.6 (18.4)	7795	23	1.00 (ref.)	10	1.00 (ref.)	13	1.00 (ref.)
19.6-21.2 (20.5)	7731	37	1.64 (0.96-2.81)	20	2.04 (0.94-4.41)	17	1.34 (0.64-2.79)
21.3-23.0 (22.0)	7964	29	1.24 (0.71-2.18)	15	1.47 (0.65-3.31)	14	1.06 (0.49-2.28)
>23.0 (24.2)	7683	30	1.34 (0.77-2.35)	14	1.42 (0.62-3.25)	16	1.28 (0.61-2.70)
p-trend			0.554		0.658		0.681
per 5 kg/m²	31173	119	1.10 (0.85-1.44)	59	1.07 (0.74-1.54)	60	1.14 (0.79-1.64)

(continued)	Person years at risk	*KRAS*$_{mut}$		*PIK3CA*$_{mut}$[d]		*BRAF*$_{mut}$[d]		dMMR[d]	
		n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)	n cases	HR (95% CI)
Quartiles of BMI at age 20 years (kg/m²): range (median)									
Men – colon									
<20.2 (19.2)	6103	50	1.00 (ref.)	32	1.00 (ref.)	19	1.00 (ref.)	19	1.00 (ref.)
20.2-21.6 (21.0)	6458	50	0.96 (0.63-1.46)	33	0.97 (0.59-1.62)	25	1.27 (0.69-2.35)	16	0.84 (0.42-1.67)
21.7-23.3 (22.4)	6308	57	1.10 (0.73-1.76)	23	0.69 (0.40-1.21)	19	0.97 (0.50-1.87)	9	0.47 (0.21-1.06)
>23.3 (24.3)	6012	57	1.20 (0.79-1.80)	35	1.13 (0.68-1.86)	21	1.16 (0.61-2.20)	15	0.86 (0.43-1.72)
p-trend			0.311		0.929		0.883		0.402
per 5 kg/m²	24881	214	1.27 (0.93-1.72)	123	1.06 (0.72-1.55)	84	0.91 (0.60-1.39)	59	0.69 (0.41-1.17)
Men – rectum									
<20.2 (19.2)	6103	11	1.00 (ref.)						
20.2-21.6 (21.0)	6458	24	2.09 (1.01-4.32)						
21.7-23.3 (22.4)	6308	12	1.06 (0.46-2.43)						
>23.3 (24.3)	6012	12	1.14 (0.50-2.62)						
p-trend			0.664						
per 5 kg/m²	24881	59	1.11 (0.71-1.74)						
Women – colon									
<19.6 (18.4)	7795	46	1.00 (ref.)	25	1.00 (ref.)	37	1.00 (ref.)	25	1.00 (ref.)
19.6-21.2 (20.5)	7731	52	1.15 (0.76-1.76)	29	1.18 (0.68-2.06)	35	0.97 (0.59-1.57)	28	1.14 (0.65-2.00)
21.3-23.0 (22.0)	7964	49	1.05 (0.69-1.61)	20	0.79 (0.43-1.44)	43	1.15 (0.73-1.83)	30	1.19 (0.69-2.06)
>23.0 (24.2)	7683	56	1.28 (0.85-1.94)	31	1.30 (0.75-2.24)	41	1.20 (0.75-1.91)	30	1.28 (0.74-2.23)
p-trend			0.330		0.624		0.345		0.370
per 5 kg/m²	31173	203	1.12 (0.88-1.42)	105	1.18 (0.87-1.59)	156	1.21 (0.93-1.56)	113	1.27 (0.94-1.72)
Women – rectum									
<19.6 (18.4)	7795	11	1.00 (ref.)						
19.6-21.2 (20.5)	7731	13	1.21 (0.54-2.73)						
21.3-23.0 (22.0)	7964	12	1.07 (0.47-2.45)						
>23.0 (24.2)	7683	14	1.31 (0.59-2.93)						
p-trend			0.591						
per 5 kg/m²	31173	50	1.08 (0.72-1.63)						

Abbreviations: HR, hazard ratio; CI, confidence interval; BMI, body mass index; CRC, colorectal cancer; NLCS, Netherlands Cohort Study; (d/p)MMR, mismatch repair (deficient/proficient); p-het, p-heterogeneity.

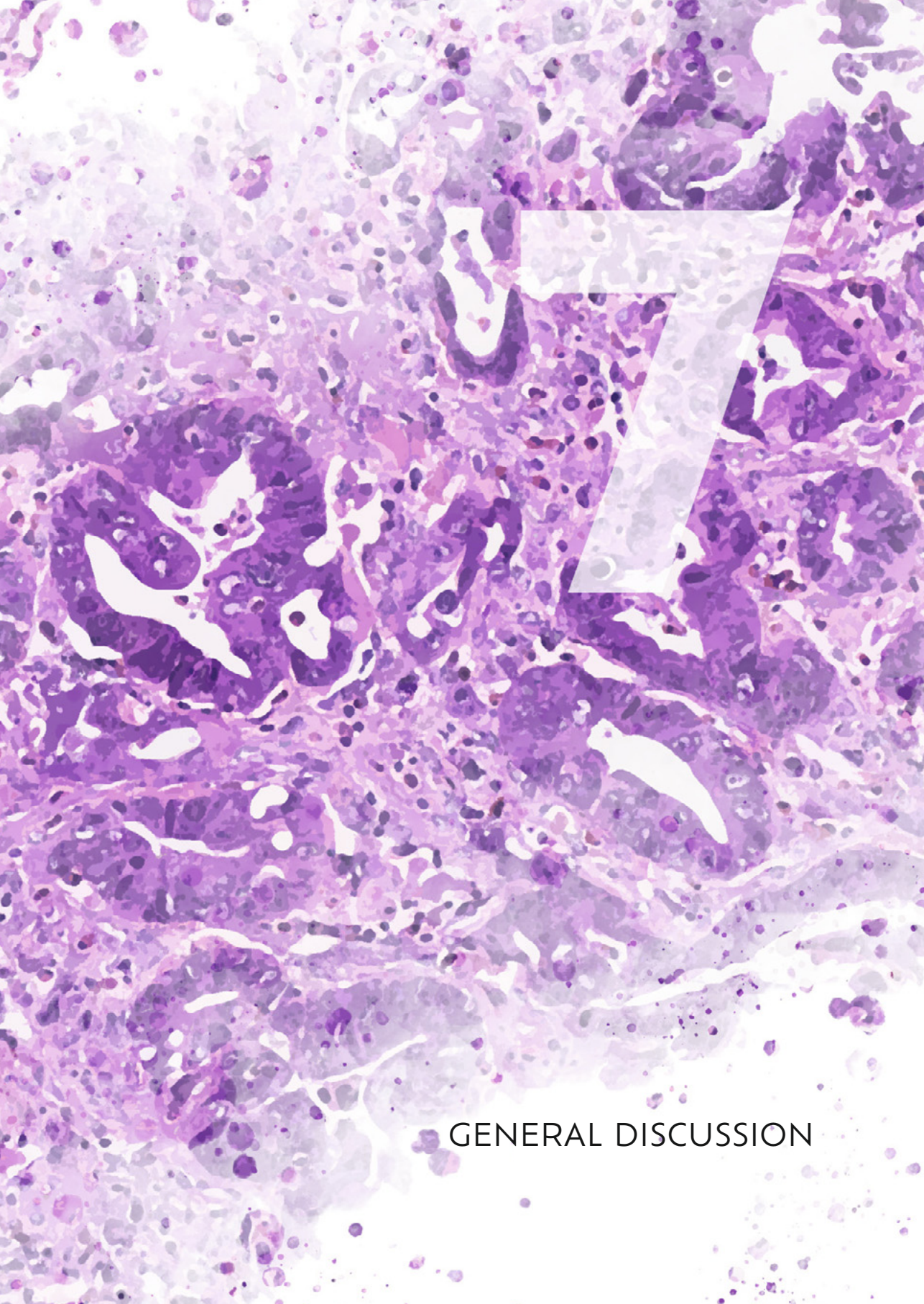
Note: p-heterogeneity tests for individual molecular features (*KRAS*$_{mut}$, *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, and dMMR) were not statistically significant (reference group for all tests: wild-type for *KRAS*, *PIK3CA*, and *BRAF*, and pMMR).

[a]Hazard Ratios were adjusted for age (years; continuous), height (cm; continuous), total energy intake (kcal/day; continuous), family history of CRC (yes/no), alcohol consumption (0; 0.1-4; 5-14; >15 g/day), non-occupational physical activity (minutes/day; continuous), processed meat intake (g/day; continuous), red meat intake (g/day; continuous). Age was included as a time-varying covariate.

[b]This group excludes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*), as well as MMR-deficient cases.

[c]This group includes cases with mutations in any of the genes (*KRAS*, *BRAF*, or *PIK3CA*) and/or cases that are MMR-deficient.

[d]Analyses for subgroups with <50 cases were not performed.



7

GENERAL DISCUSSION

Energy balance-related factors, like body mass index (BMI), physical activity, height, and energy restriction, have been associated with risk of colorectal cancer, but the mechanism behind these associations is currently unknown. Activation of the Warburg-effect (i.e. upregulated glycolysis under aerobic circumstances) via PI3K/Akt signaling is one of the proposed mechanisms. The aim of this thesis was to investigate potential involvement of the Warburg-effect in the etiological pathway between energy balance and colorectal cancer risk using a molecular pathological epidemiology (MPE) approach. We investigated associations of energy balance-related factors with colorectal cancer risk in relation to subtypes based on the estimated presence of the Warburg-effect in the tumor cells. The presence of the Warburg-effect was estimated by establishing Warburg-subtypes based on immunohistochemical (IHC) expression of six proteins involved in different levels of the Warburg-effect (LDHA, GLUT1, MCT4, PKM2, P53, PTEN). Furthermore, we investigated subgroups of colorectal cancer based on mutation status in oncogenes (*KRAS*, *PIK3CA*, and *BRAF*) which have been reported to be involved in the upstream regulation of the Warburg-effect, as well as based on mismatch repair (MMR) status, as a surrogate marker of microsatellite (in)stability (MSI/MSS). These molecular features were investigated individually (*KRAS* mutations [$KRAS_{mut}$]; $PIK3CA_{mut}$; $BRAF_{mut}$; MMR deficiency [dMMR]) as well as combined (all-wild-type+MMR-proficient [pMMR]; any-mutation/dMMR). These subgroups will hereafter be referred to as subgroups of molecular features. In addition, we investigated whether non-pathologists can generate valid and reproducible IHC scoring results.

SUMMARY OF MAIN FINDINGS

In Chapter 2, we showed that adequately trained non-pathologists were able to generate reproducible IHC scoring results that are similar to those of an experienced pathologist. Results from Chapters 3 and 5 are visually summarized in Figure 1 of this chapter (adult energy balance-related factors) and results from Chapters 4 and 6 in Figure 2 of this chapter (early-life energy balance-related factors). In these figures, a red box indicates a positive association between exposure and outcome, a green box an inverse association, and a white box indicates that no (clear) association was found.

In Chapter 3, we found that measures of adiposity (i.e. BMI and clothing size) were associated with an increased risk of Warburg-moderate and Warburg-high colon cancer in men, and with Warburg-low and Warburg-high colon cancer in women (Figure 1 – Warburg-subtypes). Furthermore, we observed that measures of physical activity were mainly associated with a decreased risk of Warburg-low and Warburg-moderate colon cancer, both in men and women. In Chapter 4, height was positively associated with colon cancer in men, regardless of Warburg-subtypes, and with Warburg-low colon and Warburg-low and -moderate rectal cancer in women (Figure 2 – Warburg-subtypes). Results on early-life energy restriction proxies (i.e. exposure to the Dutch Hunger Winter,

Figure 1 | Overview of associations between adult energy balance-related factors and risk of overall colorectal cancer, Warburg-subtypes, and subgroups of molecular features, separately for sex and tumor location. Red: positive association (increased risk); green: inverse association (decreased risk); white: no (clear) association observed (including weak and non-linear associations). A star (*) indicates that the association was statistically significant (this can be based on a significant effect estimate and/or on a significant p-trend). Analyses for subgroups with <50 cases were not performed (indicated by N.A.). Note: this figure aims to give a simplified overview of associations; cut-off points may differ according to the exposure. *Abbreviations: IHC, immunohistochemistry; mod, moderate; WT+pMMR, all-wild-type and mismatch repair proficient; Mut/dMMR, any-mutation and/or mismatch repair deficient; mut, mutated; dMMR, mismatch repair deficient; N.A., not available.*

Figure 2 | Overview of associations between early-life energy balance-related factors and risk of overall colorectal cancer, Warburg-subtypes, and subgroups of molecular features, separately for sex and tumor location. Red: positive association (increased risk); green: inverse association (decreased risk); white: no (clear) association observed (including weak and non-linear associations). A star (*) indicates that the association was statistically significant (this can be based on a significant effect estimate and/or on a significant p-trend). Analyses for subgroups with <50 cases were not performed (indicated by N.A.). Note: this figure aims to give a simplified overview of associations; cut-off points may differ according to the exposure. *Abbreviations: IHC, immunohistochemistry; mod, moderate; WT+pMMR, all-wild-type and mismatch repair (MMR) proficient; Mut/dMMR, any-mutation and/or MMR deficient; mut, mutated; dMMR, MMR deficient; N.A., not available.*

Second World War, Economic Depression) did not show clear patterns for associations with Warburg-subtypes. A high adolescent BMI was associated with an increased risk of Warburg-high colon cancer in men, and Warburg-moderate rectal cancer in women. In Chapter 5, we found that adiposity measures in women were only associated with an increased risk of $KRAS_{mut}$ colon cancer, whereas in men they were associated with all subgroups of molecular features (Figure 1 – subgroups of molecular features). Furthermore, we observed that non-occupational physical activity was associated with any-mutation/MMR-deficient (dMMR) colon cancer in both men and women, but not with all-wild-type+pMMR colon cancer. Finally, in Chapter 6, we found that height was associated with an increased risk of any-mutation/dMMR, but not all-wild-type+pMMR, colorectal cancer, with the exception of men with rectal cancer (Figure 2 – subgroups of molecular features). Again, results on early-life energy restriction proxies in relation to risk of subgroups based on molecular features did not show clear patterns. Lastly, a high adolescent BMI in men seemed to be mainly associated with an increased risk of $KRAS_{mut}$ colon cancer, and in women with $BRAF_{mut}$ and dMMR colon cancer.

