UiT The Arctic University of Norway

Faculty of Health Sciences

Department of Community Medicine

A healthy lifestyle index and cancer

Using a multifactor lifestyle exposure to estimate cancer incidence and survival among Norwegian women

Sairah Lai Fa Chen

A dissertation for the degree of *Philosophiae Doctor* May 2023

Cover illustration:

Medicine by Gustav Klimt, 1901 ©
Depicts Hygieia – goddess of health

Table of Contents

List of Tables

List of Figures

Acknowledgements

I wish to acknowledge the women who participated in the Norwegian Women and Cancer Study. We would be empty-handed without your generosity.

…

Throughout the course of this doctoral research project, I have had the privilege of receiving the guidance, encouragement, brilliance, and friendships of many people. I first wish to thank my main supervisor of the past two years, my homestretch coach, Therese H. Nøst. Words cannot express the depth and breadth of support you have provided me in many aspects of this journey. Thank you for trusting my thoughts more than I could; and for the generous discussions, outstanding guidance, and true kindness. To Torkjel M. Sandanger, my first main supervisor who launched me, for your leadership and care in lonely and uncertain times, your optimism, inclusivity, and for being thought-provoking. To Kristin B. Borch, for your remarkable compassion, willingness to be available, and sharing of expertise. To Tonje Braaten, for your invaluable assistance with the cohort and genius in statistics.

To my officemate and constant, Nikita Baiju – you are a wonder. It has been an honour to share a space with your wise soul from beginning to end of this journey. I wish to thank my other colleagues, many of whom I am glad to call friends – for their scholarly and human generosity, and for fostering a safe, positive place to research, teach, and be. Specifically, to Marie H. Lindberg and Jonas Johansson (pandemic saviours), Anja D. Norbye, Torill E. Jensen, Erlend H. Farbu, Kajsa Møllersen, Dolley D. Charles, Ilona Urbarova, Guri Skeie, Marko Lukic, and Karina S. Olsen – thank you for your humour and exchanges, big and small. To all, your welcoming, inclusive gestures have been worth more than you know.

I am so grateful for the delights, wits, and buttresses that are Alyssa Le Cornu, Lynda Hsu, Nora E. Lepri, Alvilde Vinge, and Azha L.S. Chen. To my parents, Rafeena and Steve Chen – thank you for your love, curries, calls, and needed visits. Very special thanks are owed to Eivind Sandstad, who was conned into partaking in the emotional highs and lows of this journey. Thank you for your unwavering support and for reminding me that life is, in fact, so good.

…

I dedicate this thesis to my grandmothers, Shiroon Ghani (1938-2012) and Tan Yin Ying (1940-2019) – women who endured much and gave everything.

Preface

This thesis is based on a collection of three scientific papers that are thematically bound by their use of a multifactor exposure measure, termed a *healthy lifestyle index*, to investigate cancer risk and survival. UiT The Arctic University of Norway provided funding to the Systems Epidemiology research group at the Department of Community Medicine to employ a PhD student to perform the research in this project. I am honoured to have been given the job.

Over the past century, humans have taken on epidemics of cancer, cardiovascular disease, and other metabolic syndromes at alarming rates and to wide detriment. The evidence suggests that this is a product of our changing ecologies as we modify the environment around us, and thus our behaviours. In line with this holistic thinking is the concept of the healthy lifestyle index. In the past decade, the use of a healthy lifestyle index has become increasingly popular. According to NCBI's PubMed, when searching for healthy lifestyle index in titles and abstracts, there were 9 publications before 2014. There were 68 publications from 2014 to February 2023. In this thesis, I assess the relationship between such an index and cancer-related outcomes.

This seemed a straightforward task at the outset. However, combining factors that science seemed to know much about (except for elusive diet) was challenging. The discussions that come in this thesis address the epidemiology and findings of the papers; and, importantly, the complexities of using multifactor exposure measures which have become so commonplace. I appreciate being given this opportunity to expand beyond the brevity of the papers.

Abstract

Background: Cancer is currently the leading cause of death in Norway and the number of new cancer cases is projected to continue its rise. With improvements in early detection and treatment, the number of survivors is increasing. It has been estimated that 30-50% of cancer cases are preventable and a large portion has been attributed to lifestyle behaviours.

Aim: This thesis aimed to investigate a healthy lifestyle index, as a simple multifactor exposure measure representing a gradient of healthy lifestyle behaviours, and its association to cancer incidence and survival among women in Norway.

Methods: This thesis used data from the Norwegian Women and Cancer Study (NOWAC) – a national prospective cohort of ~170 000 randomly selected women participants. A healthy lifestyle index (HLI) was constructed based on physical activity level, body mass index, smoking habits, alcohol intake, and habitual intake of major food groups. Cox proportional hazard models and restricted cubic splines were used to estimate associations between the HLI score and cancer incidence and survival, and the associations between HLI score change and cancer incidence.

Results: A higher HLI score was associated with lower risks of breast, colorectal, lung, endometrial, pancreatic, and kidney cancers. Regardless of baseline HLI score, greater positive HLI score changes were associated with lower risk of lifestyle-related cancers combined. Further, a higher prediagnostic HLI score was associated with lower all-cause mortality and, weakly, with lower breast cancer mortality among women diagnosed with breast cancer. Associations were also negative, but weak, for women diagnosed with colorectal cancer. No associations were observed for lung cancer mortality. Of all the lifestyle factors in the HLI, smoking was particularly strong in driving several associations.

Conclusion: Using a multifactor lifestyle exposure measure, Papers I-III observed associations between the HLI score, or HLI score change, and cancer-related outcomes. A healthy lifestyle, where smoking avoidance is a priority, should be promoted and facilitated throughout all adult ages to reduce the risk of cancer in the Norwegian general population of women. However, more research is required to understand the potential impact of lifestyle factors and overall lifestyle on cancer survival.

List of Papers

Paper I

Chen SLF, Braaten T, Borch KB, Ferrari P, Sandanger TM, Nøst TH. Combined Lifestyle Behaviors and the Incidence of Common Cancer Types in the Norwegian Women and Cancer Study (NOWAC). Clinical epidemiology. 2021;13:721-34.

Paper II

Chen SLF, Nøst TH, Botteri E, Ferrari P, Braaten T, Sandanger TM, et al. Overall lifestyle changes in adulthood are associated with cancer incidence in the Norwegian Women and Cancer Study (NOWAC) – a prospective cohort study. BMC Public Health. 2023;23(1):633.

Paper III

Chen SLF, Borch KB, Sandanger TM, Tinmouth J, Braaten T, Nøst TH. Combined prediagnostic lifestyle factors and survival of breast, colorectal, and lung cancer – Norwegian Women and Cancer study (NOWAC). Manuscript.

Abbreviations

ACS	American Cancer Society
AICR	American Institute for Cancer Research
AIC	Akaike Information Criterion
BMI	Body mass index
CI	Confidence interval
CUP	Continuous Update Project
EPIC	European Prospective Investigation into Cancer and Nutrition
FMI	Fraction of missing information
HDI	Human development index
HEI	Healthy Eating Index
HLI	Healthy lifestyle index
HR	Hazard ratio
ICD-10	International Classification of Diseases, 10th Revision
MAR	Missing at random
MCAR	Missing completely at random
MNAR	Missing not at random
MICE	Multiple imputation by chained equations
NOPAQ	Norwegian physical activity questionnaire
NOWAC	Norwegian Women and Cancer study

RCS	Restricted cubic splines
SEP	Socioeconomic position
SHR	Subdistribution hazard ratio
WCRF	World Cancer Research Fund
WHO	World Health Organization

1 Background

1.1 Preamble on cancer

Cancer is a large cluster of non-communicable diseases that is characterised by the uncontrolled growth, multiplication, and spread of cell masses. Through complex pathways that have yet to be completely understood, normal cells become malignant. These disruptions inactivate regulative cell cycle mechanisms that balance cell growth and functionality and promote abnormal interactions with its environment – integral components of the hallmarks of cancer (1). If left uninterrupted or untreated, cancer cells can invade other tissues, eventually impairing organ function, and cause death (2).

