

# Indicators of energy balance and risk of epigenetic instability in colorectal cancer

Citation for published version (APA):

Hughes, L. (2011). *Indicators of energy balance and risk of epigenetic instability in colorectal cancer*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.20110519lh>

## Document status and date:

Published: 01/01/2011

## DOI:

[10.26481/dis.20110519lh](https://doi.org/10.26481/dis.20110519lh)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## **English summary**

Overweight is a reflection of energy imbalance, and occurs when the calories consumed (from foods, drinks and alcohol) outnumber the calories used by the body (for normal body functions and basal metabolic rate, daily activities, and exercise). It is well accepted that a large body size (both body fat and adult attained height) increases the risk of colorectal cancer (CRC), while high levels of physical activity and caloric restriction have been shown to decrease the risk of CRC. How this risk differs according to men and women, and according to different anatomical sub-sites of the large intestine is unclear. Furthermore, CRC is not a single disease. Rather, CRC encompasses a heterogeneous complex of diseases characterized by numerous molecular abnormalities. By distinguishing underlying pathways responsible for these differences, the exposure factors that are thought to cause CRC and differential risk between individuals and populations may be better understood.

Of interest to this thesis is the serrated neoplasia pathway to CRC. Tumours arising from this pathway are characterised in part by a high degree of epigenetic instability, namely, widespread DNA methylation in the promoter region of several genes resulting in transcriptional silencing of the gene in question. This phenomenon has been labeled the CpG island methylator phenotype (CIMP). Tumours arising from this pathway also display a high degree of microsatellite instability (MSI), a form of genetic instability, and mutations in the *BRAF* proto-oncogene. This pathway was of particular interest because it is thought that timing of exposure to a given risk factor may be important for establishing epigenetic mechanisms such as methylation.

To date, epidemiologic data for associations between indicators/aspects of energy balance and risk of molecular abnormalities associated with epigenetic instability (CIMP, MSI, *BRAF* mutations) is rare. The aim of this thesis was to elucidate potential differences in risk factors for the sub-sites of the large intestine and sex, and for molecular subtypes of these tumours. This may lead to a better understanding of CRC aetiology and prevention.

All studies in this thesis were completed within the Netherlands Cohort Study on diet and cancer (NLCS). The NLCS is a prospective cohort study that includes 120,852 men and women who were between the ages of 55 and 69 at baseline (1986), and completed a self-administered

questionnaire on dietary habits, lifestyle, and other risk factors for cancer. The cohort was followed-up for cancer incidence through linkage to the Netherlands Cancer Registry, and PALGA (the nationwide registry for histopathology and cytopathology in the Netherlands). Further, a sub-cohort of 5000 individuals was randomly selected after baseline exposure measurement, to estimate accumulation of person time at risk in the cohort through biennial follow-up of vital status. Overall CRC could be investigated among 3360 cases from a follow-up period of 16.3 years after baseline. Tumour material was collected from CRC patients identified within the first 7.3 years of follow-up, excluding the first 2.3 years, and of these, 734 had sufficient DNA for molecular analyses. The research described in Chapters 3, 5 and 6 was based on data from 16.3 years of follow-up of the cohort. The research described in Chapters 4, 7 and 8 was based on data from 7.3 years of follow-up.

We begin in **Chapter 2** by describing a study in which we investigated the utility of self-reported trouser/skirt size as a predictor of cancer risk. Although epidemiological studies frequently utilize waist circumference as an indicator of intra-abdominal fat, and subsequently cancer risk, this variable was collected only from sub-cohort members of the NLCS. In weight-stable sub-cohort members, we observed that self-reported trouser/skirt size correlated well with self-reported waist circumference. Furthermore, self-reported clothing size appeared to predict risk of cancer in men and women independently of BMI. These findings indicate that self-reported trouser/skirt size can be used as a surrogate marker for waist circumference in the NLCS, and therefore, it was considered as an independent risk factor in the analyses presented in Chapters 3-8 of this thesis.