INTERPRETATION OF STUDY RESULTS

Associations between energy balance-related factors and risk of colorectal cancer varied by Warburg-subtypes, by subgroups of molecular features, as well as by sex and tumor location. Since we were the first to investigate associations of energy balance-related factors with risk of Warburg-subtypes in colorectal cancer, a direct comparison of our results with the literature is currently not possible. We were also the first to combine cases into subgroups based on a combination of molecular features. Most of the molecular features we investigated (*KRAS* and *BRAF* mutations and MMR status) have been studied individually previously, but these either showed contrasting results or are currently very limited with respect to some energy balance-related factors[1-12]. The overall associations of energy balance-related factors with colon and rectal cancer risk reported in this thesis are largely in line with the existing literature[13-24]. In the next sections, we will compare the two subtyping approaches that were applied in this thesis, describe our reasoning regarding indicated involvement of the Warburg-effect based on these approaches, and interpret our results in light of the existing literature.

WARBURG-SUBTYPES VERSUS SUBGROUPS OF MOLECULAR FEATURES

We decided to use two approaches for classifying colorectal cancer cases in order to estimate the presence of the Warburg-effect in the tumor. First, we made subtypes based on IHC expression of proteins involved in the Warburg-effect from different levels of the pathway (upstream regulation of the Warburg-effect: PTEN, P53; glucose import: GLUT1; glycolysis: PKM2; conversion of pyruvate into lactate: LDHA; lactate secretion: MCT4). Second, we made subgroups based on molecular features reported to be involved in the upstream regulation of the Warburg-effect ($KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$) as well as on MMR status.

Using two subtyping approaches allowed us to cover different aspects of the Warburg-effect. Whilst the Warburg-subtypes based on IHC protein expression levels provide a more direct view of the Warburg-effect by covering multiple levels of the pathway, the subgroups of molecular features focused on the upstream regulation of the Warburg-effect. These Warburg-effect drivers are oncogenes that are involved in other cellular carcinogenic processes as well[25-27]. Thus, presence of mutations in these oncogenes might not necessarily indicate presence of the Warburg-effect. There are multiple steps between upregulated RAS activity and increased glycolysis, for example, if tumor suppressors P53 and PTEN remain intact, they might block the RAS signaling towards increased glycolysis[28-30].

We reasoned that variation in associations across Warburg-subtypes or across subgroups of molecular features could provide an indication for the involvement of the Warburg-effect in etiological pathways between energy balance-related factors and colorectal cancer. First, we considered that associations with specifically Warburg-high colorectal cancer, or in combination with Warburg-moderate, suggest likely involvement of the Warburg-effect in the etiological pathway between the exposure of interest and colorectal cancer. Second, we considered that associations with subgroups of individual molecular features (*KRAS*$_{mut}$, *PIK3CA*$_{mut}$, *BRAF*$_{mut}$, or dMMR) and/or with the any-mutation/dMMR subgroup, but not the all-wild-type+pMMR subgroup, suggest likely involvement of the Warburg-effect in the etiological pathway between the exposure of interest and colorectal cancer.

In the following sections, we will first interpret the results of each subtyping approach individually and then give a combined conclusion based on results from both approaches. Because the Warburg-subtypes based on IHC protein expression provide a more direct way of measuring the estimated presence of the Warburg-effect, this was considered the primary subtyping approach for interpretation of Warburg-effect involvement. If specific associations with any-mutation/dMMR or individual molecular features were observed, which were not in line with associations observed for Warburg-subtypes, this was considered to probably indicate that mechanisms other than the Warburg-effect are involved.

ADIPOSITY

Based on findings for the IHC based Warburg-subtypes, involvement of the Warburg-effect was indicated for associations of adult adiposity with risk of colon cancer in men and women (Figure 1 – Warburg-subtypes), and of adolescent adiposity with risk of colon cancer in men (Figure 2 – Warburg-subtypes). For associations of adult adiposity in women, an additional mechanism involving sex hormones has been proposed previously by Gunter and colleagues[31], which might explain the association we observed with Warburg-low colon cancer. Considering subgroups of molecular features, involvement of the Warburg-effect was indicated for associations of adult adiposity with risk of colon cancer in women (Figure 1 – subgroups of molecular features). In addition, our results suggest a potential role for the involvement of the Warburg-effect in associations of adolescent BMI and risk of colon cancer in men and women

(Figure 2 – subgroups of molecular features). Thus, the results are largely, but not entirely, consistent between both subtyping approaches as well as between adiposity at both time points in life. Altogether, the results presented in this thesis seem to suggest a role for the Warburg-effect in the etiological pathway of adolescent and adult adiposity and colon but not rectal cancer.

As described in Chapter 5, previous studies have investigated the association between adult adiposity and risk of colorectal cancer in relation to individual molecular features (*KRAS* or *BRAF* mutations or MMR/MSI status)[1-12]. In line with the current results, two studies reported stronger associations between adult adiposity and $KRAS_{mut}$ compared to *KRAS* wild-type ($KRAS_{wt}$) colorectal cancer, of which one study reported this difference in women[4] and the other in men[3]. In contrast, two other studies reported stronger associations for $KRAS_{wt}$ compared to $KRAS_{mut}$ [1, 2]. For *BRAF* mutation status, our results confirm previous literature[1-3, 5, 10], reporting weaker associations for $BRAF_{mut}$ versus $BRAF_{wt}$ colorectal cancer, with the exception of one study which reported the opposite difference in women (i.e. associations with $BRAF_{mut}$ stronger than $BRAF_{wt}$)[1]. Regarding MMR/MSI status, our results are in line with a recent meta-analysis by Carr et al[9], reporting no difference in associations between adult adiposity and risk of colorectal cancer according to MSI status. As described in Chapter 6, literature on adolescent adiposity and risk of colorectal cancer in relation to these molecular features is currently very limited. Our results regarding MMR status are not in line with those from a case-control study[6], which showed a stronger association with MSI low colorectal cancer. Lastly, we were the first to investigate the associations between adiposity and colorectal cancer risk in relation to *PIK3CA* mutation status, and therefore cannot compare our results to previous literature. The different results across MPE studies regarding associations of adiposity with risk of colorectal cancer according to molecular features might be attributed to several factors such as: use of different methods for assessing molecular features; different timing and method of exposure measurements; different study designs; different approaches to (outcome) stratification (for example stratification on sex and tumor location); and/or chance findings due to multiple testing, caused by repeatedly splitting colorectal cancer into different molecular pathological subgroups.

PHYSICAL ACTIVITY

Based on findings for the IHC based Warburg-subtypes, it seems that associations between physical activity and risk of colon cancer in men and women cannot be explained by the Warburg-effect (Figure 1 – Warburg-subtypes). Similarly, subgroups of molecular features did not suggest involvement of the Warburg-effect in associations between physical activity and colon cancer risk in men (Figure 1 – subgroups of molecular features). However, in women, subgroups of molecular features did suggest involvement of the Warburg-effect in associations of physical activity with risk of colon cancer. In addition, involvement of the Warburg-effect was indicated by both subtyping approaches in associations between physical activity and increased risk of rectal cancer in men (Figure 1), we currently do not have an explanation for this counterintuitive finding. Results are thus similar between both approaches of subtyping, with the exception of the indicated Warburg-effect involvement for women based on subgroups

of molecular features. Overall, the results presented in this thesis did not indicate a role for the Warburg-effect in associations of physical activity with colorectal cancer risk.

To the best of our knowledge, associations between physical activity and risk of colon cancer in relation to molecular features have only been investigated in a case-control study by Slattery et al for $KRAS_{mut}$ [4], $BRAF_{mut}$ [10], and MSI[8] status. In contrast to our results, they observed stronger associations between physical activity and $KRAS_{mut}$ compared to $KRAS_{wt}$ colon cancer in men[4]. However, similar to our results, associations did not differ according to *KRAS* mutation status in women[4]. For *BRAF*, their results are partly in line with ours, reporting that physical activity was not associated with $BRAF_{mut}$ colon cancer[10]. Lastly, regarding MMR/MSI status, our results are in contrast to those reported by Slattery et al[8], showing no difference in associations according to MSI status in men, but an association between physical activity and MSS colon cancer only in women. As there are no previous studies on physical activity and risk of colorectal cancer in relation to *PIK3CA* status, we cannot compare our results to the literature.

HEIGHT

Based on results for IHC based Warburg-subtypes, the Warburg-effect does not seem to be involved in associations of adult-attained height with colon or rectal cancer risk neither in men nor in women (Figure 2 – Warburg-subtypes). In contrast, subgroups of molecular features did indicate involvement of the Warburg-effect in associations of height with colon and rectal cancer risk, both in men and in women (Figure 2 – subgroups of molecular features). It is interesting that, even though the two subtyping approaches give a contrasting indication of Warburg-effect involvement, the results are similar across sex and tumor location, which was often not the case for other studied exposures. Overall, a role for the Warburg-effect was not indicated for associations of adult-attained height with colorectal cancer risk by the results presented in this thesis.

Previous studies on the association between height and risk of colorectal cancer in relation to molecular features studied in this thesis are currently very limited, as described in Chapter 6. Our results are in line with those of a pooled analysis of the Netherlands Cohort Study (NLCS), using 7.3 years of follow-up, and the Melbourne Collaborative Cohort study, which showed stronger associations between height and $BRAF_{mut}$ compared to $BRAF_{wt}$ colorectal cancer in both cohorts, and stronger associations for MSI compared to MSS colorectal cancer, again in both cohorts[5]. In contrast, our results on height are not in line with those of Brändstedt et al[3, 12], who did not observe clear (differences between) associations based on $KRAS_{mut}$, $BRAF_{mut}$, or MMR status in colorectal cancer, which could have been related to limited statistical power. Studies on height and risk of colorectal cancer in relation to *PIK3CA* mutations are currently lacking.

ENERGY RESTRICTION

The three proxy measures for energy restriction (exposure to the Dutch Hunger Winter; Second World War [WWII]; Economic Depression) did not show clear patterns with

Warburg-subtypes or subgroups of molecular features. Nevertheless, we will discuss the observed associations and potential explanations of differences across proxy measures.

Regarding exposure to the Dutch Hunger Winter, neither the Warburg-subtypes nor the subgroups of molecular features indicated a role for the Warburg-effect in the association between energy restriction and risk of colorectal cancer. In contrast, involvement of the Warburg-effect was indicated for the association between exposure to energy restriction during WWII and risk of colon cancer in men, based on both subtyping approaches. However, this was not the case for rectal cancer in men, nor for colon or rectal cancer in women. Lastly, based on Warburg-subtypes, involvement of the Warburg-effect is indicated for the positive association observed between energy restriction during the Dutch Economic Depression and risk of adult colon cancer in women. Subgroups of molecular features showed a similar pattern, but these associations did not reach statistical significance. Furthermore, this same pattern was observed for rectal cancer in women for both subtyping approaches, but again associations did not reach statistical significance.

We have described the differences between the used proxy measures for energy restriction in Chapter 4, which might explain why no clear pattern was observed across the three proxies. First, participants were older during the Hunger Winter (age range: 12-28 years) and WWII (age range: 8-28 years) than during the Economic Depression (age range: 0-23 years). Second, the severity of energy restriction differed across proxy measures. The place of residence during the Hunger Winter is used to identify cases with the most severe energy restriction equivalent to a caloric intake of 400-800 kcal per day at the height of famine[32-34]. The place of residence during WWII reflected a period of chronically impaired nutrition, with a different ratio between dietary nutrients for cities compared to rural areas (i.e. more carbohydrates and less fats in cities compared to rural areas)[34-36]. Lastly, even though the energy available in unemployed families was somewhat lower compared to employed families (3000 versus 3400 calories per day), a difference was observed especially in dietary variation and (micro)nutrient intake[34, 37]. Contrasting results have also been reported in previous studies on early-life energy restriction and risk of colorectal cancer[20-24]. This probably has to do with the timing, duration, and severity of energy restriction, which have been reported to influence the association between energy restriction and (colorectal) cancer risk[23, 24]. To the best of our knowledge, studies on (early-life) energy restriction and risk of colorectal cancer in relation to molecular features that were investigated in this thesis are currently lacking.

Overall, even though the differential associations observed for the different early-life energy restriction proxies might be explained by abovementioned factors, we deem it impossible to draw conclusions on indication of Warburg-effect involvement in associations of energy restriction with colorectal cancer. Nevertheless, our proxy measures individually show similar patterns for both subtyping approaches, which strengthens the indication for suggested presence of the Warburg-effect.