The rate and degree to which cancer impairs physiological function is dependent on many factors, including the site of the primary tumour, the type of tissue in which it is invading, and the environment (2). Cancers are often referred by their primary site of tumour development – i.e. breast cancer, colorectal cancer, and lung cancer. In an advanced disease state, cells from the primary tumour can acquire the ability to move and grow in different sites, an occurrence known as *metastasis.* Cancer that has spread from the initial tumour site, i.e. breast to the lung, is not lung cancer, rather, metastatic breast cancer. At this advanced stage, prognosis is far poorer compared to breast cancer that is in-situ (3).

What causes a normal, healthy cell to become a malignant cell? The vast body of scientific literature on cancer and oncology is a testament that the answer(s) to this question is manifold and the focus of many research fields where the unit of interest may be small to very large – genetic, chemical, microbiological, and physiological, to name a few. This thesis uses an epidemiological approach to study the impact of behavioural factors on cancer outcomes at the population level. It focuses on whether and the degree to which human behaviour impacts the cancer risk and the risk of death after a cancer diagnosis. Its broad vantage point is not separate from other fields, but provides a complement to the greater picture of cancer; and can contribute to the knowledge of patterns that are observed in populations and hypotheses generated by real world data.

Cancer can occur in any sex and at any age, even congenitally. There are large sex differences in the genetics, aetiology, and epidemiology of cancer (4). Further, the aetiology and sites of paediatric cancers are distinct from that of adult cancers (5). Importantly, even within these categories, there are diverse risk factor profiles across cancer sites, which is an underpinning

of this thesis. This thesis focuses on cancers occurring among adult females, which, for the sake of traditional classification, will be referred to as *women*.

1.2 Cancer burden

Cancer is currently a leading contributor to the global disease burden. In 2019, it was estimated that cancer was globally responsible for the second highest number of deaths, years of life lost, and disability-adjusted life years, only after cardiovascular disease (6). According to GLOBOCAN estimates, there were 19.3 million people diagnosed with cancer and 10 million cancer deaths in 2020 (7). Almost one in six deaths worldwide are caused by cancer (7).

Between countries, the ranking of cancer among causes of premature death is directly correlated with the ranking of socioeconomic development (7). Current and projected distributions of cancer incidence, mortality, and survival are inseparable from human development index (HDI) levels, reflecting trends in exposure to risk factors such as lifestyle behaviours and environmental contaminants, and differential access to screening and treatment (8, 9). There are thus considerable differences in the distribution of cancer sites and cancer causes of death across countries (10). Among women, breast cancer is most commonly diagnosed, except for in a cluster of low HDI countries, where cervical cancer is most commonly diagnosed. These most incident cancers are also largely the leading causes of cancer mortality, by country, with some exceptions. Several high HDI countries in North America and northern Europe report lung cancer as responsible for the most cancer deaths, despite breast cancer being the most incident (7).

In Norway, a high HDI country, the pattern of cancer incidence and mortality is largely consistent with other high HDI countries. Cancer is the leading cause of death, having surpassed that of cardiovascular disease in 2017 (11). Lung cancer was responsible for the greatest number of cancer deaths among women, followed by colon, breast, and pancreatic cancer (12). According to the Cancer Registry of Norway, there were 17 319 new cancer cases (12) diagnosed among 2 672 110 total females in 2021 (13). The majority (86%) of female cancer cases were diagnosed among women age 50 and over. The most frequent cancer sites for Norwegian women in the period 2017-2021 were breast, lung, colon, skin (non-melanoma), and melanoma of the skin (12).

Globally, the cancer burden is expected to be higher in 2040 compared to 2020, especially in low and medium HDI countries (7). However high HDI countries are projected to have the greatest absolute increase in new cancer cases on account of the aging population distribution. In line with global trends, the annual cancer incidence in Norway has been and continues on an upward trend (12).

The gap in cancer survival rates between countries again reflects differences in socioeconomic development. Poor survival is related to late detection of cancer and a lack of access to treatment. Norway is among the most fortunate of countries in this regard. Survival in Norway has been improving steadily across all cancer sites, which the Cancer Registry of Norway has largely attributed to advances in treatments (12). Further, the improved survival observed for breast and cervical cancers in Norway has been made partially attributable to the mammographic and cervical screening programmes that were initiated in 1995 and 1996, respectively (12). This, coupled with increased incidence, has led to an increasing number of cancer survivors. By the end of 2021, there were 316 145 cancer survivors in Norway. This represents a sharp increase from 217 977 cancer survivors at the end of 2011.

1.3 Factors related to cancer incidence

Cancer is a multifactorial disease, with a complex set of relevant host factor and environmental exposures that interact, including the timing and duration of those exposures. The latent period for cancers – the duration from the time of exposure to a carcinogen (an agent capable of causing cancer) to cancer diagnosis (14) – cannot easily be defined due to its many sufficient and component causes (15). Further, these risk factor profiles differ widely across cancer sites and types, as does our current level of understanding of the components of cancer-specific risk factor profiles. Nevertheless, it is estimated that 5-10% of cancer cases are heritable, whereby an allele variant is passed down to offspring that promotes the inhibition of tumour suppressor genes or activation of an oncogene (2, 16). However, cancer causing allele variants are heterogenous across tumour sites and types, as is their probability of being expressed. Rather, there is complex interplay between genes and the environment that gives rise to tumour development.

Non-hereditary factors are responsible for the majority of cancers (16). Cancer risk increases with age and age-specific cancer risk is higher for men compared to women (7). For women, reproductive system characteristics that impact their exposure to sex hormones are an

important part of the risk factor profile for many cancers, including the very commonly diagnosed breast, endometrial, and ovarian cancers. Socioeconomic development of the country (8, 9) and socioeconomic position (SEP) within the national population (17, 18) are strong determinants of exposure to environmental and lifestyle factors, which are important risk factors for cancer.

In several landmark publications, lifestyle and environment have been implicated as responsible for large proportions of cancer incidence. The Comparative Risk Assessment project estimated that 35% of cancers diagnosed globally in 2001 were attributable to 9 lifestyle and environmental exposures . Parkin et al. (20) estimated that 14 lifestyle and environmental factors were responsible for 42.7% of cancer cases in the UK in 2010 (20). There is also a large proportion of cancer that is unaccounted for after considering the estimated cancers attributable to hereditary and non-hereditary components. This gap continues to puzzle researchers in cancer epidemiology and oncology (21).

Although the amount of cancer attributable to known risk factors depends on the cancer site and type in question, several lifestyle factors, including tobacco smoke, overweight and obesity, alcohol intake, and fruit and vegetable intake have emerged as dominant. This has motivated the wide acceptance of primary prevention through lifestyle modification to reduce the cancer burden. The World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) and the International Agency for Research on Cancer (IARC), a division of World Health Organization (WHO), have established recommendations for cancer prevention based on their synthesis of single studies and meta-analyses. By and large, other public health authorities worldwide synchronise their recommendations to WCRF/AICR and WHO, including the Norwegian Institute of Public Health. The following will summarise the current evidence and consensus of the impact of lifestyle factors – physical activity, body fatness, smoking, alcohol intake, and dietary habits – on cancer risk.

1.3.1 Physical activity

Physical activity is defined as movement produced by skeletal muscles that requires energy expenditure. However, it is a complex lifestyle factor to capture as there are many components of physical activity and a diverse array of instruments that can be used to estimate physical activity levels. Duration, intensity, frequency, and domain – whether occupational, household, transport, or recreational – are all recognized components of

physical activity (22). The literature available on physical activity and cancer is largely observational with self-reported physical activity levels.