In **Chapter 3**, we investigated the association between exposure to energy restriction during childhood and adolescence and risk of CRC later in life. Energy restriction appears to be one of the most convincing interventions for lowering cancer risk in laboratory animals. When investigating risk in humans, epidemiological researchers often depend on historical data, for example exposure to a period of known energy restriction, such as a famine or war, as a proxy measure for energy restriction earlier in life. The NLCS is comprised of individuals who grew up during the years of the Economic Depression (1932-40), World War 2 (1940-44) and the Hunger Winter of

1945. A number of proxy measures were collected from cohort members that reflect their exposure to energy restriction during these three time periods. We observed no associations between exposure to the Economic Depression, WW2 years and CRC risk, and hypothesized that energy restriction was not severe enough during these time periods to influence risk. We did however observe that men who resided in a city exposed to the Hunger Winter had a lower risk than those who did not. This was true for total CRC, and tumours of the proximal colon and rectum. Inverse associations were also observed for women who were exposed to the Hunger Winter, but these did not reach statistical significance.

We then investigated the association between these exposures and CRC risk according to molecular endpoints, specifically CIMP and MSI, in **Chapter 4**. Adolescence is a period of rapid growth and change, and we hypothesized that exposure to a severe environmental exposure during this period of life, such as what would have been experienced during the Hunger Winter, may influence the development of epigenetic mechanisms. We observed that men and women exposed to the Hunger Winter had a lower risk of developing a tumour characterized by CIMP compared to those not exposed. We did not observe such an association with respect to MSI.

In **Chapter 5**, we considered adult aspects of body size and risk of CRC, specifically BMI at age 20, BMI at baseline, BMI change, trouser/skirt size and height. We observed that risk appears to differ according to sex. In men, risk of CRC increased with increasing BMI, BMI at age 20, BMI change and trouser size. This was especially true for tumours of the distal colon. In women, body fat was not associated with CRC risk in women unless considered simultaneously with physical activity; a high trouser/skirt size and low level of physical activity increased risk of all tumor subtypes. Height was associated with increased risk of CRC in women, but not men.

We also observed sex specific differences with respect to physical activity and CRC risk, as described in **Chapter 6**. Occupational physical activity appeared to be more protective of CRC in men, especially tumours of the distal colon, and non-occupational physical activity appeared to be more protective in women, especially for tumours of the distal colon and rectum.

Next, we wanted to investigate whether BMI, BMI at age 20, trouser/skirt size, height and physical activity differentially influenced risk of molecular subtypes of CRC tumours. In **Chapter 7**, we considered this with

respect to CIMP status. We observed that individuals with a high BMI and/or trouser/skirt size were at greater risk of developing CRC regardless of CIMP status, and physical activity was protective regardless of CIMP status. However, individuals who reported a higher BMI at age 20 had a greater risk of developing a CIMP tumour than a non-CIMP tumour. This observation builds on the hypothesis that methylation is an early event in CRC progression. However, we did not observe a difference in risk with respect to height, which is a proxy for nutritional exposures early in life.

For the study presented in **Chapter 8**, we used data from both the NLCS and the Melbourne Collaborative Cohort Study (MCCS) to assess the association between body size and risk of CRC characterized by *BRAF* mutations and MSI. We pooled data from the studies together to improve the precision of risk estimates, and observed that BMI and waist measurements were associated with all tumour subtypes (*BRAF* mutated and *BRAF* non-mutated; MSI and MS-stable). However, taller individuals were at greater risk of developing a tumour characterized by either a *BRAF* mutation or MSI. These findings provide more evidence that early life exposures may influence (epi) genetic mechanisms later in life. We hypothesize that there was not enough statistical power to observe a differential association in Chapter 7 with respect to height and CIMP.

This thesis concludes with a summary of the main findings, a discussion about the limitations and insights offered by molecular pathological epidemiology, a proposed hypothesis to explain our observations, and implications for future research (**Chapter 9**). We first acknowledge that CRC is a complex disease and that heterogeneity and associations with a given risk factor are far from black and white. Another issue is that in any molecular epidemiological study, sample size will be a problem. One way to overcome this issue is to pool data from similar studies, as we did in Chapter 8 of this thesis. We also address the issue that CIMP as an indicator of epigenetic instability in CRC research has not been without controversy. In order to overcome this controversy, a consensus about gene panels, threshold values and laboratory methods used to define CIMP must be reached.

Based on our observations, we suggest that timing of exposure is important for risk of CRCs characterized by epigenetic instability. This hypothesis may influence future research questions, as the incidence of

childhood overweight and obesity continues to increase and the long-term consequences of being obese during this critical period of growth and development remain largely undetermined.