METHODOLOGICAL CONSIDERATIONS

THE NETHERLANDS COHORT STUDY

The NLCS has several strengths, including its large prospective cohort design with coverage throughout the Netherlands, its long and complete follow-up, and the availability of information on potential confounders. These assets reduce selection bias, recall bias, reverse causation, bias related to loss to follow-up, and confounding bias. The size of the cohort along with its long follow-up resulted in a large number of colorectal cancer cases, enabling investigation of associations between energy balance-related factors and risk of subgroups of colorectal cancer. A unique element of the NLCS is the opportunity to investigate exposure to (severe) early-life energy restriction in relation to adult colorectal cancer risk in a large prospective cohort.

A potential limitation is the use of exposure data obtained from a self-administered questionnaire at baseline only. Because of the absence of repeated measurements, potential exposure changes throughout follow-up are missed. However, the NLCS cohort that consists of an elderly Dutch population in 1986 is likely to have had relatively stable (dietary) habits[38]. Another consideration regarding the questionnaire is that it was filled out by the participants themselves, which might have led to information bias for anthropometry measures. However, the overall associations for height and BMI with colorectal cancer that we reported in this thesis were in line with those of studies where these anthropometric measures were assessed by trained personnel[13]. Furthermore, the usage of a one-time self-administered questionnaire enabled the inclusion of a large number of participants (n=120,852) from all over the Netherlands, thereby increasing the statistical power.

Such a large number of participants is essential for MPE research, as this type of research often involves many steps between study initiation and final analyses where cases might be excluded due to lack of tumor tissue, failing TMA quality control, incomplete confounder data, and other reasons. Through these steps, the sample size can be considerably reduced. Although the initial number of incident colorectal cancer cases (n=4,597) was considerably reduced after all exclusion steps, the total number of cases with complete molecular and exposure data was still very large (n ~ 1900) in the studies presented in this thesis. Especially for rectal cancer, however, the final statistical analyses in this thesis included relatively small groups ($n_{smallest\ group}$=37) which was also due to classification into subtypes and stratification on sex and tumor location. However, this stratification was considered essential to ensure sensible statistical analyses, further supported by the previously described (etiological) difference between colon and rectal cancer[39-42] and between men and women[13,41].

TISSUE MICROARRAYS AND IMMUNOHISTOCHEMISTRY

Since the introduction of tissue microarrays (TMAs) in 1998 by Kononen and colleagues[43], this method has been widely used as it allows for investigation of a large number of patients at reasonable costs and time. In addition, the usage of TMAs in large-scale studies is a powerful method to decrease run-to-run variability in IHC assays[44, 45].

However, the use of TMAs has some limitations. In the current study, formalin-fixed paraffin-embedded (FFPE) tissue blocks from different time periods and different hospitals were combined into the same TMA. It has been shown that differences in tissue fixation, storage time, and temperature of FFPE blocks can affect tissue quality, and thereby potentially influence IHC staining results[46-48]. These factors may vary over time and by hospital, whereby the quality of original FFPE tissue blocks may not be equivalent within TMAs. To minimize variation in IHC results based on follow-up time and/or hospital, we used internal controls whenever possible as IHC quality control measure to assess adequacy of the staining.

It has been reported that TMA cores are more prone to tissue loss during cutting, transferring onto glass slides, and IHC staining compared to whole tissue slides[49]. To prevent loss of complete cases during these steps, three cores per case were included in TMA blocks whenever possible. If a case had to be excluded because of insufficient available tumor tissue, this was unlikely to be related to either exposure or outcome, in other words, exclusion is expected to be non-differential. Furthermore, it might happen that during TMA construction, the tumor tissue is missed when punching a core from the whole tissue block. Therefore, we performed a strict quality control of all TMA blocks using an H&E stained section close to the tissue section used for IHC staining. Cores that were missing, lacking tumor tissue, or uninterpretable were excluded from all further analyses. Lastly, it is important that IHC staining protocols of antibodies, especially those that are not routinely used in the clinic, are sufficiently optimized. For the optimization, we initially used multi-tissue blocks (containing tissues from different organs) and whole tissue sections containing colorectal cancer tissue. It has been reported that TMA cores sometimes stain with a different intensity than whole tissue sections from the same tissue block[49]. Therefore, the IHC staining protocols were refined for TMAs specifically by using spare TMA sections including normal colon tissue and colorectal cancer tissue cores.

ESTIMATING PRESENCE OF THE WARBURG-EFFECT

While the classification of a disease into molecular subgroups in MPE research provides a unique opportunity of investigating potential mechanisms underlying etiological pathways, there is also risk of measurement errors and misclassification in outcome variables (i.e. the subtypes). In this section, we will discuss the issue of measurement error and misclassification of our outcome variables, as well as our approach to minimize this, separately for the Warburg-subtypes and for the subgroups of molecular features.

For Warburg-subtypes establishment (Chapters 3 and 4), we used a sum score of six proteins involved in the Warburg-effect (PTEN, P53, GLUT1, PKM2, LDHA, MCT4). Arguably, this approach did not include all proteins potentially involved in the Warburg-effect. However, we tried to make sure that proteins from different levels of the pathway were included in the IHC panel and subsequent sum score (i.e. from upstream regulators, glucose import, glycolysis, and lactate secretion). There were some proteins that we initially planned to include in our study, like HIF-1, MYC, and PDK1, because of

their well-known roles in the Warburg-effect[30]. However, these stainings did not pass IHC quality standards. Still, we believe that the Warburg-subtypes based on the expression of six proteins should be indicative of the presence or absence of the Warburg-effect in the tumor. In accordance with what we expected based on previous literature[50], we observed that Warburg-high colorectal cancer cases had a worse prognosis compared to Warburg-low and Warburg-moderate cases in a study not included in this thesis[51].

Our subgroups based on molecular features that have been associated with upstream regulation of the Warburg-effect (Chapters 5 and 6), $KRAS_{mut}$, $PIK3CA_{mut}$, $BRAF_{mut}$, and dMMR, were assessed by the existing ColoCarta panel (Agena Bioscience, Hamburg) for mutation status and by IHC of MLH1 and MSH2 for MMR status. It might have been interesting to include mutation status of other drivers of the Warburg-effect, like AKT1 or mTOR[30]. However, considering budgetary constraints due to the high number of cases, using the existing and validated ColoCarta panel appeared to be the best option. As described in Chapters 5 and 6, this panel covered the majority of known *KRAS* and *BRAF* mutations (99% and 98%, respectively), but only 78% of known *PIK3CA* mutations[52]. Therefore, misclassification might have occurred especially in *PIK3CA* mutation assessment. However, as the most common *PIK3CA* mutations are included in the panel[53], it seems unlikely that detection of additional cases with less common mutations would alter the current results. In addition, MSI status estimation by assessing MMR status using IHC expression of MLH1 and MSH2 might have led to misclassification of some of the cases. However, it has been shown that loss of MLH1 or MSH2 expression was observed in ~90% of MSI cases[54], which also probably did not alter our results since the number of additional cases would be rather small. The percentages of *KRAS*, *BRAF*, and *PIK3CA* mutations observed in our colorectal cancer cases were in line with the Catalogue Of Somatic Mutations In Cancer (COSMIC) database[55] and the percentage of dMMR cases, as an indication for MSI, was in line with previous literature[25].

FUTURE RESEARCH

The results presented in this thesis indicate potential involvement of the Warburg-effect in etiological associations between specific energy balance-related factors and colorectal cancer risk. We were the first to investigate a potential role of the Warburg-effect in associations between energy balance-related factors and colorectal cancer risk in a large prospective cohort. Therefore, replication of the current findings is required for validation. The approach that is least prone to bias in MPE studies is the prospective cohort design, as was used in the current thesis. However, this requires a large number of participants, long follow-up time, available funding, and availability of tumor material. Some examples of large prospective cohort studies with available tumor tissue that would have the potential to replicate (part of) the current results include the Health Professionals Follow-up Study[56], the Nurses' Health Study[57], the Iowa

Women's Health Study[58], the European Prospective Investigation into Cancer and Nutrition[59], and the Melbourne Collaborative Cohort Study[60]. However, the unique opportunity in the NLCS to investigate (severe) early-life energy restriction is often not available in existing cohorts.

In addition to replicating the current results, evaluating other methods to assess presence of the Warburg-effect and/or combining different methods would be of interest because our subgroups probably do not reflect all factors involved in the Warburg-effect. Examples could be: including Warburg-effect markers that we were not able to include in the current thesis (for example: HIF-1, MYC, AKT1, PDK1), metabolomics[61], microRNAs[62, 63], or epigenetics[62, 64]. Eventually, machine learning algorithms could be applied based on (a combination of) these methods.

In addition to validation of the current results in colorectal cancer cases, it would be of interest to study these associations in other cancer types. The proposed mechanism behind the association between energy balance and colorectal cancer risk that was studied in this thesis (i.e. Warburg-effect activation via PI3K/Akt through altered adipokine, insulin, and IGF-1 signaling) has also been proposed for breast cancer, hepatocellular carcinoma, and ovarian cancer[65].

Even though the results presented in this thesis provide further insights into potential involvement of the Warburg-effect in associations between energy balance and colorectal cancer risk, additional research is necessary to confirm a causal relationship. The previously described adenoma-carcinoma sequence (Chapter 1) of colorectal cancer provides a unique opportunity to distinguish molecular characteristics that occur early or later in the carcinogenic process[66, 67]. It would be of interest to repeat the current analyses in colorectal adenomas, which could provide additional evidence for a potential role of the Warburg-effect in carcinogenesis. Another way to further decipher the potentially causal relationship between energy balance-related factors, the Warburg-effect, and colorectal cancer would be to include pre-diagnostic serum levels of signaling molecules related to energy balance (i.e. adipokines, insulin, and IGF-1) in investigating the associations between energy balance-related factors and Warburg-subtypes of colorectal cancer.

Lastly, several other potential mechanisms have been proposed for the etiological pathway between energy balance and colorectal cancer. These include, amongst others, sex hormone signaling and obesity-related inflammation[68-70]. We believe it would be of interest to investigate these pathways as well, preferably in context to each other (including the proposed pathway studied in this thesis), in order to further unravel the mechanism(s) behind etiological pathway(s) between energy balance-related factors and colorectal cancer.

CONCLUDING REMARKS

The aim of this thesis was to investigate potential involvement of the Warburg-effect in the etiological pathway between energy balance and colorectal cancer risk using an MPE approach. We investigated associations of energy balance-related factors with Warburg-subtypes in colorectal cancer as well as with colorectal cancer subgroups based on molecular features that have been associated with the upstream regulation of the Warburg-effect. Overall, the results presented in this thesis seem to indicate a role for the Warburg-effect in the etiological pathway between adiposity and colon cancer, but not rectal cancer, both for adolescent and adult adiposity. A role for the Warburg-effect is not indicated by our results for associations of physical activity or adult-attained height with colorectal cancer risk. We did not observe clear patterns for the three energy restriction proxies and therefore do not draw a conclusion regarding early-life energy restriction. Since we were the first to investigate associations of energy balance-related factors with risk of colorectal cancer risk in relation to the estimated presence of the Warburg-effect in the tumor, confirmation in additional large prospective MPE studies is necessary.