According to the studies reviewed in the WCRF/AICR 2018 Continuous Update Project (CUP), higher physical activity level is associated with lower risk for several cancers (23). They reported convincing evidence for colorectal cancer, probable evidence for postmenopausal breast and endometrial cancer, and limited evidence for oesophageal, lung, liver, and premenopausal breast cancer. The U.S. Department of Health and Human Services (HHS) Physical Activity Guidelines Advisory Committee (PAGAC) additionally reported that there is strong evidence that physical activity reduces the risk of cancers of the bladder, stomach, kidney, and lung (24). Studies on other cancer sites have been too limited to enable concrete interpretations of the evidence.

1.3.2 Body fatness

Adult body fat mass in excess, termed *overweight* and *obesity* in order of increasing health detriment (25), is considered a major risk factor for cancer. Various instruments have been used in studies to measure body fatness, with the most widely applied being BMI, defined as kg/m^2. The WHO defines adult overweight as BMI $\geq$ 25 and obese as BMI $\geq$ 30 (25).

In their 2016 review of the epidemiological (largely observational), experimental animal, and mechanistic literature on the potential impact of body fatness on cancer risk, the IARC Working Group identified 13 cancer sites as have sufficient strength of evidence in humans for the preventive effect of the absence of body fatness (26). This includes: postmenopausal breast, colorectum, endometrium, pancreas, kidney, liver, oesophagus, stomach, gallbladder, ovary, meningioma, thyroid, and multiple myeloma cancers. The WCRF/AICR 2018 CUP is largely in agreement with the conclusions from IARC, reporting that there exists strong evidence to support the hypothesis that greater adult body fatness increases the risk of cancer in the first 10 sites reported by IARC (23). The WCRF/AICR 2018 CUP additionally reported the existence of sufficient evidence for mouth, pharynx, larynx cancer sites. Large cohort studies (n > 100 000) and pooled analyses published in 2018 onwards have largely corroborated the findings from the WCRF/AICR 2018 CUP and 2016 IARC Working Group meta-analyses (27, 28, 29).

1.3.3 Smoking

The act of cigarette smoking, which will be referred to as *smoking*, has been repeatedly associated with higher cancer risk (30, 31). Smoke from tobacco use is considered a carcinogenic agent with sufficient evidence in humans – a Group 1 carcinogen – by IARC (32), whereby the only safe level of smoking is complete avoidance. According to the IARC Monograph, last updated in 2022, the evidence for causality is conclusive for 18 main sites, including lung, oral cavity, pharynx, stomach, liver, oesophagus, colorectum, pancreas, cervix, ovary, nasal cavity and paranasal sinus, larynx, kidney, bladder, renal pelvis and ureter cancers, and leukaemia (33). It has been estimated that 70% of lung cancer deaths and 21% of all cancer deaths worldwide would have been prevented if the population avoided smoking (19).

The evidence for breast cancer – the most common cancer among women in Norway – is weaker, and defined as "suggestive" by IARC (31). Nevertheless, a moderate and statistically significant dose-response association between smoking, measured as a function of duration and/or intensity, and higher breast cancer risk has been identified in meta-analyses (34). In contrast, smoking has consistently been associated with lower risk of endometrial cancer (Terry, 2002). The mechanisms for this association are unclear and could be spurious (Dimou, 2022).

1.3.4 Alcohol

The intake of alcoholic beverages has been strongly and consistently associated with higher cancer risk (32, 35). Consequently, alcohol is defined as a group 1 carcinogenic substance by IARC and as having sufficient evidence to be classified as a cause of cancer in nine main sites, including oral cavity, pharynx, larynx, upper aerodigestive tract, liver, oesophagus, colorectum, and breast (33). Sites listed by the WCRF/AICR 2018 CUP for which there is strong evidence for increased cancer risk due to alcohol intake is in agreement with IARC (23).

According to the meta-analysis performed by the WCRF/AICR 2018 CUP panel, a 10 g/day increment in alcohol consumed was associated with a range of 2% higher risk for stomach cancer to 25% higher risk for oesophageal cancer (23). Assuming a causal association between alcohol intake and cancer, it was estimated that alcohol intake was responsible for 4% of the total cancers that occurred globally in 2020 (36) based on the GLOBOCAN 2020

cancer incidence data (7) and risk estimates from the WCRF/AICR 2018 CUP meta-analyses (23).

1.3.5 Diet

Diet is perhaps the least understood lifestyle factor as it relates to cancer risk. Intake of foods is challenging to measure with accuracy or precision. Further, diets consist of innumerable components that can be examined anywhere from individual nutrient exposures to holistic dietary patterns, which are all part of total energy intake. The preparation of foods presents an additional modification to the already complex dietary exposures. In terms of food groups, IARC has listed processed meat and red meat intake as carcinogens with sufficient and limited evidence, respectively, for causing cancer (32). The WCRF/AICR 2018 CUP is somewhat in agreement, listing both processed- and red meat as having strong evidence for increasing the risk of colorectal cancer. The WCRF/AICR 2018 CUP has also concluded that a diet with higher glycaemic load increases the risk of endometrial cancer, and foods contaminated by aflatoxins increases the risk of liver cancer (23). In terms of food groups that prevent cancer, the WCRF/AICR 2018 CUP has listed dairy, whole grains, and foods containing fibre as having strong evidence for reducing the risk of colorectal cancer; and non-starchy fruits and vegetables products as having strong evidence for reducing the risk of aerodigestive cancers (23).

Overall diet studied through dietary patterns – the quantity, proportions, and varieties of food and drinks consumed, and their frequencies – have also been heavily examined in observational studies. Consensus is scarce and there is no singular dietary pattern that can be recommended for reducing all cancers (23). The WCRF/AICR 2018 CUP identified limited-suggestive risk decreases for oral cavity and laryngeal cancer in terms of the American Cancer Society (ACS) diet score; head and neck cancer in terms of the Healthy Eating Index-2005 score and alternate Mediterranean diet score score; and upper aerodigestive tract cancer in terms of the WCRF/AICR score (23). However, these reports were based on single prospective cohort studies and the European Prospective Investigation into Cancer and Nutrition (EPIC) – a multicentre study. A systematic review on the Mediterranean diet score reported that increasing adherence was associated with lower breast, colorectal, gastric, liver, head and neck, and prostate cancer (37). In another systematic review, the Healthy Eating Index, alternative Healthy Eating Index, and Dietary Approaches to Stop Hypertension score were associated with lower cancer incidence according to 31 studies (38). Overall, diets that

tend to be dominated by whole grains, fruits, vegetables, nuts, and unsaturated fats; and contain lower amounts of processed meat, red meat, sugar, saturated and trans fats have been association with lower cancer risk in some sites.

1.3.6 Altering risk by lifestyle modifications

While there is a wealth of evidence relating lifestyle factors to cancer incidence, there are far fewer studies that have been designed specifically to investigate the impact of lifestyle modifications on cancer risk. Smoking cessation is universally recommended by health authorities, with benefits for both avoiding tobacco-related cancer incidence and death (39, 40, 41, 42). Weight gain has also been widely recognised as associated with higher postmenopausal breast cancer risk in particular (23), but also with higher colon, oesophageal (adenocarcinoma), kidney, and endometrial cancer risk (26, 43). However, the risk lowering benefits of avoiding weight gain has not been clearly extended to weight loss (44, 45, 46). Increasing physical activity levels during adulthood has been associated with lower cancer risk in single studies (47, 48); although, the existing evidence is weak. Increasing alcohol intake has been associated with higher risk for postmenopausal breast cancer (49), while alcohol cessation has been associated with lower risk for aerodigestive cancers, including oesophageal cancer (50, 51). There appears to be little to no research conducted on diet-specific modification and cancer risk in the general population.