## **Nederlandse samenvatting**



Overgewicht is een reflectie van een verstoorde energie balans en vindt plaats wanneer de geconsumeerde calorieën (uit voedsel) talrijker zijn dan de calorieën gebruikt door het lichaam (normale lichaamsfuncties en basaalmetabolisme, dagelijkse activiteiten en lichaamsbeweging). Over het algemeen wordt aangenomen dat een groot lichaamsvolume (zowel lichaamsvet en -lengte van een volwassen persoon) het risico op dikkedarmkanker<sup>1</sup> (DDK) verhoogt, terwijl een hoge fysieke activiteit en calorie beperking het risico op DDK verlaagt. Hoe dit risico op kanker verschilt tussen mannen en vrouwen en in verschillende anatomische delen van de dikke darm is onduidelijk. Bovendien staat DDK niet op zichzelf. Beter gezegd is DDK een heterogene ziekte gekarakteriseerd door talrijke moleculaire afwijkingen. Door de onderliggende metabole routen te karakteriseren voor deze verschillen zouden de bekende oorzaken van DDK en de verschillende risico's tussen individuen en populaties beter begrepen kunnen worden.

Van belang in het kader van dit proefschrift is de ontwikkeling van colorectaal carcinoom via de geserreerde route. Tumoren die voortkomen uit dit proces worden gedeeltelijk gekarakteriseerd door een hoge mate van epigenetische instabiliteit. Zo kan genexpressie geblokkeerd worden door DNA methylering in de promotorregio van het gen in kwestie. Dit verschijnsel wordt aangeduid als *CpG island methylator phenotype* (CIMP). Tumoren die via deze route ontstaan correleren met een hoge mate van microsatelliet instabiliteit (MSI), een vorm van genetische instabiliteit, en mutaties in het proto-oncogen *BRAF*. Ongeveer 15% van alle DDKs ontstaan via dit proces en deze tumoren zijn interessant omdat men denkt dat de periode van blootstelling aan de gegeven risicofactor belangrijk kan zijn voor het vaststellen van epigenetische mechanismen als methylering.

Vandaag de dag zijn epidemiologische data over associaties tussen indicatoren van de energie balans en het risico op DDK gekarakteriseerd door moleculaire afwijkingen geassocieerd met epigenetische instabiliteit (CIMP, MSI, *BRAF* mutaties) zeldzaam. Het doel van dit proefschrift was het onderzoeken van potentiële verschillen in risicofactoren voor kanker in verschillende delen van de dikke darm, naar geslacht, en voor moleculaire subtypes van deze tumoren. Dit zou kunnen leiden tot een beter inzicht in de etiologie en preventie van DDK.

<sup>1</sup>Onder DDK wordt verstaan kanker van de dikke (colon) en endel (rectum) –darm

Alle studies in dit proefschrift zijn uitgevoerd binnen de Nederlandse Cohort Studie naar voeding en kanker (NLCS). De NLCS bestaat uit 120.852 mannen en vrouwen die bij aanvang van de studie in 1986 een leeftijd tussen de 55 en 69 hadden en een vragenlijst over eetgewoonte, levensstijl en andere risicofactoren voor kanker, ingevuld hebben. Het cohort werd gevolgd voor het optreden van incidentie kankergevallen door een koppeling met de Nederlandse Kankerregistratie en PALGA (Pathologisch-Anatomisch Landelijk Geregistreerd Archief). Bij aanvang van de studie, na de beginmeting, is een subcohort van 5000 personen willekeurig geselecteerd voor het schatten van de opgebouwde persoonstijd door middel van de follow-up van de vitale status. Na een follow-up periode van 16,3 jaar, werd DDK vastgesteld bij 3360 personen uit het totale cohort. Tumormateriaal werd verzameld van DDK patiënten geregistreerd in de eerste 7,3 jaar van de follow-up, met uitzondering van de eerste 2,3 jaar. Van 734 patiënten was voldoende DNA materiaal aanwezig voor moleculaire analyses. Het onderzoek dat beschreven wordt in de hoofdstukken 3, 5 en 6 was gebaseerd op data verkregen gedurende 16,3 jaar na aanvang van het cohort. Het onderzoek beschreven in de hoofdstukken 4, 7 en 8 was gebaseerd op data verkregen gedurende 7.3 jaar na aanvang van de studie.