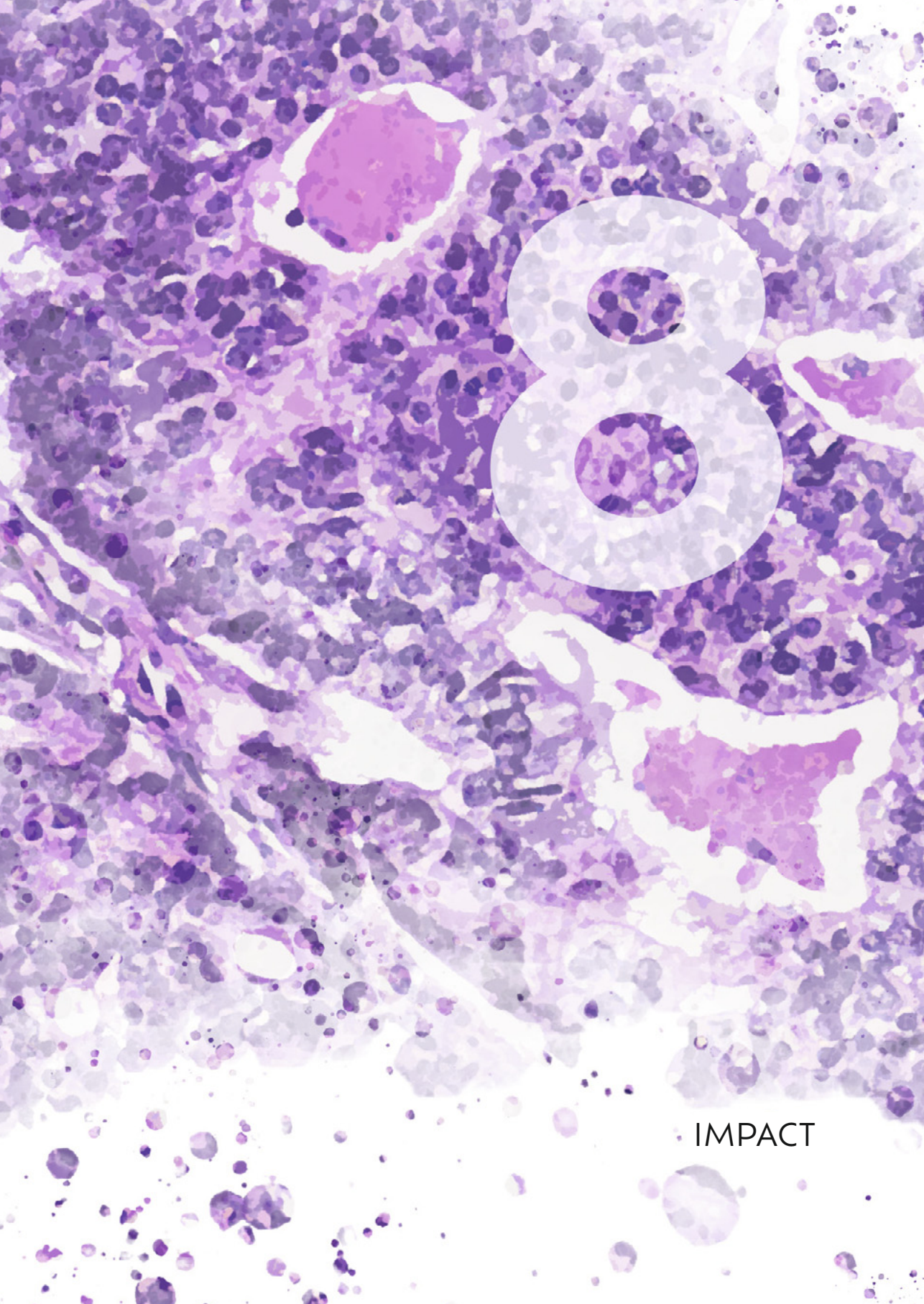
REFERENCES

1. Carr PR, Amitay EL, Jansen L, Alwers E, Roth W, Herpel E, Kloor M, Schneider M, Bläker H, Chang-Claude J. Association of BMI and major molecular pathological markers of colorectal cancer in men and women. The American Journal of Clinical Nutrition. 2020.
2. Myte R, Gylling B, Häggström J, Häggström C, Zingmark C, Löfgren Burström A, Palmqvist R, Van Guelpen B. Metabolic factors and the risk of colorectal cancer by KRAS and BRAF mutation status. Int J Cancer. 2019;145(2):327-37.
3. Brändstedt J, Wangefjord S, Nodin B, Eberhard J, Sundström M, Manjer J, Jirström K. 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(6):e98964.
4. Slattery ML, Anderson K, Curtin K, Ma K-N, Schaffer D, Edwards S, Samowitz W. Lifestyle factors and Ki-ras mutations in colon cancer tumors. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2001;483(1-2):73-81.
5. Hughes LA, Williamson EJ, van Engeland M, Jenkins MA, Giles GG, Hopper JL, Southey MC, Young JP, Buchanan DD, Walsh MD. Body size and risk for colorectal cancers showing BRAF mutations or microsatellite instability: a pooled analysis. International journal of epidemiology. 2012;41(4):1060-72.
6. Campbell PT, Jacobs ET, Ulrich CM, Figueiredo JC, Poynter JN, McLaughlin JR, Haile RW, Jacobs EJ, Newcomb PA, Potter JD. Case–control study of overweight, obesity, and colorectal cancer risk, overall and by tumor microsatellite instability status. Journal of the National Cancer Institute. 2010;102(6):391-400.
7. Hoffmeister M, Bläker H, Kloor M, Roth W, Toth C, Herpel E, Frank B, Schirmacher P, Chang-Claude J, Brenner H. Body mass index and microsatellite instability in colorectal cancer: a population-based study. Cancer Epidemiology and Prevention Biomarkers. 2013;22(12):2303-11.
8. Slattery ML, Curtin K, Anderson K, Ma K-N, Ballard L, Edwards S, Schaffer D, Potter J, Leppert M, Samowitz WS. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. Journal of the National Cancer Institute. 2000;92(22):1831-6.
9. Carr P, Alwers E, Bienert S, Weberpals J, Kloor M, Brenner H, Hoffmeister M. Lifestyle factors and risk of sporadic colorectal cancer by microsatellite instability status: a systematic review and meta-analyses. Annals of Oncology. 2018;29(4):825-34.
10. Slattery ML, Curtin K, Sweeney C, Levin TR, Potter J, Wolff RK, Albertsen H, Samowitz WS. Diet and lifestyle factor associations with CpG island methylator phenotype and BRAF mutations in colon cancer. International journal of cancer. 2007;120(3):656-63.
11. Hanyuda A, Ogino S, Qian ZR, Nishihara R, Song M, Mima K, Inamura K, Masugi Y, Wu K, Meyerhardt JA. Body mass index and risk of colorectal cancer according to tumor lymphocytic infiltrate. International journal of cancer. 2016;139(4):854-68.
12. Brändstedt J, Wangefjord S, Borgquist S, Nodin B, Eberhard J, Manjer J, Jirström K. Influence of anthropometric factors on tumour biological characteristics of colorectal cancer in men and women: a cohort study. Journal of translational medicine. 2013;11(1):1-13.
13. Abar L, Vieira AR, Aune D, Sobiecki JG, Vingeliene S, Polemiti E, Stevens C, Greenwood DC, Chan DS, Schlesinger S. Height and body fatness and colorectal cancer risk: an update of the WCRF–AICR systematic review of published prospective studies. European journal of nutrition.

2018;57(5):1701-20.
14. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. Cancer Epidemiology and Prevention Biomarkers. 2007;16(12):2533-47.
15. Garcia H, Song M. Early-life obesity and adulthood colorectal cancer risk: a meta-analysis. Revista Panamericana de Salud Pública. 2019;43:e3.
16. Robsahm TE, Aagnes B, Hjartåker A, Langseth H, Bray FI, Larsen IK. Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. European Journal of Cancer Prevention. 2013;22(6):492-505.
17. Wolin KY, Yan Y, Colditz GA, Lee I. Physical activity and colon cancer prevention: a meta-analysis. British journal of cancer. 2009;100(4):611-6.
18. Samad A, Taylor R, Marshall T, Chapman MA. A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. Colorectal disease. 2005;7(3):204-13.
19. Gunnell D, Okasha M, Davey Smith G, Oliver S, Sandhu J, Holly J. Height, leg length, and cancer risk: a systematic review. Epidemiologic reviews. 2001;23(2):313-42.
20. Hughes LA, van den Brandt PA, Goldbohm RA, De Goeij AF, De Bruine AP, Van Engeland M, Weijenberg MP. Childhood and adolescent energy restriction and subsequent colorectal cancer risk: results from the Netherlands Cohort Study. International journal of epidemiology. 2010;39(5):1333-44.
21. Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age–period–cohort analysis. European journal of cancer prevention. 2002;11(5):489-95.
22. Svensson E, Møller B, Tretli S, Barlow L, Engholm G, Pukkala E, Rahu M, Tryggvadottir L. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. Cancer Causes & Control. 2005;16(3):215-23.
23. Keinan-Boker L, Vin-Raviv N, Liphshitz I, Linn S, Barchana M. Cancer incidence in Israeli jewish survivors of World War II. Journal of the National Cancer Institute. 2009;101(21):1489-500.
24. Xie SH, Lagergren J. A possible link between famine exposure in early life and future risk of gastrointestinal cancers: Implications from age-period-cohort analysis. International journal of cancer. 2017;140(3):636-45.
25. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology. 2010;138(6):2073-87. e3.
26. Li W, Qiu T, Dong L, Zhang F, Guo L, Ying J. Prevalence and characteristics of PIK3CA mutation in mismatch repair-deficient colorectal cancer. Journal of Cancer. 2020;11(13):3827.
27. Haluska F, Pemberton T, Ibrahim N, Kalinsky K, editors. The RTK/RAS/BRAF/PI3K pathways in melanoma: biology, small molecule inhibitors, and potential applications. Seminars in oncology; 2007: Elsevier.
28. Bensaad K, Tsuruta A, Selak MA, Vidal MNC, Nakano K, Bartrons R, Gottlieb E, Vousden KH. TIGAR, a p53-inducible regulator of glycolysis and apoptosis. Cell. 2006;126(1):107-20.
29. Stambolic V, MacPherson D, Sas D, Lin Y, Snow B, Jang Y, Benchimol S, Mak T. Regulation of PTEN transcription by p53. Mol Cell. 2001;8(2):317-25.
30. Cairns RA, Harris IS, Mak TW. Regulation of cancer cell metabolism. Nature Reviews Cancer. 2011;11(2):85.
31. Gunter MJ, Hoover DR, Yu H, Wassertheil-Smoller S, Rohan TE, Manson JE, Howard BV, Wylie-Rosett J, Anderson GL, Ho GYJCr. Insulin, insulin-like growth factor-I, endogenous estradiol, and risk of colorectal cancer in postmenopausal

women. 2008;68(1):329-37.
32. Trienekens GM. Tussen ons volk en de honger: de voedselvoorziening, 1940-19451985.
33. Burger G, Drummond JC, Sandstead H. Malnutrition and starvation in western Netherlands: September 1944-July 1945: General State Print. Office.; 1948.
34. Dirx MJ, van den Brandt PA, Goldbohm RA, Lumey L. Energy restriction early in life and colon carcinoma risk. Cancer. 2003;97(1):46-55.
35. Breunis J. The food supply. The Annals of the American Academy of Political and Social Science. 1946;245(1):87-92.
36. Dols M, van Arcken D. De voedselvoorziening in Nederland tijdens en onmiddellijk na den tweeden wereldoorlog 1940-19451946.
37. Den Hartog A. Werklozen en hun voeding in de jaren 1930-1939. Voeding. 1983;44:92-8.
38. van den Brandt PA, Goldbohm RA, Veer PVT, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. Journal of clinical epidemiology. 1990;43(3):285-95.
39. Li M, Li J, Zhao A, Gu J. Colorectal cancer or colon and rectal cancer? Oncology. 2007;73(1-2):52-7.
40. Kapiteijn E, Liefers G, Los L, Klein Kranenbarg E, Hermans J, Tollenaar R, Moriya Y, Van De Velde C, Van Krieken J. Mechanisms of oncogenesis in colon versus rectal cancer. The Journal of pathology. 2001;195(2):171-8.
41. Wei EK, Giovannucci E, Wu K, Rosner B, Fuchs CS, Willett WC, Colditz GA. Comparison of risk factors for colon and rectal cancer. International journal of cancer. 2004;108(3):433-42.
42. Li F-y, Lai M-d. Colorectal cancer, one entity or three. Journal of Zhejiang University Science B. 2009;10(3):219-29.
43. Kononen J, Bubendorf L, Kallionimeni A, Bärlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallionimeni O-P. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nature medicine. 1998;4(7):844-7.
44. Sherman ME, Howatt W, Blows FM, Pharoah P, Hewitt SM, Garcia-Closas M. Molecular pathology in epidemiologic studies: a primer on key considerations. Cancer Epidemiology and Prevention Biomarkers. 2010;19(4):966-72.
45. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut. 2010:gut. 2010.217182.
46. Grillo F, Pigozzi S, Ceriolo P, Calamaro P, Fiocca R, Mastracci L. Factors affecting immunoreactivity in long-term storage of formalin-fixed paraffin-embedded tissue sections. Histochemistry and cell biology. 2015;144(1):93-9.
47. Werner M, Chott A, Fabiano A, Battifora H. Effect of formalin tissue fixation and processing on immunohistochemistry. The American journal of surgical pathology. 2000;24(7):1016-9.
48. Grillo F, Bruzzone M, Pigozzi S, Prosapio S, Migliora P, Fiocca R, Mastracci L. Immunohistochemistry on old archival paraffin blocks: is there an expiry date? Journal of Clinical Pathology. 2017;70(11):988-93.
49. Ilyas M, Grabsch H, Ellis IO, Womack C, Brown R, Berney D, Fennell D, Salto-Tellez M, Jenkins M, Landberg G, Byers R, Treanor D, Harrison D, Green AR, Ball G, Hamilton P, National Cancer Research Institute B, Imaging Clinical Studies G. Guidelines and considerations for conducting experiments using tissue microarrays. Histopathology. 2013;62(6):827-39.
50. Yu M, Chen S, Hong W, Gu Y, Huang B, Lin Y, Zhou Y, Jin H, Deng Y, Tu L, Hou B, Jian Z. Prognostic role of glycolysis for cancer outcome: evidence from 86 studies. Journal of Cancer Research and Clinical Oncology. 2019.
51. Offermans K, Jenniskens JCA, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ,

Weijenberg MP, Grabsch HI, van den Brandt PA. Expression of proteins associated with the Warburg-effect and survival in colorectal cancer. J Pathol Clin Res. 2021.
52. Fumagalli D, Gavin PG, Taniyama Y, Kim S-I, Choi H-J, Paik S, Pogue-Geile KL. A rapid, sensitive, reproducible and cost-effective method for mutation profiling of colon cancer and metastatic lymph nodes. BMC cancer. 2010;10(1):101.
53. Gray RT, Cantwell MM, Coleman HG, Loughrey MB, Bankhead P, McQuaid S, O'neill RF, Arthur K, Bingham V, McGready C. Evaluation of PTGS2 expression, PIK3CA mutation, aspirin use and colon cancer survival in a population-based cohort study. Clinical and translational gastroenterology. 2017;8(4):e91.
54. Lanza G, Gafà R, Maestri I, Santini A, Matteuzzi M, Cavazzini L. Immunohistochemical pattern of MLH1/MSH2 expression is related to clinical and pathological features in colorectal adenocarcinomas with microsatellite instability. Modern Pathology. 2002;15(7):741.
55. Catalogue Of Somatic Mutations In Cancer (COSMIC): large intestine carcinoma. [Available from: https://cancer.sanger.ac.uk/cosmic].
56. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. N Engl J Med. 2007;356(21):2131-42.
57. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. Nature Reviews Cancer. 2005;5(5):388-96.
58. Folsom AR, Kaye SA, Prineas RJ, Potter JD, Gapstur SM, Wallace RB. Increased incidence of carcinoma of the breast associated with abdominal adiposity in postmenopausal women. American journal of epidemiology. 1990;131(5):794-803.
59. Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol. 1997;26(suppl_1):S6.
60. Milne R, Fletcher A, MacInnis R, Hodge A, Hopkins A, Bassett J, Bruinsma F, Lynch B, Dugué P, Jayasekara H. Cohort profile: the Melbourne collaborative cohort study (health 2020). Int J Epidemiol. 2017;46(6):1757-i.
61. Chen X, Yi C, Yang M-J, Sun X, Liu X, Ma H, Li Y, Li H, Wang C, He Y. Metabolomics study reveals the potential evidence of metabolic reprogramming towards the Warburg effect in precancerous lesions. Journal of Cancer. 2021;12(5):1563.
62. Bensinger SJ, Christofk HR, editors. New aspects of the Warburg effect in cancer cell biology. Semin Cell Dev Biol; 2012: Elsevier.
63. Wai Hon K, Zainal Abidin SA, Othman I, Naidu R. Insights into the role of microRNAs in colorectal cancer (CRC) metabolism. Cancers. 2020;12(9):2462.
64. Oermann EK, Wu J, Guan K-L, Xiong Y, editors. Alterations of metabolic genes and metabolites in cancer. Semin Cell Dev Biol; 2012: Elsevier.
65. Park J, Morley TS, Kim M, Clegg DJ, Scherer PE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. Nature Reviews Endocrinology. 2014;10(8):455.
66. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.
67. Cho KR, Vogelstein B. Genetic alterations in the adenoma–carcinoma sequence. Cancer. 1992;70(S4):1727-31.
68. Park J, Morley TS, Kim M, Clegg DJ, Scherer PEJNRE. Obesity and cancer—mechanisms underlying tumour progression and recurrence. 2014;10(8):455.
69. Hopkins BD, Goncalves MD, Cantley LCJJoco. Obesity and cancer mechanisms: cancer metabolism. 2016;34(35):4277.
70. Friedenreich CM, Ryder-Burbidge C, McNeil JJMO. Physical activity, obesity and sedentary behavior in cancer etiology: Epidemiologic evidence and biologic mechanisms. 2020.