1.4 Factors related to cancer survival

As the number of cancer cases is expected to continue to rise globally and in Norway, and with improvements in early detection and treatment, there will be a growing number of individuals who are diagnosed with cancer, undergo treatment, and are in remission. As is the case with risk factors for cancer occurrence, there is a complex set of factors that can impact cancer survival outcomes, including tumour-, treatment-, and host-related characteristics.

The anatomical extent of disease, or *stage*, at the time of diagnosis is the most important factor that impacts survival (3). Regardless of the primary tumour site or cell type, survival worsens with advancing stage. Survival is also impacted by the degree of abnormality of the tumour cells, operationalised as *tumour grade*, which provides an indication of how fast the tumour cells are likely to multiply and grow (2). Further, breast cancer cells can express a combination of oestrogen receptors (ER), progesterone receptors (PR), and human epidermal growth receptor 2 (HER2). Their classification based on positive or negative expression is

known as *molecular subtype*, which is a determinant of tumour aggressiveness. The type of treatment administered to the patient plays a major role in survival (3). The choice of treatment is largely dependent upon the tumour characteristics described above, where late stage cancers typically require the most invasive treatments. Molecular subtype is particularly relevant for breast cancer diagnosis as it predicts the responsiveness to hormone therapy treatments (52). Further, quality of care, access to healthcare, and healthcare policy are closely tied to the effectiveness of treatment and rehabilitation, and thus survival (3).

It has been consistently documented that cancer survival varies according host characteristics, including, sex and age at diagnosis, depending on the cancer type (12, 53). The presence of comorbidities and poor immune status – also functions of age, sex, and SEP – can limit treatment options and has been shown to interfere with treatment completion (3, 54). Lower SEP has been consistently associated with lower cancer survival across most cancer sites and has been hypothesized to be related to inequitable access to healthcare and healthcare navigation (55). SEP is also related to lifestyle behaviours and given known gaps in SEP in Norway (56) and the increasing number of survivors, lifestyle may be important for survival.

There is increasing evidence to support the hypothesis that the primary preventive effects of following cancer prevention recommendations may also extend to cancer survival**.** Physical activity may be associated with better physical functions after surgery (57). Research findings have suggested that several metabolic factors including obesity, insulin resistance, and hyperinsulinemia are positively associated with the rate of cancer progression and tumour grade (54). Further, obesity (58) and smoking (30) have been linked to worse immune function. Given the obesity-reducing effects of higher physical activity levels and dietary habits that adhere to cancer recommendations, it could be hypothesized that this metabolic-related cluster of lifestyle factors – physical activity level, diet, and body fatness – and smoking can impact survival by extension. Furthermore, well-established risk factors for cancer are also known to be associated with risk for chronic diseases, including type II diabetes mellitus and cardiovascular diseases. These factors can ultimately impact overall survival rates among individuals who have been diagnosed with cancer (23). Challengingly, the relevant time window of exposure for lifestyle factors to take effect on survival could vary based on the lifestyle factor and cancer site in question (23). In this next section, current epidemiological evidence on the impact of lifestyle factors, measured in the pre- and postdiagnostic periods, on survival among adults diagnosed with cancer will be outlined.

1.4.1 Prediagnostic lifestyle

There are a limited number of studies investigating the effect of prediagnostic lifestyle on cancer survival, with some consistent findings. Higher prediagnostic BMI was shown to be consistently associated with poorer survival among women diagnosed with breast cancer (59) and suggested for people diagnosed with colorectal cancer (60). Obesity compared to normal weight before lung cancer diagnosis was consistently associated with better survival (61). There is also evidence that current compared to never smoking is associated with poorer survival among women diagnosed with breast cancer according to a meta-analysis (62); colorectal cancer according to several single studies (63, 64, 65, 66); and lung cancer according to one study (67).

It is unclear whether the risk reducing benefits of higher physical activity levels are extended to survival after cancer diagnosis as the number of studies targeting this hypothesis has been small. Several studies have indicated that higher prediagnostic physical activity levels have a protective effect on survival among people diagnosed with cancer in several sites (64, 68, 69, 70, 71, 72, 73). However, null associations have also been observed (63, 66, 74, 75, 76). Studies have also reported inconsistent findings for prediagnostic alcohol intake (63, 66, 71, 74, 77, 78, 79, 80, 81, 82, 83). Prediagnostic dietary habits have been seldom investigated as it relates to cancer survival. Higher vegetable intake was associated with improved survival among adults diagnosed with colorectal (66) and lung cancer (67). Findings are not convincing for the intake of selenium (84), polyphenols (85), meat and fibre (86), vitamin D (66, 87), and calcium and dairy products (87).

1.4.2 Postdiagnostic lifestyle

Lifestyle after diagnosis is often altered in response to tertiary prevention, which aims to prevent further morbidity, promote a better quality of life, and improve survival. Refraining from smoking and smoking cessation postdiagnosis is universally recommended and has been shown to be associated with improved survival among cancer patients compared to continued smoking (88). The WCRF/AICR 2018 CUP summarised that current evidence is not strong enough to establish specific recommendations for cancer survivors, with respect to physical activity, body fatness, alcohol intake, and diet (23). Nevertheless, the panel advises that it is not likely to be harmful to cancer survivors who have completed treatment to adhere to cancer prevention recommendations. The ACS has been more forthwith with recommendations for cancer survivors, which fall largely in line with recommendations for cancer prevention (89).

1.5 Multifactor lifestyle exposure

Most studies investigating the effect of lifestyle factors on cancer outcomes have aimed to isolate the effect of a single factor by adjusting for all covariate lifestyle factors. A multifactor approach to estimating the impact of overall lifestyle on chronic disease, mortality, and cancer outcomes has increased in popularity over the past two decades for its hypothesized benefit in addressing the synergistic effects of lifestyle risk factors (90) and for providing a simple measure to assess population health behaviours (91). Indeed, it is widely accepted that, for example, exposure to smoking and alcohol interact to pose a greater risk than simply the two risks combined for aerodigestive cancers (31). Several studies have documented that it is common for lifestyle behaviours to cluster among individuals, supporting the need to assess the impact of lifestyle factors on health outcomes in combination (92, 93, 94). An exposure that combines several evidence-based risk factors into a multifactor exposure measure that represents a gradient of healthy behaviours and overall healthy lifestyle has been proposed as one way of addressing this (90, 95).

1.5.1 Constructing multifactor lifestyle exposure measures

There is substantial diversity in the construction of multifactor lifestyle exposure measures in the literature. However, all share the characteristic of being represented by a single numeric score, often referred to as an *index*, with maximum and minimum scores representing either the most or least healthy combination of factors. The majority of indices are designed a-priori, meaning the selection and weighting of components are defined before data exploration. These are typically based on current knowledge for primary prevention and aim to capture different combinations of physical activity level, body fatness, smoking habits, alcohol intake, and diet. Most multifactor lifestyle exposures attribute equal weight to their components.

There are two main classes of multifactor exposures that have been used in cancer epidemiology – 1) scores operationalising adherence to an established set of recommendations and 2) generalised lifestyle indices, which are based on evidence for risk factors and the distribution of data. Regarding the first instance, the standardised score based on WCRF/AICR recommendations is the most prominent in the literature (96). It includes 8 equally weighted components: i) body fatness (assessed through BMI and waist circumference), ii) physical activity level, iii) intake of wholegrains, vegetables, fruits, beans, iv) intake of processed foods, v) intake of red meat and processed meat, vi) sugar sweetened

drinks, vii) alcohol intake, and viii) breastfeeding for mothers. Notably, the WCRF/AICR score does not include a smoking component as the WCRF/AICR does not provide specific recommendations for tobacco use.