Als eerste, wordt in **hoofdstuk 2** een studie beschreven over de bruikbaarheid van zelfgerapporteerde broek/rok maat als voorspeller van het risico op kanker. Ondanks het feit dat epidemiologische studies vaak taille omtrek gebruiken als een indicator voor intra-abdominaal vet en vervolgens het risico op kanker, werd deze variabele slechts verzameld voor een deel, de subcohortleden, van de NLCS. We observeerden in subcohortleden met een stabiel gewicht dat zelfgerapporteerde broek/rok maat goed correleert met zelfgerapporteerde taille omtrek. Daarbij bleek dat de zelfgerapporteerde kledingmaat het risico op kanker in mannen en vrouwen voorspelt onafhankelijk van BMI. Deze bevindingen geven aan dat zelfgerapporteerde broek/rok maat als een surrogaat marker voor taille omtrek gebruikt kan worden in de NLCS en werd daardoor beschouwd als een onafhankelijke risicofactor in de analyses beschreven in hoofdstukken 3 tot en met 8 van dit proefschrift.

In **hoofdstuk 3** onderzochten we de associatie tussen energie beperking gedurende de kinderjaren en adolescentie en het risico op DDK op latere leeftijd. Energie beperking blijkt een van de meest overtuigende

interventies te zijn voor het verlagen van het risico op kanker in proefdieren. Voor het onderzoeken van deze verbanden bij mensen zijn epidemiologische onderzoekers voornamelijk afhankelijk van historische data. De blootstelling aan een periode van energie beperking, zoals een hongersnood of oorlog, kan daarbij dienen als een indirecte maat voor energie beperking op jongere leeftijd. NLCS deelnemers groeiden op ten tijde van de Grote Depressie (1932-1940), de Tweede Wereldoorlog (1940-1944) en de Hongerwinter van 1945. Deelnemers waren gevraagd te rapporteren of hun vader werk had tijdens de Depressie en wat het woonadres was tijdens de Oorlog en de Hongerwinter. Deze informatie werd beschouwd als proxy maten voor de mate van energie beperking gedurende de jeugd. We zagen geen associaties tussen de blootstelling aan de Grote Depressie en de Tweede Wereldoorlog en het risico op DDK en veronderstelden dat de energie beperking niet hevig genoeg was gedurende deze periodes om van invloed te zijn op het risico. We observeerden echter wel dat mannen die in een stad woonden die getroffen was door de Hongerwinter een lager risico op DDK hadden dan degene die daar niet woonden. Dit was in het bijzonder het geval voor totale DDK en tumoren in het proximale deel van het colon en het rectum. Inverse associaties werden eveneens geobserveerd voor vrouwen die de Hongerwinter ondervonden, deze waren echter niet statistisch significant.

Vervolgens onderzochten we in **hoofdstuk 4** de associatie tussen deze periodes en het risico op DDK gekarakteriseerd door moleculaire veranderingen, in het bijzonder CIMP en MSI. Adolescentie is een periode van snelle groei en verandering en we stelde de hypothese op dat de gevolgen van blootstelling aan heftige omstandigheden gedurende deze periode van het leven, zoals dat wat men ervaart gedurende de Hongerwinter, de ontwikkeling van epigenetische mechanismen beïnvloedt. We observeerden dat mannen en vrouwen die blootgesteld waren aan de Hongerwinter een lager risico hadden om tumoren te ontwikkelen die door CIMP gekarakteriseerd werden in vergelijking met diegene die niet blootgesteld waren aan de Hongerwinter. We vonden geen associatie met MSI.

In **hoofdstuk 5** beschouwden we de volwassen aspecten van lichaamsgrootte en het risico op DDK, in het bijzonder BMI op 20-jarige leeftijd, BMI bij aanvang van de studie, BMI verandering, broek/rok maat en lengte. We observeerden dat het risico afhankelijk lijkt te zijn van geslacht. In mannen leidt een hogere BMI, BMI op 20-jarige leeftijd, BMI verandering

en broekmaat, tot een hoger het risico op DDK. Dit was voornamelijk het geval bij tumoren in de distale deel van de dikke darm. Lichaamsvet was bij vrouwen niet geassocieerd met het risico op DDK, tenzij het samen met lichamelijke activiteit werd beschouwd. Een grote broek/rok maat en een lage maat van fysieke activiteit verhoogden het risico op alle tumor types. Lichaamslengte was geassocieerd met een verhoogd risico op DDK in vrouwen, echter niet in mannen.