8
IMPACT

This section addresses the relevance and the potential scientific and social impact of the research presented in this thesis.

RELEVANCE

Colorectal cancer has been amongst the five most common cancers worldwide for decades[1-4], with an estimated 1.9 million new cases in 2020[4]. It has been estimated that more than 50% of colorectal cancers are attributable to modifiable risk factors and thus potentially preventable[5]. Organizations like the World Cancer Research Fund (WCRF) and the International Agency for Research on Cancer (IARC) regularly publish updated reports on scientific evidence for cancer prevention[6, 7]. Both reports indicate convincing evidence for associations of adiposity, height, and physical activity with colorectal cancer risk. However, while many studies reported associations between these energy balance-related factors and colorectal cancer, associations often appear to be weak and are not always consistent between studies[8]. This might be caused by disease heterogeneity, i.e. when exposure factors are differentially associated with subgroups of a disease[9]. This issue is being addressed by using a molecular pathological epidemiology (MPE) approach, as was done in the current thesis.

MPE is an emerging transdisciplinary field incorporating molecular pathology into epidemiological research[10]. Classification of (colorectal) cancers into subgroups based on specific molecular characteristics addresses disease heterogeneity, whereby weak or masked associations can be revealed. In addition, MPE research can strengthen the evidence for causal relationships by providing further insights into etiology and pathogenesis of a disease. In the current thesis, we investigated whether energy balance-related factors (i.e. adiposity, physical activity, height, and energy restriction) are differentially associated with risk of colorectal cancer subgroups based on estimated presence of the Warburg-effect.

After the discovery of the Warburg-effect (i.e. increased glycolysis under aerobic conditions) in the 1920s[11], a long period passed with lack of interest in this phenomenon. In recent decades, however, the Warburg-effect regained interest in the scientific community. We were the first to investigate whether the Warburg-effect might be involved in the etiological pathway between energy balance-related factors and colorectal cancer. The results presented in this thesis indicate involvement of the Warburg-effect in the etiological pathway of adiposity with colon cancer. The etiological pathway of physical activity and adult-attained height are probably explained by mechanisms other than the Warburg-effect. No clear patterns were observed for the three energy restriction proxies. Since we were the first to investigate these associations, confirmation in additional large prospective MPE studies is required. Nevertheless, the suggested involvement of the Warburg-effect in the association between adiposity and colon cancer contributes additional insights

into the underlying mechanisms of this etiological pathway. A better understanding of how exposures (such as energy balance-related factors) affect disease initiation may ultimately improve preventive strategies. In addition, in a study that was not included in the current thesis, we observed that colorectal cancer cases with estimated presence of the Warburg-effect had a worse survival compared to cases without[12]. Therefore, this new way of colorectal cancer subtyping based on estimated presence of the Warburg-effect seems to have both etiological as prognostic significance.

More generally, the results presented in this thesis support the evidence that energy balance-related factors like adiposity, height, and physical activity are associated with colorectal cancer risk. In addition, our results revealed stronger associations for specific subgroups of colorectal cancer, further demonstrating how important these factors are in the prevention of colorectal cancer. Over the past decades, overweight and obesity has been rising amongst adults as well as children and adolescents. Among Dutch adults (age ≥20 years), overweight has increased from 35.3% in 2001 to 36.8% in 2020, and obesity from 9.6% to 14.2%[13]. In Dutch children and adolescents (age 4-19 years), overweight has increased from 9.2% in 2001 to 12.5% in 2020, whereas the percentage of obese children and adolescents remained the same (2.5%)[13]. These numbers are alarming and highlight the need of prevention with a focus on a healthy energy balance throughout life.

KNOWLEDGE TRANSFER

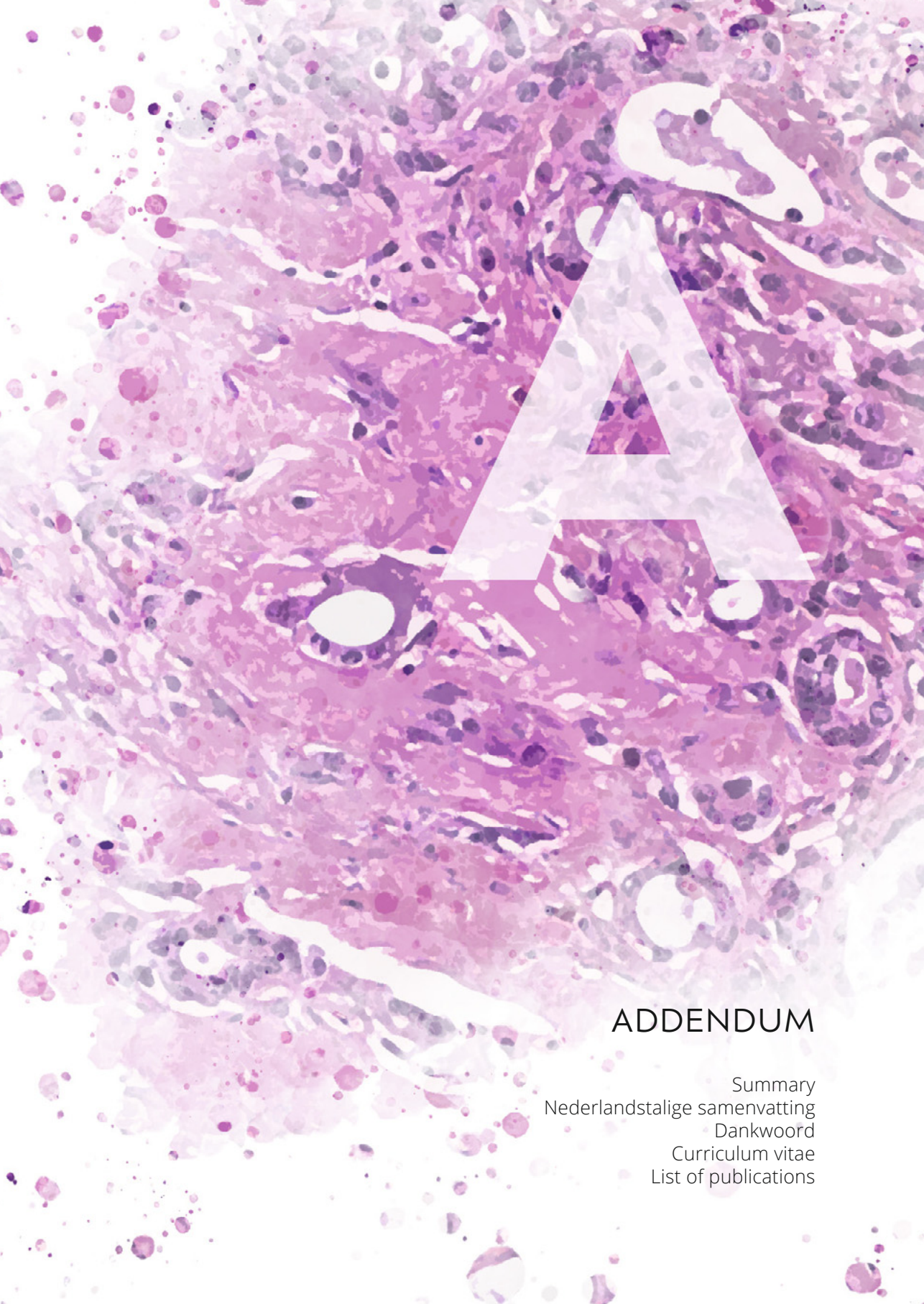
The scientific knowledge presented in this thesis has been shared with fellow researchers through publication in several international scientific journals. In addition, our results have been presented at various conferences and symposia for different audiences. Our results have been presented online at the Virtual Annual Meeting (2021) of the American Association for Cancer Research (AACR), which is a large international conference with a broad audience of cancer researchers. Furthermore, our results have been presented online at the 13th Joint Meeting of the British Division of the International Academy of Pathology and the Pathological Society of Great Britain & Ireland (Manchester Pathology 2021) and online at the Dutch Epidemiological conference (WEON 2021). The audience of these conferences mainly consists of pathologists and epidemiologists, respectively. Lastly, our results were presented live at the Science Day of the Maastricht University Medical Centre+, of which the audience consisted of a broad audience of researchers and clinicians.

CONCLUSION

In this thesis, we investigated associations between energy balance-related factors and colorectal cancer risk in relation to subtypes based on the estimated presence of the Warburg-effect in the tumor. Our results underline the importance of preventive strategies aimed at reaching and/or maintaining a healthy energy balance throughout life. Since we were the first to investigate these associations in relation to the Warburg-effect, confirmation of the results presented in this thesis is necessary. Nevertheless, these results provide additional insights into underlying mechanisms of the etiological pathway between energy balance-related factors and colorectal cancer risk, which may ultimately improve preventive strategies.

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer. 2010;127(12):2893-917.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015;65(2):87-108.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
4. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2021;71(3):209-49.
5. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, McCullough ML, Patel AV, Ma J, Soerjomataram I. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. CA: a cancer journal for clinicians. 2018;68(1):31-54.
6. World Cancer Research Fund/American Institute for Cancer Research. Diet, nutrition, physical activity and cancer: a global perspective. Continuous Update Project Expert Report 2018.
7. Wild C, Weiderpass E, Stewart B. World cancer report 2020. Lyon: International Agency for Research on Cancer. 2020;199:512.
8. Abar L, Vieira AR, Aune D, Sobiecki JG, Vingeliene S, Polemiti E, Stevens C, Greenwood DC, Chan DS, Schlesinger S. Height and body fatness and colorectal cancer risk: an update of the WCRF–AICR systematic review of published prospective studies. European journal of nutrition. 2018;57(5):1701-20.
9. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut. 2010:gut. 2010.217182.
10. Ogino S, Stampfer M. Lifestyle factors and microsatellite instability in colorectal cancer: the evolving field of molecular pathological epidemiology. Oxford University Press; 2010. p. 365-7.
11. Warburg O. The metabolism of carcinoma cells. The Journal of Cancer Research. 1925;9(1):148-63.
12. Offermans K, Jenniskens JCA, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Expression of proteins associated with the Warburg-effect and survival in colorectal cancer. The Journal of Pathology: Clinical Research. 2022;8(2):169-80.
13. Centraal Bureau voor de Statistiek (2021). Gezondheid, leefstijl, zorggebruik en -aanbod, doodsoorzaken; kerncijfers. [Available from: https://www.cbs.nl/nl-nl/cijfers/detail/81628NED?q=overgewicht].



ADDENDUM

SUMMARY

Energy balance-related factors have been associated with risk of colorectal cancer. Even though the mechanism behind these associations is currently unknown, Warburg-effect activation via PI3K/Akt signaling is one of the proposed mechanisms. The Warburg-effect is a metabolic phenotype characterized by increased aerobic glycolysis and is considered an important step in carcinogenesis.

In this thesis, we investigated potential involvement of the Warburg-effect in the etiological pathway between energy balance and colorectal cancer risk using a molecular pathological epidemiology (MPE) approach. We investigated associations of energy balance-related factors in adulthood (i.e. body mass index [BMI]; lower body clothing size, as a proxy for waist circumference; physical activity) and early in life (i.e. height; energy restriction proxies of exposure to the Dutch Hunger Winter, World War II, and the Dutch Economic Depression; BMI at age 20 years) with colorectal cancer risk in relation to the estimated presence of the Warburg-effect in the tumor. The presence of the Warburg-effect was estimated by establishing Warburg-subtypes based on the immunohistochemical (IHC) expression of six proteins involved in different levels of the Warburg-effect (PTEN, P53, GLUT1, PKM2, LDHA, MCT4). Furthermore, we investigated subgroups of colorectal cancer based on mutations in oncogenes that have been associated with the upstream regulation of the Warburg-effect (*KRAS*, *PIK3CA*, and *BRAF* mutations) as well as mismatch repair (MMR) status, hereafter referred to as subgroups of molecular features. In addition, we aimed to investigate whether non-pathologists can generate valid and reproducible IHC scoring results.