The second instance has acquired several names, including *protective lifestyle factor index* and *modifiable lifestyle index*, with the most common being *healthy lifestyle index (HLI)*. The term *healthy lifestyle index (HLI)* will be used to refer to any index composed of multiple lifestyle factors that does not specifically intend to operationalise official recommendations, rather a gradient of healthy overall lifestyle. The construction of HLIs differ across publications in terms of their components and the number of graded categories for each component. For example, there have been several HLI publications within the EPIC study that have embraced multiple versions of an HLI, including the WCRF/AICR score, as described in Table 1.

Table 1. Comparison of multifactor lifestyle score construction across select publications used in the EPIC study

	WCRF/AICR score (97)	McKenzie et al. 2016 (98)	Botteri et al. 2022 (99)	Buckland et al. 2015 (100)
Score range	0-7 (0-8 for mothers)	0-20	0-16	0-4
Physical activity	*Weight: 1* 3 categories: METs	*Weight: 4* 5 categories: METs	*Weight: 4* 5 categories: METs	N/A
Body fatness	*Weight: 1* 3 categories: BMI or waist circumference	*Weight: 4* 5 categories: BMI	*Weight: 4* 5 categories: BMI	*Weight: 1* Dichotomous: BMI
Smoking	N/A	*Weight: 4* 5 categories: status, time since smoking cessation, smoking intensity	*Weight: 4* 5 categories: status, time since smoking cessation, smoking intensity	*Weight: 1* Dichotomous: status, time since smoking cessation
Alcohol	*Weight:1* 3 categories	*Weight: 4* 5 categories: g/day	*Weight: 4* 5 categories: g/day	*Weight: 1* Dichotomous: g/day
Diet	*Weight: 4* 3 categories: wholegrains, vegetables, fruit, and beans 3 categories: fast foods 3 categories: red and processed meat 3 categories: sugar-sweetened drinks	*Weight: 4* 5 categories: Score based on 6 dietary factors	N/A	*Weight: 1* Dichotomous: Relative Mediterranean diet score

1.5.2 Overall lifestyle and cancer risk

In a systematic review on prospective cohort studies investigating the association between multifactor lifestyle exposures and cancer incidence, overall healthier lifestyle was associated with lower total cancer incidence and the incidence of several site-specific cancers, including

breast, colorectal, lung, bladder, endometrial, oesophageal, kidney, liver, and gastric cancers (101). The systematic review reported that the healthiest compared to least healthy lifestyles were associated with 29% lower cancer incidence in 16 studies with a total of 1.9 million participants. Greater adherence to the WCRF/AICR score was associated with lower breast and colorectal cancer risk in 21 observational studies (102). Overall lifestyle and cancer risk has been investigated once among women in Norway (n = 17 145) with a focus on breast cancer, observing that a greater number of unfavourable lifestyle factors was associated with higher postmenopausal breast cancer incidence among non-users of hormone replacement therapy (103). In a sample of 6 315 Norwegian men and women who participated in a pilot screening study, Bowel Cancer Screening in Norway (BCSN), overall healthier lifestyle was associated with lower risk for advanced colorectal neoplasia (104).

With respect to the effect of changing overall lifestyle on cancer risk, this topic has only been assessed in two publications, to our knowledge. Both studies indicate that adopting a healthier lifestyle during adulthood is associated with lower colorectal cancer incidence (105) and lower lifestyle-related cancer incidence (105). Neither study included a dietary component in their multifactor exposure measures.

1.5.3 Overall lifestyle and cancer survival

Fewer studies have investigated the impact of overall lifestyle on cancer survival. Of the studies that do exist, there is substantial heterogeneity in the type of multifactor lifestyle exposure used with respect to when the exposure measurement took place (pre- or postdiagnosis), whether prediagnostic lifestyle was retrospectively assessed, or whether the study intended to capture lifetime exposure to risk factors or exposure in a more specific time frame. Regardless of the timing of the exposure, all studies that were encountered have observed that a healthier lifestyle before or after cancer diagnosis was associated with improved survival for those diagnosed with breast (prediagnostic (103); postdiagnostic (106, 107)), colorectal (prediagnostic (63, 71, 108); postdiagnostic (108)), colon (postdiagnostic (109)), ovarian (postdiagnostic (110)), and overall cancer (postdiagnostic (111)).

2 Aims of the thesis

While an appreciative pool of research has explored hypotheses relating single lifestyle behaviours to cancer outcomes, fewer studies have investigated lifestyle behaviours in combination. This has seldom been undertaken in sample of Norwegians. There is relatively little known about the effects of lifestyle modification, studied as a specific hypothesis, on cancer risk in general, regardless of country sample. When the project was initiated, there were no published population-based studies that focused on the impact of changes in combined lifestyle behaviours on cancer outcomes. Further, cancer survival is typically explored as function of clinical and healthcare system interactions (i.e. diagnosis, treatment, follow-up). However, cancer treatments, while improved, have not been as effective compared to other chronic diseases (19). There is evidence from epidemiological and biological studies that lifestyle behaviours may modify prognosis, but little research has been conducted on this topic.

In light of these areas where knowledge is lacking, this thesis aimed to investigate a healthy lifestyle index, representing a simple multifactor exposure measure for a gradient of healthy lifestyle behaviours, and its association to cancer incidence and survival among women in Norway.

Specific aims:

- Define an HLI combining five lifestyle risk factors, including physical activity level, smoking habits, body fatness, alcohol intake, and dietary habits. It should:
    - be a simple measure
    - represent a gradient of overall healthy lifestyle as defined by the inclusion of selected evidence-based risk factors
- Estimate the associations between HLI scores measured during adulthood and the incidences of common cancer types (Paper I).
- Estimate the associations between changes in HLI scores between two timepoints during adulthood and the incidences of lifestyle-related cancer types (Paper II).
- Estimate the associations between prediagnostic HLI score during adulthood and survival of breast, colorectal, and lung cancers (Paper III).

3 Materials and methods

3.1 Study sample – the Norwegian Women and Cancer study

The study samples for Papers I-III comprising this PhD project were obtained from the Norwegian Women and Cancer study (NOWAC). The NOWAC study is a national, population-based cohort consisting of 172 526 adult women, born in 1927-1957 and between the ages of 30 to 70 at recruitment, living in Norway. With the initial aim of investigating the association between oral contraceptives and breast cancer risk, NOWAC has since expanded its focus to other health outcomes and risk factors. Follow-up data collection, amended questionnaires, and the development of the Post-genome Biobank housing blood and tissue samples for a sub-cohort has enabled a wide range of investigation into women's health (112).

The participants were sampled randomly from the National Population Register of Norway. Recruitment took place in several calendar periods mainly represented by three enrolment waves: 1991-92, 1995-97, and 2003-07. At enrolment, randomly selected women were mailed an invitation with an information letter (Appendix) and a questionnaire (see Paper II, Supplementary File 1 and 2 for two example questionnaires). Most enrolment waves have been administered between one and three follow-up questionnaires.

All questionnaires have been either four or eight pages in length. Common to the four- and eight-page questionnaires was the inclusion of questions on use of oral contraceptives and hormone replacement therapy, reproductive history, age at menarche, age at menopause, smoking habits, physical activity, alcohol intake, height, weight, education, breast cancer screening, family history of breast cancer, sunbathing habits, and self-reported diseases. Eight-page questionnaires included an additional four page section consisting of a detailed food frequency questionnaire (FFQ) to measure dietary habits.

3.2 Designs and study samples

Papers I-III employed prospective cohort designs and drew individual study samples from NOWAC depending on the research question. Figure 1 describes the sample selection from NOWAC for Papers I-III. Further exclusions are described in more detail in text.