We namen eveneens geslachtspecifieke verschillen waar met betrekking tot lichamelijke activiteit en het risico op DDK, zoals beschreven in **hoofdstuk 6**. Beroepsmatige lichamelijke activiteit bleek beter beschermend tegen DDK in mannen, vooral tegen tumoren in het distale deel van de dikke darm, en niet-beroepsmatige fysieke activiteit bleek beter beschermend in vrouwen, vooral tegen tumoren in het distale deel van de dikke darm en het rectum.

Vervolgens wilden we onderzoeken in welke mate BMI, BMI op 20-jarige leeftijd, broek/rok maat, lengte en fysieke activiteit van verschillende invloed zijn op het risico op moleculaire subtypes van DDK tumoren. In **hoofdstuk 7** beschouwden we dit met betrekking tot CIMP status. We observeerden dat individuen met een hoge BMI en/of een grote broek/rok maat een hoger risico liepen om DDK te ontwikkelen ongeacht de CIMP status. Lichamelijke activiteit acteerde als een beschermende factor ongeacht de CIMP status. Individuen die een hoger BMI op 20-jarige leeftijd rapporteerden liepen echter een groter risico op het ontwikkelen van een CIMP tumor dan een non-CIMP tumor. Deze observatie past bij de hypothese dat methylering optreedt in een vroeg stadium in de ontwikkeling van DDK. We namen echter geen verschil in risico waar met betrekking tot lengte, welke als een indirecte maat voor de voedingsstatus in de jonge levensjaren gezien kan worden.

Voor de beschreven studie in **hoofdstuk 8** gebruikten we data van zowel de NLCS als de *Melbourne Collaborative Cohort Study* (MCCS) om de associatie tussen lichaamsgrootte en het risico op DDK gekarakteriseerd door *BRAF* mutaties en MSI te bepalen. De data van beide studies zijn samengevoegd om de nauwkeurigheid van de risicoschattingen te vergroten en we zagen dat BMI en omtrekmetingen geassocieerd waren met alle tumor types (*BRAF* gemuteerd en niet gemuteerd, MSI en MS-stabiel). Langere mensen liepen echter een groter risico op het ontwikkelen van een tumor die

gekaracteriseerd werd door of een *BRAF* mutatie of MSI. Deze bevindingen verschaffen extra bewijs dat blootstellingen gedurende de jonge jaren (epi) genetische mechanismen kunnen beïnvloeden. We stellen dat er niet genoeg statistische power was om verschillende associaties waar te nemen in hoofdstuk 7 met betrekking tot lengte en CIMP.

Dit proefschrift sluit af met een samenvatting van de belangrijkste bevindingen, een discussie over de beperkingen en inzichten verschaft door de moleculaire pathologische epidemiologie, een voorgestelde hypothese om onze waarnemingen te verklaren en suggesties voor toekomstig onderzoek (**hoofdstuk 9**). Als eerste erkennen we dat DDK een complexe ziekte is en dat de heterogeniteit en associaties met een gegeven risicofactor complex zijn. Een andere kwestie is dat in iedere moleculaire epidemiologische studie de grootte van de steekproef een probleem zal zijn. Een manier om deze kwestie te overwinnen is door de data van verschillende studies te poolen, zoals we deden in hoofdstuk 8 van dit proefschrift. We brengen eveneens de kwestie naar voren dat CIMP als een indicator voor epigenetische instabiliteit niet oncontroversieel is in DDK onderzoek. In een poging deze controverse te ondervangen, moet overeenstemming bereikt worden de gebruikte markers, afkappunten en laboratorium methodes die gebruikt worden om CIMP te definiëren.

Gebaseerd op onze waarnemingen suggereren we dat de timing van blootstelling belangrijk is voor het risico op DDK gekarakteriseerd door epigenetische instabiliteit. Deze hypothese kan toekomstige onderzoeksvragen beïnvloeden, vooral zolang de incidentie van overgewicht gedurende de kinderjaren blijft stijgen en de lange termijn gevolgen van obesitas tijdens deze kritieke periode van groei en ontwikkeling grotendeels onduidelijk blijven.