All studies presented in this thesis were conducted using data from the Netherlands Cohort Study (NLCS). The NLCS is a large prospective cohort study that was initiated in 1986, and included 120,852 subjects aged 55-69 years at baseline. Information on energy balance-related factors and other cancer risk factors were collected through a mailed, self-administered questionnaire at baseline. A case-cohort approach was used, in which cases were derived from the entire cohort, whereas person-years at risk for the entire cohort were estimated from a subcohort (n=5,000) randomly sampled at baseline. Cancer cases from the total cohort were identified via record linkage with the Netherlands Cancer Registry and the Dutch Pathology Registry, PALGA, covering 20.3 years of follow-up for the current thesis. After excluding cases and subcohort members with a history of cancer (except skin cancer) at baseline, a total of 4,597 incident colorectal cancer cases and 4,774 subcohort members were available.

For the Rainbow-TMA project (2012-2017), formalin-fixed paraffin-embedded (FFPE) tissue blocks from primary tumor and matched normal tissue were requested from 3,872 incident colorectal cancer cases. From these blocks, three tumor cores per case were sampled and combined into tissue microarrays (TMAs). In total, tumor tissue of 2,694 colorectal cancer cases was successfully assembled in 78 TMAs. For the Warburg-subtypes, TMAs were subjected to IHC in order to establish expression levels of six proteins involved in the Warburg-effect (PTEN, P53, GLUT1, PKM2, LDHA, MCT4).

All IHC stained TMAs were scored by three non-pathologist assessors and a random 10% was additionally scored by an experienced pathologist. The expression levels were combined into a pathway-based sum score and categorized into three Warburg-subtypes (Warburg-low, -moderate, -high). For subgroups of molecular features, two slices were cut from FFPE tissue blocks containing primary tumor. DNA was isolated from these tissue slices. Then, tumor DNA was screened for *KRAS*, *PIK3CA*, and *BRAF* mutations. In addition, MMR status was assessed using IHC staining of MLH1 and MSH2 on TMAs. These molecular features were investigated individually (*KRAS* mutations [$KRAS_{mut}$]; $PIK3CA_{mut}$; $BRAF_{mut}$; MMR deficiency [dMMR]) as well as combined (all-wild-type+MMR-proficient [pMMR]; any-mutation/dMMR).

After exclusion of cases and subcohort members with incomplete covariate data, 3,911 subcohort members were available for analyses, 1,972 colorectal cancer cases with complete IHC expression data for Warburg-subtypes (Chapters 3 and 4), and 1,934 cases with complete data on molecular features (Chapters 5 and 6). Multivariable Cox regression analyses were used to estimate associations of energy balance-related factors with Warburg-subtypes and with subgroups of molecular features in colorectal cancer.

In Chapter 2, we investigated whether non-pathologists can generate valid and reproducible IHC scoring results. This was done by assessing interobserver agreement between trained non-pathologists and an experienced pathologist and by assessing intraobserver agreement within non-pathologists. We found that trained non-pathologists can generate reproducible IHC scoring results that are similar to those of an experienced pathologist. Combining the scores of at least two non-pathologist assessors yielded optimal results.

In Chapter 3, we investigated associations of adult energy balance-related factors with Warburg-subtypes in colorectal cancer. We found that measures of adiposity (i.e. BMI and clothing size) were associated with an increased risk of Warburg-moderate and Warburg-high colon cancer in men, and with Warburg-low and Warburg-high colon cancer in women. Furthermore, we observed that measures of physical activity were mainly associated with a decreased risk of Warburg-low and Warburg-moderate colon cancer, both in men and women.

In Chapter 4, we investigated associations of early-life energy balance-related factors with Warburg-subtypes in colorectal cancer. We found that height was positively associated with colon cancer in men, regardless of Warburg-subtypes, and with Warburg-low colon and Warburg-low and -moderate rectal cancer in women. We did not observe clear patterns across associations of early-life energy restriction proxies with Warburg-subtypes. A high adolescent BMI was associated with an increased risk of Warburg-high colon cancer in men, and Warburg-moderate rectal cancer in women.

In Chapter 5, we investigated associations of adult energy balance-related factors with risk of colorectal cancer subgroups of molecular features. We found that adiposity

measures in women were only associated with an increased risk of $KRAS_{mut}$ colon cancer, whereas in men they were associated with all subgroups of molecular features. Furthermore, we observed that non-occupational physical activity was associated with any-mutation/dMMR colon cancer in both men and women, but not with all-wild-type+pMMR colon cancer.

In Chapter 6, we investigated associations of early-life energy balance-related factors with risk of colorectal cancer subgroups of molecular features. We found that height was associated with an increased risk of any-mutation/dMMR, but not all-wild-type+pMMR, colorectal cancer, with the exception of men with rectal cancer. Again, results on early-life energy restriction proxies in relation to risk of subgroups based on molecular features did not show clear patterns. A high adolescent BMI in men seemed to be mainly associated with an increased risk of $KRAS_{mut}$ colon cancer, and in women with $BRAF_{mut}$ and dMMR colon cancer.

Chapter 7 concludes this thesis with a summary of the main findings, interpretation of study results, a discussion of methodological considerations, and recommendations for future research. Overall, the results presented in this thesis seem to indicate a role for the Warburg-effect in the etiological pathway between adiposity and colon cancer, but not rectal cancer, both for adolescent and adult adiposity. A role for the Warburg-effect is not indicated by our results for associations of physical activity or adult-attained height with colorectal cancer risk. We did not observe clear patterns for the three energy restriction proxies and therefore do not draw a conclusion regarding early-life energy restriction. Since we were the first to investigate associations of energy balance-related factors with risk of colorectal cancer in relation to the estimated presence of the Warburg-effect in the tumor, confirmation in additional large prospective MPE studies is necessary.

NEDERLANDSTALIGE SAMENVATTING

Factoren gerelateerd aan de energiebalans, zoals overgewicht en fysieke activiteit, zijn geassocieerd met het risico op colorectaalkanker. Het mechanisme achter deze associaties is momenteel onbekend. Activatie van het zogenaamde Warburg-effect via de PI3K/Akt pathway is een van de voorgestelde mechanismen. Het Warburg-effect is een metabool fenotype dat wordt gekarakteriseerd door verhoogde aerobe glycolyse. Het Warburg-effect lijkt een belangrijke rol te spelen in de carcinogenese.

In dit proefschrift hebben we de mogelijke betrokkenheid van het Warburg-effect in de etiologische associatie tussen energiebalans-gerelateerde factoren en colorectaalkanker onderzocht. Hiervoor hebben we gebruik gemaakt van een moleculair pathologische epidemiologische (MPE) aanpak. We onderzochten associaties tussen energiebalans-gerelateerde factoren, zowel op volwassen leeftijd als vroeg in het leven, en het risico op colorectaalkanker in relatie tot aanwezigheid van het Warburg-effect in de tumor. Energiebalans-gerelateerde factoren op volwassen leeftijd betroffen body mass index (BMI), kledingmaat (als indicator voor tailleomtrek) en fysieke activiteit. Energiebalans-gerelateerde factoren vroeg in het leven betroffen lengte, indicatoren van energierestrictie (blootstelling aan de Nederlandse Hongerwinter, Tweede Wereldoorlog en Nederlandse Economische Depressie) en BMI op 20-jarige leeftijd. De geschatte aanwezigheid van het Warburg-effect was gebaseerd op immunohistochemische (IHC) expressie van zes eiwitten die op verschillende niveaus van het Warburg-effect een rol spelen (PTEN, P53, GLUT1, PKM2, LDHA, MCT4). Met deze data hebben we Warburg-subtypen gemaakt. Daarnaast hebben we subgroepen van colorectaalkanker onderzocht gebaseerd op mutaties in oncogenen die gelinkt zijn aan de aansturing van het Warburg-effect (*KRAS*-, *PIK3CA*- en *BRAF*-mutaties) en op de mismatch repair (MMR) status. Tenslotte hebben we onderzocht of niet-pathologen valide en reproduceerbaar IHC-kleuringen kunnen beoordelen.

De studies beschreven in dit proefschrift zijn uitgevoerd met data van de Nederlandse Cohort Studie (NLCS). De NLCS is een grote prospectieve cohortstudie waarin 120.852 mannen en vrouwen deelnemen die tussen de 55 en 69 jaar oud waren in 1986 tijdens de start van de studie. Alle deelnemers hebben bij de start van de studie een vragenlijst ingevuld over energiebalans-gerelateerde factoren en andere risicofactoren voor kanker. Het gehele cohort werd opgevolgd om nieuwe kankergevallen te identificeren. Daarnaast werd een willekeurig geselecteerd subcohort van 5.000 deelnemers gevolgd voor een schatting van de opgebouwde persoonstijd. Nieuwe kankergevallen uit het totale cohort werden geïdentificeerd via koppeling met de Nederlandse Kankerregistratie en de Nederlandse Pathologie Registratie (PALGA), met een totale follow-up van 20,3 jaar voor het huidige proefschrift. Na exclusie van kankergevallen en subcohortleden met een eerdere kankerdiagnose (m.u.v. huidkanker) waren er 4.597 nieuwe colorectaalkankergevallen en 4.774 subcohortleden beschikbaar voor deze studie.

Voor het Rainbow-TMA project (2012-2017) werd formaline-gefixeerd paraffine-ingebed weefsel, ook wel bekend als FFPE-weefsel, van de primaire tumor en het daarbij behorende normale weefsel van 3.872 nieuw gediagnosticeerde colorectaalkankergevallen opgevraagd. Uit de verzamelde FFPE-blokken werden drie tumorkernen per patiënt genomen en gecombineerd in zogenaamde "tissue microarrays" (TMA's). Tumorweefsel van 2.694 colorectaalkankergevallen werd succesvol samengevoegd in 78 TMA's. Voor de Warburg-subtypen werden TMA's onderworpen aan IHC-kleuring om de expressieniveaus van zes eiwitten vast te stellen die een rol spelen in het Warburg-effect (PTEN, P53, GLUT1, PKM2, LDHA, MCT4). Alle IHC-gekleurde TMA's werden gescoord door drie niet-pathologen. Daarnaast werd een willekeurige 10% van de TMA's ook gescoord door een ervaren patholoog. Vervolgens werden de expressieniveaus gecombineerd tot een somscore en gecategoriseerd in drie Warburg-subtypen (Warburg-laag/Warburg-matig/Warburg-hoog). Voor het bepalen van subgroepen van moleculaire factoren werden twee coupes van FFPE-blokken gesneden die primair tumorweefsel bevatten. Uit deze weefselcoupes werd DNA geïsoleerd en de *KRAS*-, *PIK3CA*- en *BRAF*-mutatie status bepaald. Verder werd de MMR-status beoordeeld met behulp van IHC-kleuringen op TMA's van de eiwitten MLH1 en MSH2. Deze moleculaire factoren werden zowel afzonderlijk (*KRAS*-mutaties [$KRAS_{mut}$]; $PIK3CA_{mut}$; $BRAF_{mut}$; MMR deficiënt [dMMR]) als gecombineerd (all-wild-type+MMR-proficiënt [pMMR]; any-mutation/dMMR) onderzocht.

Na de exclusie van colorectaalkankergevallen en subcohort leden met incomplete covariaat data waren er 3.911 subcohort leden beschikbaar voor statistische analyses, 1.972 colorectaalkankergevallen met volledige IHC expressiegegevens voor Warburg-subtypen (Hoofdstukken 3 en 4) en 1.934 colorectaalkankergevallen met volledige gegevens over moleculaire factoren (Hoofdstukken 5 en 6). Multivariabele Cox-regressieanalyses werden gebruikt om associaties tussen energiebalans-gerelateerde factoren en colorectaalkanker te schatten in relatie tot Warburg-subtypen en subgroepen van moleculaire factoren.

In Hoofdstuk 2 hebben we onderzocht of niet-pathologen valide en reproduceerbare IHC-scores kunnen genereren. Hiervoor hebben we naar de interbeoordelaarsovereenkomst tussen getrainde niet-pathologen en een ervaren patholoog gekeken en naar de intrabeoordelaarsovereenkomst van niet-pathologen. We zagen dat getrainde niet-pathologen reproduceerbare IHC-scores kunnen genereren die vergelijkbaar zijn met de IHC-scores van een ervaren patholoog. Het combineren van de IHC-scores van ten minste twee niet-pathologen leverde optimale resultaten op.

In Hoofdstuk 3 hebben we de associaties tussen energiebalans-gerelateerde factoren op volwassen leeftijd en Warburg-subtypen in colorectaalkanker onderzocht. We zagen dat adipositasmetingen (d.w.z. BMI en kledingmaat) geassocieerd waren met een verhoogd risico op Warburg-matige en Warburg-hoge colonkanker in mannen en met Warburg-lage en Warburg-hoge colonkanker in vrouwen. Verder zagen we dat fysieke activiteit voornamelijk was geassocieerd met een verlaagd risico op Warburg-lage en Warburg-matige colonkanker, zowel in mannen als vrouwen.

In Hoofdstuk 4 hebben we de associaties tussen energiebalans-gerelateerde factoren op jonge leeftijd en Warburg-subtypen in colorectaalkanker onderzocht. We zagen dat lengte was geassocieerd met een verhoogd risico op colonkanker in mannen, ongeacht Warburg-subtype, en met Warburg-lage colonkanker en Warburg-lage en –matige rectumkanker in vrouwen. We zagen geen duidelijke patronen in associaties tussen indicatoren van energierestrictie en Warburg-subtypen. Een hoog adolescent BMI was geassocieerd met een verhoogd risico op Warburg-hoge colonkanker in mannen en Warburg-matige rectumkanker in vrouwen.