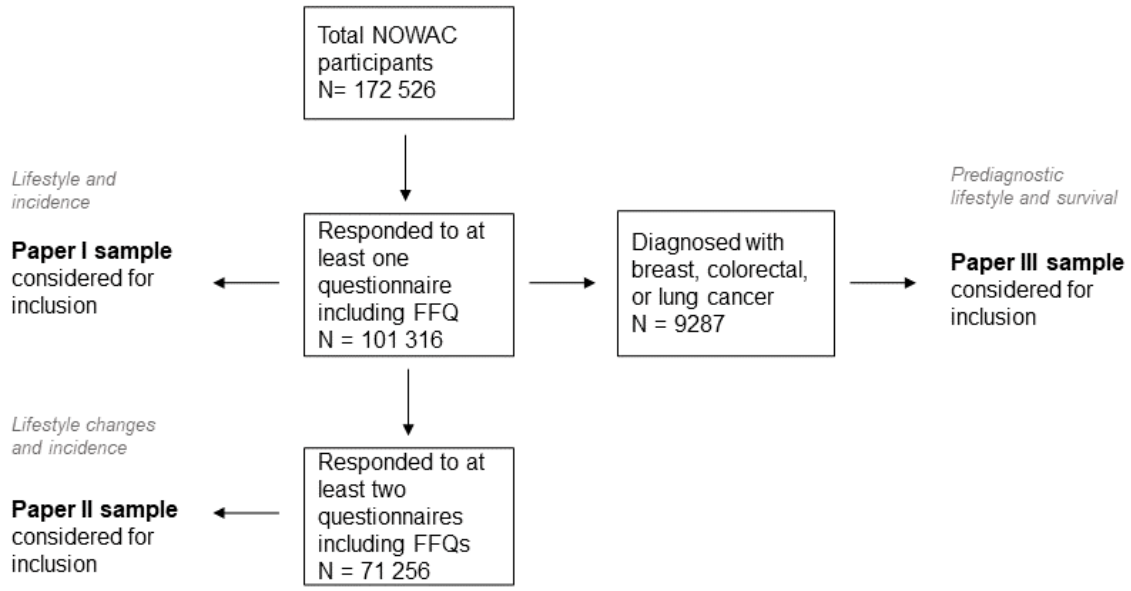


Figure 1 Sampling for Papers I-III from the Norwegian Women and Cancer Study (NOWAC)

Paper I study sample

Paper I considered all NOWAC participants who had responded to at least one questionnaire including an FFQ section for inclusion in the study sample (n = 101 316). Respondents' earliest questionnaire including an FFQ section served as baseline. As such, baseline information was collected from the first mailing of women enrolled in 1996-97 (mean age: 55.6) and 2003-04 (mean age: 53.0) and the second mailing of women enrolled in 1991-92, which was administered in 1998-99 (mean age: 47.7). The year at baseline therefore ranged from 1996 to 2004. There were 96 869 women remaining the analytical sample after conducting the following further exclusions: women with prevalent cancer, recorded as dead or emigrated before or in the same month as baseline, and women with extreme energy intakes (<2 100 or >15 000 KJ/day).

Paper II study sample

Paper II investigated HLI score change as the exposure, thus requiring information from two timepoints. NOWAC participants who answered at least two eight-page questionnaires were included in the sample for Paper II. Therefore, the sample for Paper II is a subset of the sample for Paper I, where the earliest eight-page questionnaire was used to obtain baseline

information and the subsequent eight-page questionnaire was used to obtain follow-up information.

A total of 71 256 women were eligible for inclusion. We excluded 5 018 prevalent cancer cases and women who were registered as deceased (n = 3) or emigrated (n = 2) before or at the date of the second eight-page questionnaire. The analytical sample consisted of 66 233 women. Figure 2 displays the timing of baseline, Questionnaire 1 (Q1) and follow-up, Questionnaire 2 (Q2) for the analytical sample and mean ages for each group.

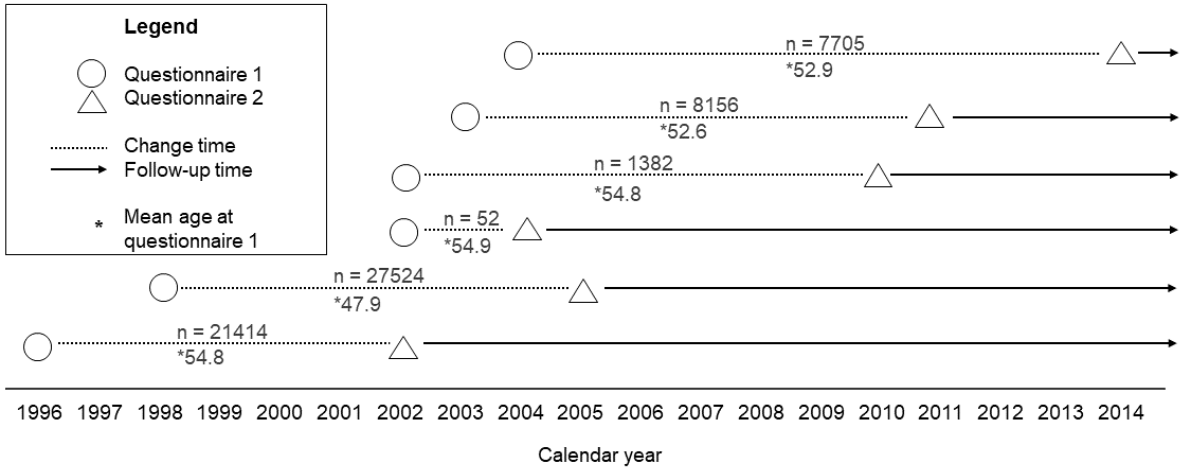


Figure 2. Timing of data collections and start of follow-up for the analytical sample. Norwegian Women and Cancer Study (n = 66 233).

Paper III study sample

The sample for Paper III included NOWAC participants who had responded to at least one eight-page questionnaire (n = 101 316) before a diagnosis of breast [C50], colorectal [C18-C20], or lung cancer [C34] up until the last update from the Cancer Registry of Norway in December 2020 (n = 9 287). Women registered as deceased (n = 153) or emigrated (n = 1) prior or within the same month as cancer diagnosis were excluded. Women reported as deceased, but with a missing value for month of death were excluded (n = 38). The analytical sample therefore consisted of 5032, 2468, and 1594 women diagnosed with breast, colorectal, and lung cancers, respectively.

3.3 Exposure assessment

The HLI score represented the exposure measure in Papers I-III. The intention of the HLI score was to capture a multifactorial exposure measure of a gradient of overall healthy

lifestyle as defined by the inclusion of selected evidence-based risk factors. In Papers I and III, a single HLI score assessed at a defined baseline provided the exposure. The closest questionnaire prior to diagnosis was used to obtain baseline information in Paper III. In Paper II, HLI score change provided the exposure measure and was defined as the difference between the HLI score at baseline and follow-up.

3.3.1 Constructing the Healthy lifestyle index (HLI) in NOWAC

The HLI was constructed by combining five lifestyle factor components – physical activity level, BMI, smoking, alcohol intake, and dietary habits. Given the availability of information on each eight-page NOWAC questionnaire, this HLI enabled participants to be assigned a score at baseline timepoints for Papers I, II, and III, and at the follow-up timepoint for Paper II. Figure 3 displays a summary matrix for the HLI score.