In Hoofdstuk 5 hebben we de associaties tussen energiebalans-gerelateerde factoren op volwassen leeftijd en subgroepen van moleculaire factoren in colorectaalkanker onderzocht. We zagen dat adipositasmetingen bij vrouwen alleen geassocieerd waren met een verhoogd risico op $KRAS_{mut}$ colonkanker. In mannen waren adipositasmetingen geassocieerd met alle subgroepen van moleculaire factoren. Verder zagen we dat niet-beroepsmatige fysieke activiteit was geassocieerd met any-mutation/dMMR colonkanker in zowel mannen als vrouwen, maar niet met all-wild-type+pMMR colonkanker.

In Hoofdstuk 6 hebben we de associaties tussen energiebalans-gerelateerde factoren op jonge leeftijd en subgroepen van moleculaire factoren in colorectaalkanker onderzocht. We zagen dat lengte was geassocieerd met een verhoogd risico op any-mutation/dMMR colorectaalkanker, maar niet all-wild-type+pMMR colorectaalkanker, met uitzondering van mannen met rectumkanker. Ook met de subgroepen van moleculaire factoren lieten de indicatoren van energierestrictie indicatoren geen duidelijke patronen in associaties zien. Een hoog adolescent BMI bij mannen leek vooral geassocieerd te zijn met een verhoogd risico op $KRAS_{mut}$ colonkanker, en bij vrouwen met $BRAF_{mut}$ en dMMR colonkanker.

Dit proefschrift sluit af met een samenvatting van de belangrijkste bevindingen, interpretatie van onderzoeksresultaten, bespreking van methodologische overwegingen en aanbevelingen voor toekomstig onderzoek in Hoofdstuk 7. Over het algemeen lijken de in dit proefschrift gepresenteerde resultaten te wijzen op een rol voor het Warburg-effect in de etiologische associatie tussen adipositas en colonkanker, maar niet rectumkanker. Dit lijkt zowel voor adipositas tijdens adolescentie als op latere leeftijd te gelden. Het Warburg-effect lijkt geen rol te spelen in de etiologische associatie tussen fysieke activiteit of lengte met het risico op colorectaalkanker. We hebben geen duidelijke patronen waargenomen voor de drie indicatoren van energierestrictie, waardoor we geen conclusie kunnen trekken over een al dan niet verwachte rol voor het Warburg-effect in de etiologische associatie. Wij waren de eersten die associaties van energiebalans-gerelateerde factoren met colorectaalkanker onderzochten in relatie tot de geschatte aanwezigheid van het Warburg-effect in de tumor. Daarom is bevestiging van onze resultaten in aanvullende grote prospectieve MPE-studies noodzakelijk.

DANKWOORD

Hier ligt 'ie dan, mijn proefschrift! Maar alleen had ik dit nooit kunnen doen en daarom wil ik graag iedereen bedanken die me de afgelopen vier jaar heeft gesteund.

Mijn promotieteam wil ik in het bijzonder bedanken voor alle tijd en energie die jullie in mij en mijn proefschrift hebben gestoken. Toen ik zo'n vier jaar geleden tijdens mijn stage binnen de NLCS hoorde dat er een promotieplek vrij zou komen op het gebied van moleculair pathologische epidemiologie was ik direct enthousiast. Ik ben ontzettend blij dat jullie genoeg vertrouwen in mij hadden en mij de mogelijkheid hebben gegeven om onder jullie begeleiding aan de slag te gaan met dit interessante en uitdagende project.

Piet, bedankt voor de kansen die je me hebt gegeven om me verder te kunnen ontwikkelen. Je motiveerde me altijd om interessante cursussen te volgen of congressen te bezoeken, ook al waren ze niet altijd gerelateerd aan het project. Het project zelf was ook voor jou nog deels onbekend terrein, met veel potentiële beren op de weg. Dankzij onze wekelijkse overleggen en het vertrouwen dat je in Kelly en mij had hebben we samen met iedereen dit project tot een mooi einde laten komen. Ik heb ontzettend veel geleerd van je deskundigheid en kritische blik. Ik heb vooral veel geleerd van hoe geordend en gestructureerd je werkt, waardoor ook de chaotische ik steeds gestructureerder te werk ging. Onze wekelijkse overleggen begonnen en/of eindigden altijd met een gezellig praatje over onze vakanties, de ins en outs van de onderzoekswereld of over de actualiteiten, zelfs toen de overleggen online plaatsvonden. Bedankt voor alles!

Heike, thank you for your knowledge and patience while teaching us the basics of colorectal cancer histopathology. I always enjoyed our meetings where we discussed new stainings or scoring protocols. Once the period of data analysis and writing started, you joined our weekly meetings and gave input from a different perspective. Even though the pathological view sometimes differed quite a bit from the epidemiological view, we always came to an agreement which definitely improved our work. Once a paper was accepted you always wanted to celebrate with cake, which unfortunately happened quite late because of corona. Thank you for all your time, effort, interest, and especially enthusiasm during this project.

Colinda, naast co-promotor ken ik je vooral ook als een gezellige en toegewijde collega. Je was er regelmatig bij als er weer een borrel was en maakte altijd tijd voor een praatje. Voor het 'DNA deel´ heb jij ontzettend veel werk gedaan en hebben we jouw expertise goed kunnen benutten, bijvoorbeeld door jouw ervaring met wat er allemaal mis kan gaan met dataverwerking etc. In het laatste jaar sloot ook jij aan bij onze wekelijkse overleggen en heb ik veel aan je kennis en kritische blik gehad. Bedankt voor alles!

Graag wil ik ook de rest van het projectteam, ***Gregorio***, ***Harry***, ***Iryna***, ***Jaleesa***, ***Kim***, ***Leo*** en ***Matty***, bedanken voor de prettige samenwerking en de kritische blik op mijn manuscripten, maar ook voor de enthousiaste reacties als er weer nieuwe resultaten waren of als er een artikel was gepubliceerd!

Gregorio, je hebt ontzettend veel tijd en energie in ons project gestoken. We hebben heel wat uren samen op het lab gestaan of achter de microscoop/computer gezeten (is dit stroma nou desmoplastisch of niet!?). Maar gelukkig hebben we ook vaak genoeg met een kop koffie wat gekletst om onze ogen wat rust te gunnen, vooral het scoren van al die de TMA's liet ons duizelen… Bedankt voor alles wat je me hebt geleerd en voor de interesse die je bleef houden voor ons project nadat je 'weg was'. Maar vooral ook bedankt voor de altijd gezellige borrels, etentjes en koffiepauzes!

Harry, misschien wel de meest toegewijde van ons allemaal, zelfs in je pensioen stond je nog voor ons klaar! Wat hebben wij ontzettend veel gehad aan het software programma dat je voor ons hebt gemaakt om gemakkelijk, snel en overzichtelijk ál die TMA's te scoren. Intussen weet ik enigszins wat het verschil tussen de front- en back-end is, en je hebt me zelfs geleerd om kleine aanpassingen in het programma door te voeren. Ik heb het idee dat je daadwerkelijk van je hobby je werk hebt kunnen maken en dat je van verschillende aspecten van ons project echt hebt genoten, bijvoorbeeld het 'spelen' met QuPath. Harry, ontzettend bedankt voor alles wat je voor ons hebt gedaan, met jouw ICT hulp verliep alles een stuk soepeler! Ik hoop dat je geniet van je pensioen en je prachtige tuin (en vooral van het programmeren van je grasmaaier ;)).

Iryna, wat heb je, ondanks je drukke rooster, ontzettend veel tijd en moeite in ons project gestoken! We hebben heel wat tijd samen achter de microscoop of het computerscherm doorgebracht om kleurings- en scoringsprotocollen te optimaliseren. Maar hoe druk je ook was we bleven altijd even doorkletsen over vanalles en nog wat. Ik wil je ontzettend bedanken voor je tijd, maar vooral ook voor je interesse en je enthousiasme!

Leo, wie had er verwacht dat ons skype gesprek vijf jaar geleden over stagemogelijkheden mij uiteindelijk naar een promotietraject in Maastricht zou leiden? Ik heb ontzettend veel geleerd van mijn periode als stagiaire bij jou en Jeroen! Toen het einde van mijn stage naderde en je me tipte op een KWF grant die Piet had ontvangen voor een gaaf moleculair onderzoek binnen de NLCS heb ik natuurlijk direct gesolliciteerd. Ook toen ik mijn promotietraject was gestart bleef je erg betrokken bij het project en stond je voor me klaar wanneer er bijvoorbeeld iets misging met de heterogeniteitsanalyses ;). Bedankt!

De ***leden van de beoordelingscommissie*** wil ik bedanken voor de tijd en moeite die ze hebben gestoken in het kritisch lezen en beoordelen van dit proefschrift.

Collega's van de ***afdeling epidemiologie***, bedankt voor de leuke tijd de afgelopen vier jaar! Helaas vond ongeveer de helft van mijn promotie meer online dan op DEB plaats, maar gelukkig werden er ook in deze online tijd allerlei dingen georganiseerd om in contact te blijven. Ik voelde me altijd erg thuis bij jullie en vond het ook erg leuk om nog een laatste keer mee te gaan met het dagje uit, ook al was ik officieel al uit dienst.

Secretariaat, bedankt voor alle hulp de afgelopen vier jaar. ***Mariëlle***, jou wil ik in het bijzonder bedanken dat je altijd voor me klaarstond als ik weer een vraag had in de afrondingsfase van mijn promotie en om me te herinneren aan volgende stappen.

Conny, je staat altijd voor iedereen klaar met vragen maar ook voor een gezellig praatje, bedankt!

Harry, ***Jos***, ***Ron*** en ***Dominiek***, jullie stonden altijd voor me klaar als ik een programma nodig had of als mijn computer niet meer werkte, hartstikke bedankt hiervoor, jullie hulp was onmisbaar!

Jolanda, bekend van de verjaardagsposters en meesteres van de memes! Bedankt voor al je enthousiasme en je interesse, en voor het meedenken over een Maastrichtse stelling!

Christel, jouw blok was één van de eerste waar ik een tutorrol had. Ik moest mijn weg nog vinden in het onderwijs en jij hebt me hier ontzettend in geholpen met je belangstelling en tips. Daarnaast maakte je ondanks je drukke schema vaak tijd om even een praatje te maken, bedankt!

Adri, bedankt voor je interesse en de altijd gezellige praatjes bij het koffie apparaat!

IJmert, ook jij hield totaal niet van het thuiswerken waardoor wij elkaar in tijden van corona vaak op DEB tegenkwamen in de lege wandelgangen. Ik waardeer alle gezellige gesprekken en je interesse in mijn project erg. Laten we nog eens proosten met een ´Kessels blond´!

Laure, naast dat ik ontzettend veel van jou heb geleerd vond ik je ook altijd een erg gezellige en geïnteresseerde collega. Ik vond het dan ook erg leuk om samen de journal club te organiseren!

Eline, ik heb bewondering voor jouw positieve houding, harde werk en oprechte interesse in anderen. Ik heb ontzettend veel aan jou gehad tijdens mijn promotietraject! Je was er vaak bij als we gingen borrelen of naar feestjes als de boat party of ambtenarencarnaval. Ook tijdens corona spraken we soms af om gezellig een rondje te wandelen, zelfs nog toen je hoogzwanger was :). Hartstikke bedankt voor alles!

Maria, in September we heard that someone would join our office, and I couldn't have been happier that it was you! I love your positive energy and your sense of humour. Thank you very much for always cheering me up when I was so sick of writing and for all the laughs and coffees we shared!

Mede (oud) PhD-ers, bedankt voor alle gezellige lunches, borrels, etentjes en feestjes, met toch wel als hoogtepunt de karaoke avond in Groningen! ***Annaleen***, ***Jeroen***, ***Lisette***, ***Lloyd*** en ***Maya***, bedankt dat jullie deur altijd voor me openstond als ik weer vragen had of gewoon voor een gezellig praatje! ***Linda*** en ***Ariane***, bedankt voor de interesse en gezellige praatjes op de gang! ***Sophie*** en ***Jessica***, bedankt voor de gezellige feestjes en lunches! ***Karlijn***, bedankt voor de altijd gezellige etentjes en (werk)tripjes! ***Jacqueline***, met jou valt er altijd wat te lachen, maar je staat ook altijd voor me klaar met een luisterend oor. Na mijn sollicitatie ging ik direct bij jou naar binnen en ik ben blij dat we collega's zijn geworden, bedankt voor alles! ***Danielle***, ***Marvin*** en ***Nadira***, helaas was ik al (bijna) klaar toen jullie bij de vakgroep kwamen, maar alsnog bedankt voor de gezellige momenten en succes met jullie projecten!

Ook wil ik de ***afdeling pathologie*** bedanken voor de tijd die wij op het lab mochten doorbrengen voor alle kleuringen. Ik vond het erg fijn dat we ook werden uitgenodigd voor borrels en dergelijke, waardoor we een soort bonus afdeling hadden! In het speciaal wil ik ***Emil*** bedanken voor de vele (iets te lange) koffiepauzes, ook al was je

daar eigenlijk vééél te druk voor. En natuurlijk bedankt voor de altijd lekkere Lindt chocolade! Helaas zien en spreken we elkaar de laatste tijd wat minder nu je in Ulm woont, maar toch wil ik je graag bedanken voor de gezellige tijd in Maastricht!