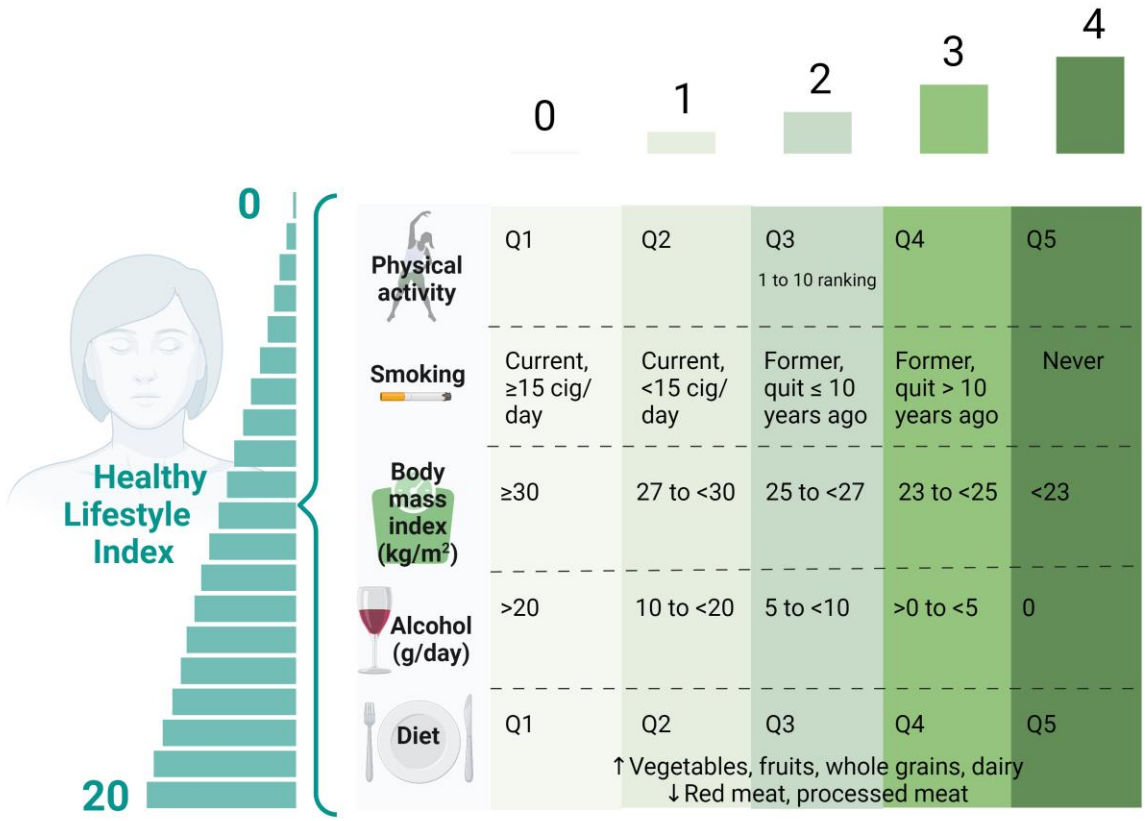


Figure 3. Healthy lifestyle index scoring matrix.

Abbreviations: Q, quintile; cig, cigarettes; g, grams.

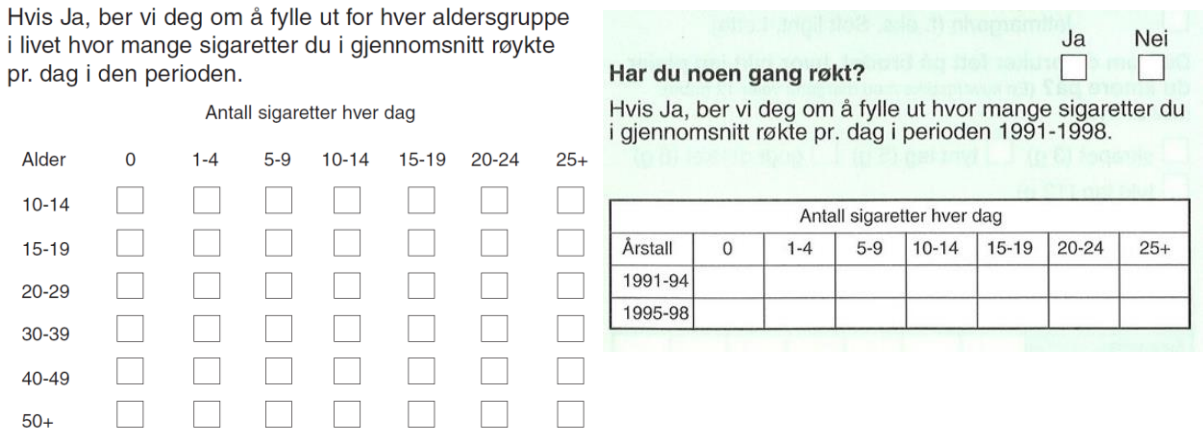
The HLI score ranged from zero to 20 points, where 20 points was considered the healthiest. The five lifestyle factors were equally weighted, and thus allocated between zero and 4 points each. Scoring cut-offs were based on a combination of scientific evidence of cancer risk factors and population distributions (percentiles) where natural and absolute cut-offs were not feasible. For each lifestyle factor, the data collection tool and transformation into component scores for the HLI are described in detail below.

Physical activity

Physical activity was measured using the Norwegian physical activity questionnaire (NOPAQ). In NOPAQ, participants ranked their current physical activity level on a scale from 1 to 10, where 1 was specified to mean "very low" and 10, "very high" physical activity level. Participants were asked to consider their self-assessed physical activity level as a measure of all physical activity, including domestic work, occupation, exercise, and other activities, such as going for a walk. The validity of NOPAQ has been previously assessed, observing that the tool provides a valid ranking of global physical activity among Norwegian adult women (113). To transform the NOPAQ scale to a five-level component of the HLI, five ordinal categories were created by computing quintiles, which were scored 0 to 4 from least active to most active.

Smoking

Questionnaires across waves and mailings differed in the number and types of questions on smoking habits. All participants were asked if they had ever smoked and if they smoked currently every day. Obtaining information on smoking intensity differed most drastically, employing a matrix query of the average number of cigarettes smoked daily either within four-year age ranges or three-year calendar periods (Figure 4).



Hvis Ja, ber vi deg om å fylle ut for hver aldersgruppe i livet hvor mange sigaretter du i gjennomsnitt røykte pr. dag i den perioden.

Antall sigaretter hver dag

Alder	0	1-4	5-9	10-14	15-19	20-24	25+
10-14	☐	☐	☐	☐	☐	☐	☐
15-19	☐	☐	☐	☐	☐	☐	☐
20-29	☐	☐	☐	☐	☐	☐	☐
30-39	☐	☐	☐	☐	☐	☐	☐
40-49	☐	☐	☐	☐	☐	☐	☐
50+	☐	☐	☐	☐	☐	☐	☐

Ja ☐ Nei ☐

Har du noen gang røkt?

Hvis Ja, ber vi deg om å fylle ut hvor mange sigaretter du i gjennomsnitt røkte pr. dag i perioden 1991-1998.

	Antall sigaretter hver dag						
Årstall	0	1-4	5-9	10-14	15-19	20-24	25+
1991-94							
1995-98							

Figure 4. Excerpts from NOWAC questionnaire on smoking intensity.
Left: first mailing for participants enrolled in 2003 (n = 13 950). Translated from Norwegian: "If yes [smoked more than 100 cigarettes in lifetime], we ask that you fill in how many cigarettes, on average, you smoked per day for each age group". Right: second mailing in 1998 for participants enrolled in 1991 (n = 38 184). Translated from Norwegian: "Have you ever smoked? [yes/no] If yes, we ask that you fill in how many cigarettes, on average, you smoked per day in the period 1991-1998".

For current smokers, current smoking intensity was defined by matching participants' age at the time of responding to the questionnaire to the corresponding age group in the matrix or by the most recent four-year calendar period. Ever smokers who reported not being current smokers were considered former smokers. Former smokers provided the age at which they quit smoking.