Silvia and ***Giovanni***, thank you for all that I have learned during my time in beautiful Torino! I really appreciated you as my internship supervisors, my time with you actually inspired me to do a PhD. Grazie mille!!

Kiran, thank you for being a great friend in Torino but also afterwards! You helped me change my view on the world and I am forever grateful for that. Wishing you the best and hope to see you soon!

Lieve ***Amy***, ***Juul*** en ***Merel***, hier issie dan, het 'boekje'! Ik ben blij dat we elkaar (als restje) nog zo regelmatig zien. Dankjulliewel voor jullie interesse en de gezellige bezoekjes aan Bree, Breda en Rotterdam!

Dear ***Marlijne***, ***Marlou-Floor***, and ***Sophie***, I really enjoyed our little book club where we had great discussions on (sometimes not so) great books. I also really loved our dinners, drinks, parties and hikes! And Marlijne, our long SUP sessions with great talks in the sun always bring me vacation feelings. Thank you all for everything!

Lieve ***Roula***, ***Thaaer***, ***Nouran*** en ***Ez Aldeen***, شكراً voor alle heerlijke etentjes en gezellige momenten! Ik ben blij dat ik jullie mijn nieuwe familie mag noemen.

Lieve ***ooms***, ***tantes***, ***nichten*** en ***neven***, bedankt voor jullie interesse en alle gezellige momenten! ***Bart***, ***Gabi***, ***Lian***, ***Marcel***, ***Pascal***, ***Peter***, ***Renée***, ***Silja*** en ***Wil***, bedankt voor alle gezellige skivakanties en andere tripjes of dagjes weg die mij altijd, zelfs in de laatste weken van mijn PhD, goed lieten afschakelen! ***Peter*** en ***Gabi***, bedankt dat ik altijd welkom ben bij jullie, Zwitserland voelt als een tweede thuis voor me.

Lieve ***Dian***, ***Evelien***, ***Iris***, ***Kelly***, ***Megan***, ***Michelle*** en ***Nienke***, hoewel we elkaar kennen van onze bacheloropleiding zijn we allemaal toch wel verschillend terechtgekomen. Na zo´n vijf jaar Nijmegen onveilig te hebben gemaakt zijn we ook nog eens bijna allemaal uitgevlogen naar andere delen van Nederland, van Amsterdam tot Maastricht. Ik ben super blij dat ik jullie nog steeds zo regelmatig zie en dat ik altijd bij jullie terechtkan, bedankt!

Lieve ***Hangovers***, een leukere vriendengroep had ik me niet kunnen wensen! Met jullie kan ik altijd goed afschakelen, kan ik 100% mezelf zijn (en nog een beetje meer) en kan ik vooral ook heel veel lachen. Ook nu iedereen z'n eigen weg gaat en we niet meer allemaal in 'os Kessels dörp' wonen blijven we elkaar gelukkig regelmatig zien. Bedankt dat jullie er altijd voor me zijn! ***Kimberly***, bedankt voor de gezellige koffiepauzes waar we allebei even aan ons werk konden ontsnappen en bijpraten! ***Marieke*** en ***Julie***, ik vind het heel fijn dat ik jullie als 'mede-vriendloze-vriendinnen' nog zo vaak zie, ik heb ontzettend veel aan jullie gehad de afgelopen maanden!

Natuurlijk wil ik de ***beste paranimfen*** Marlou-floor en Kelly bedanken! Jullie zijn naast super fijne collega's ook allebei ontzettend goede vriendinnen van mij geworden!

Lieve ***Marlou***, toen ik je nog niet eens zo lang kende kwam op een borrel het random idee om mee te gaan naar Schotland om je broertje te bezoeken en te hiken. Dit heeft toen al snel de toon gezet voor verschillende wandeldagen en weekends. Daarnaast zijn we vorig jaar samen gaan surfen en city-trippen in Portugal, wat een heerlijke eerste échte vakantie na corona was dat! We hadden het allebei niet makkelijk tijdens de eerste lockdown en ik ben daarom erg blij dat we allebei in Maastricht waren voor wandelingen, sportsessies, etentjes en borrels in de buitenlucht. Dankjewel voor alle gezellige momenten, maar ook dat ik altijd bij je terecht kan om mijn hart te luchten. Nu je in Nijmegen woont zal ik je helaas wat minder zien, maar ik hoop dat we contact blijven houden en nog verschillende tripjes zullen maken!

Lieve ***Kelly***, ik had me geen betere match kunnen voorstellen als mede-PhD-er voor dit project (ook hier wil ik mijn promotieteam bij deze graag nog voor bedanken ;)). Wat begon met gezellige borrels groeide al snel uit tot (onder andere) jaarlijkse bezoekjes aan de dameszitting in Voerendaal, 11-11 in Maastricht en Auw Wieverbal in Valkenburg. Naast de carnavalsmuziek hadden we ook een gezamelijke passie voor Nederlandse rap, wat zorgde voor verbaasde gezichten als iemand ons kantoor bezocht en wij lekker met muziek op zaten te scoren. Toen we thuis moesten gaan werken mistte ik mijn kantoorgenootje, maar gelukkig deden we zeer regelmatig zoomen en was het weer alsof we samen op kantoor zaten. Ik ben ontzettend trots op wat jij in de afgelopen jaren hebt bereikt! Vooral dit jaar is een jaar van mijlpalen voor jou, wat al begon met de geboorte van lieve Luna. Naast dit alles stond (en sta) je altijd voor me klaar als ik er even doorheen zat, maar ook nog steeds voor gezellige borrels en feestjes! Kelly, ik kan nog vijf kantjes volschrijven maar zal het toch maar kort houden, ontzettend bedankt voor alles!!

Lieve ***pap***, ***mam***, ***Kristel***, ***Wouter***, ***Bart*** en ***Suze***, bedankt voor jullie interesse maar vooral voor alle gezellige momenten!

Pap, wanneer we samen naar open dagen gingen werd je zelfs zelf vaak enthousiast over de studies, wat mij dan ook weer extra enthousiast maakte. Onze reis naar India was een hele mooie maar ook confronterende gebeurtenis. Ik ben ontzettend blij dat we dit nog samen hebben kunnen doen en deze mooie herinneringen delen. Ik ben heel trots op hoe positief je in het leven staat en er zo veel mogelijk uit haalt.

Mam, mede dankzij jou raakte ik geïnteresseerd in het lab en medisch onderzoek omdat je ons op jonge leeftijd al meenam naar je werk. Je staat altijd voor ons klaar waar en wanneer dan ook. Weet dat het ook andersom is, ook al wonen we niet om de hoek. Je bent een groot voorbeeld voor me, bedankt voor alles!

Kristel, bedankt dat jij altijd geïnteresseerd bent in mij en mijn werk! Je vroeg altijd enthousiast voor een linkje als één van mijn artikelen was geaccepteerd. Ik wens je heel veel succes en plezier met jouw mooie PhD project!

Wouter, bedankt voor de altijd fijne gesprekken in de auto van en naar Kessel, tenzij er natuurlijk Formule 1 of wielrennen bezig is ;). Ik wens je heel veel succes met de laatste twee jaar co-schappen en in je verdere carrière.

Van politieke discussies (waar ik niet graag aan deelneem) tot gezellige

spelletjes (waar ik wel graag aan deelneem), het is altijd gezellig als we met z'n allen zijn. Bedankt voor alles!

Lieve ***Naser***, ik kende je net toen ik na mijn master op zoek ging naar een promotieplek. Na een afwijzing vertelde jij me dat er iets beters op me wachtte. Natuurlijk had je daar helemaal gelijk in want vervolgens kwam dit project op mijn pad. Tijdens mijn promotie stond je altijd voor me klaar en hielp je me altijd relativeren als het tegenzat. De afgelopen jaren waren niet altijd even gemakkelijk, maar ik heb ontzettend veel zin om met jou het leven tegenmoet te gaan. Jouw positieve kijk op het leven, rust en flexibiliteit hebben mij veel geleerd. Ik vind het mooi om te zien hoe we allebei van elkaar en elkaars cultuur leren en zo samen iets heel speciaals opbouwen. Ik ben benieuwd wat de toekomst ons brengt, إن شاء الله.

Liefste ***Bobbie*** en ***Loulou***, bedankt dat jullie altijd voor me klaarstonden tijdens vergaderingen door voor mijn beeld langs te lopen, om aandacht te miauwen en aan de kabel van mijn koptelefoon te trekken. Maar vooral bedankt voor de vele vele knuffels.

CURRICULUM VITAE

Josien Jenniskens was born on the 2nd of January 1994 in Venlo, the Netherlands. After graduating secondary school at Bouwens van der Boijecollege in Panningen in 2012, Josien studied Biomedical Sciences at Radboud University in Nijmegen. She obtained her bachelor's degree in 2015.

During her master's education, she specialized in the field of epidemiology. In addition, she expanded her knowledge on nutrition and health at Wageningen University & Research. As an intern, she investigated the association of processed food intake with DNA methylation and clinical manifestations of accelerated ageing at the Italian Institute for Genomic Medicine in Turin, Italy. For her second internship, she studied the associations of smoking and BMI with promoter CpG island methylation in clear-cell renal cell cancer at the Department of Epidemiology, Maastricht University.

After obtaining her master's degree in 2018, Josien started working as a PhD student at the Department of Epidemiology, Maastricht University, within GROW (School for Oncology and Reproduction). Under the supervision of professor Piet A. van den Brandt, professor Heike I. Grabsch, and dr. Colinda C.J.M. Simons she studied the potential role for cancer cell metabolism in the association of energy balance with colorectal cancer risk using data from the Netherlands Cohort Study (NLCS). The scientific results presented in this thesis have been published in international peer-reviewed journals and have been presented at (inter)national congresses.

Currently, Josien is working as a researcher at Statistics Netherlands.

LIST OF PUBLICATIONS

PUBLICATIONS IN SCIENTIFIC JOURNALS

Schoutrop ELM, Brouwer MAE, **Jenniskens JCA**, Ferro BE, Mouton JW, Aarnoutse RE, van Ingen J. The stability of antimycobacterial drugs in media used for drug susceptibility testing. Diagn Microbiol Infect Dis, 2018, 92.4: 305-308.

Echle A, Grabsch HI, Quirke P, van den Brandt PA, West NP, Hutchins GGA, Heij LR, Tan X, Richman SD, Krause J, Alwers E, **Jenniskens JCA**, Offermans K, Gray R, Brenner H, Chang-Claude J, Trautwein C, Pearson AT, Boor P, Luedde T, Gaisa NT, Hoffmeister M, Kather JN. Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. Gastroenterology 2020; 159: 1406-16.

Jenniskens JCA, Offermans K, Samarska I, Fazzi GE, Simons CCJM, Smits KM, Schouten LJ, Weijenberg MP, van den Brandt PA, Grabsch HI. Validity and reproducibility of immunohistochemical scoring by trained non-pathologists on Tissue MicroArrays. Cancer Epidemiol Biomarker Prev 2021; 30: 1867-74.

Krause J, Grabsch HI, Kloor M, Jendrusch M, Echle A, Buelow RD, Boor P, Luedde T, Brinker TJ, Trautwein C, Pearson AT, Quirke P, **Jenniskens JCA**, Offermans K, van den Brandt PA, Kather JN. Deep learning detects mutations in cancer histology generated by adversarial networks. J Pathol 2021; 254: 70-79.

Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors and risk of colorectal cancer expressing different levels of proteins involved in the Warburg-effect. Cancer Epidemiol Biomarker Prev 2022; 31: 633-46.

Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors in childhood and adolescence and risk of colorectal cancer expressing different levels of proteins involved in the Warburg-effect. Int J Cancer 2022; 150: 1812-24.

Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, van der Meer JRM, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors and risk of colorectal cancer based on *KRAS*, *PIK3CA* and *BRAF* mutations and MMR status. J Cancer Res and Clin Onc 2022; 148: 2723-42.

Jenniskens JCA, Offermans K, Simons CCJM, Samarska I, Fazzi GE, van der Meer JRM, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Energy balance-related factors in childhood and adolescence and risk of colorectal cancer based on *KRAS*, *PIK3CA* and *BRAF* mutations and MMR status. Molecular Carcinogenesis, 2022.

Offermans K, **Jenniskens JCA**, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Expression of proteins associated with the Warburg-effect and survival in colorectal cancer. J Pathol Clin Res 2022; 8: 169-80.

Offermans K, **Jenniskens JCA**, Simons CCJM, Samarska I, Fazzi GE, van der Meer J, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Association between mutational subgroups, Warburg-subtypes, and survival in patients with colorectal cancer. Cancer Medicine 2022.

Echle A, Ghaffari Laleh N, Quirke P, Grabsch HI, Muti HS, Saldanha OL, Brockmoeller SF, van den Brandt PA, Hutchins GGA, Richman SD, Horisberger K, Galata C, Ebert MP, Eckardt M, Boutros M, Horst D, Reissfelder C, Alwers E, Brinker TJ, Langer R, **Jenniskens JCA**, Offermans K, Mueller W, Gray R, Gruber SB, Greenson JK, Rennert G, Bonner JD, Schmolze D, Chang-Claude J, Brenner H, Trautwein C, Boor P, Jaeger D, Gaisa NT, Hoffmeister M, West NP, Kather JN. Artificial intelligence for detection of microsatellite instability in colorectal cancer – a multicentric analysis of a pre-screening tool for clinical application. ESMO Open 2022; 7: 100400.

SUBMITTED

Offermans K, **Jenniskens JCA**, Simons CCJM, Samarska I, Fazzi GE, Smits KM, Schouten LJ, Weijenberg MP, Grabsch HI, van den Brandt PA. Association between adjuvant therapy and survival in colorectal cancer patients according to metabolic Warburg-subtypes.