The HLI smoking component score was based on smoking status, time since smoking cessation if a former smoker, and smoking intensity if a current smoker. Never smokers received the highest score, 4. Participants who reported being former smokers were divided into those who quit smoking ≥ 10 years ago – receiving a score of 3 – and those who quit smoking < 10 years ago – receiving a score of 2. Participants who reported being current smokers were divided into those smoking < 15 cigarettes/day – receiving a score of 1 – and those smoking ≥ 15 cigarettes/day – receiving a score of 0.

Body mass index

BMI was calculated as $\frac{kilograms}{meters^2}$, using self-reported body weight (kg) and body height (cm). Lower BMI corresponded to higher HLI BMI component scores. Under- to normal weight

received the highest scores (BMI <23 = 4, 23 to <25 = 3), while overweight (25 to <27 = 2, 27 to <30 = 1) and obesity (≥30 = 0) received the lowest scores.

Diet and alcohol

The FFQ invited NOWAC participants to recall average eating habits over the past year and to record how often they consumed listed food items and alcoholic drinks within defined response categories. Since the FFQ was created for NOWAC, it contained food items commonly found in Norwegian stores and traditional Norwegian foods. Food items ranged from raw/unprocessed (ie. carrots, cabbage) to prepared/processed foods (ie. sausages, chocolate pudding). As an example, an excerpt from the FFQ as shown in Figure 5, participants were asked "*How often do you eat potatoes (boiled, baked, mashed)?*" to which response categories were "*don't eat/rarely eat potatoes*", "*1-4 times per week*", "*5-6 times per week*", "*once per day*", "*twice per day*", "*3 times per day*", or "*4+ times per day*"

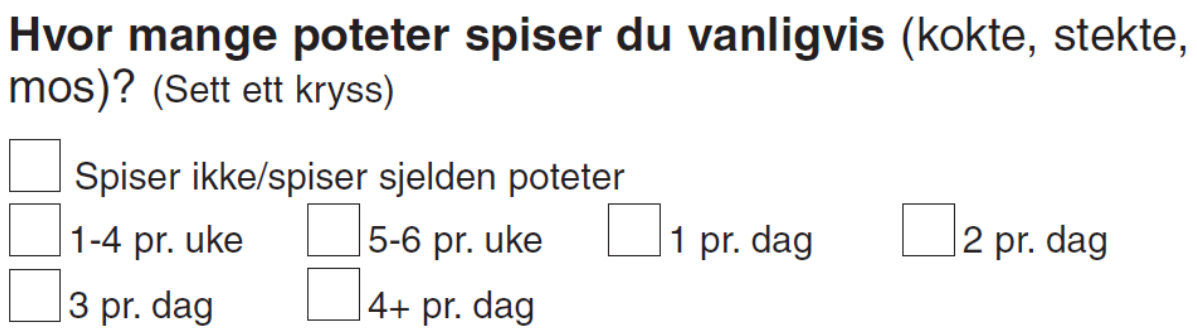


Figure 5. Excerpt from the Norwegian Women and Cancer Study food frequency questionnaire. First mailing for women enrolled in 2003 (n = 13 950).

Participants could report on the number of alcoholic drinks consumed per week from options including beer, wine, and spirits. In FFQs mailed in 2003 and after, an extra drink category – liqueur/fortified wine – was included.

Daily intake of food groups and ethanol, in grams, and energy, in kilojoules were calculated using an analysis programme developed at the Department of Community Medicine, UiT The Arctic University of Norway. The Norwegian food composition table provided reference conversions for frequencies and portions (114). Alcohol intake was scored according to absolute daily grams of ethanol cut-offs: no alcohol intake = 4, >0 to <5 g/day = 3, 5 to <10 g/day = 2, 10 to <20 g/day = 1, >20 g/day = 0).

Dietary habits were assessed by a diet-specific score that ranged from 0 to 18, with 18 considered the healthiest dietary pattern. It included six, equally-weighted food groups. Whole grains, fruits, vegetables, and dairy contributed positively to the score with higher intake. Red meat and processed meat contributed positively with lower intake. Daily grams for each food group was standardised by energy intake, where daily grams of intake was divided by daily energy intake. The energy-standardised food groups were each categorised into quartiles and scored from 0 (lowest quartile) to 3 (highest quartile). Scores from each food group were summed and then transformed for inclusion in the HLI by categorising the 0-18 diet-specific score into quintiles scored from 0 (lowest quintile) to 4 (highest quintile).

3.3.2 HLI score change

HLI score change was measured as the difference between the HLI score at Q1 and Q2, computed as:

$$Q2\ HLI\ score - Q1\ HLI\ score$$

Negative values therefore represented overall lifestyle worsening, while positive values represented overall lifestyle improvement. Both HLI score change as a continuous measure and in categories (≤-3, -2, -1, 0, 1, 2, ≥3) were examined. Where the HLI established score cut-offs were based on percentiles, the distributions from Q1 were used to determine absolute cut-offs when scoring participants at Q2.

3.4 Outcomes

The linkage of NOWAC participants to the Norwegian national registries through the Norwegian personal number supplied all endpoint information used for Papers I-III. The National Population Register provided information on date of death and emigration. Cause and month/year of death was provided by the Cause of Death Register. The cancer diagnosis according to the International Classification of Diseases, 10th Revision (ICD-10) codes and date of cancer diagnosis were provided by the Cancer Registry of Norway.

Paper I

Incidence was the event of interest in Paper I. Primary sites of cancer invasion with over 300 cases in the sample were investigated as outcomes. As such, breast cancer [C50], colorectal cancer [C18-20], lung cancer [C34], endometrial cancer [C54], ovarian cancer [C56],

pancreatic cancer [C25], and kidney cancer [C64] were included. Breast, endometrial, and ovarian cancers were restricted to postmenopausal women due to 1) known etiological and risk factor profile differences between pre- and postmenopausal statuses and 2) inadequate numbers of premenopausal cancer cases to investigate separately.

Paper II

Paper II engaged in a broader focus of cancer incidence, investigating all lifestyle-related cancers as identified by the *IARC Monographs* and *IARC Handbooks* Working Group (115). Four lifestyle-related cancer subgroups were specified, also by IARC, including alcohol-related, tobacco-related, obesity-related, and reproductive-related cancers. Table 2 presents the ICD-10 codes included in each exposure-related cancer subgroup. Site-specific analysis was also conducted for breast and colorectal cancer incidence due to their high incidence in the Norwegian population and in NOWAC.

Table 2. Exposure-related cancer subgroups defined by IARC.

Shaded regions denote cancer diagnosis belonging to the exposure-related cancer subgroup.

		Exposure-related cancer subgroups				
ICD-10 code	**Site**	Lifestyle	Alcohol	Tobacco	Obesity	Reproductive
C01	Upper aerodigestive	■	■	■		
C02		■	■	■		
C03		■	■	■		
C04		■	■	■		
C05		■	■	■		
C06		■	■	■		
C07		■	■	■		
C09		■	■	■		
C10		■	■	■		
C11	Pharynx	■	■	■		
C12		■	■	■		
C13		■	■	■		
C14		■	■	■		
C15	Esophagus	■	■	■	■	
C16	Stomach	■		■	■	
C18	Colorectum	■	■	■	■	
C19		■	■	■	■	
C20		■	■	■	■	
C22	Liver	■	■	■	■	
C23		■	■	■	■	
C24		■	■	■	■	
C25	Pancreas	■		■	■	
C31	Accessory sinus	■		■		
C32	Larynx	■	■	■		
C33	Trachea	■		■		
C34	Lung	■		■		
C50	Breast	■	■	■	■	
C51	Vulva	■				■
C52	Vagina	■				■
C53	Cervix	■		■		■
C54	Uterine	■			■	■
C55		■			■	■
C56	Ovarian	■		■	■	■
C57	Other female genital organs	■				■
C58		■				■
C64	Kidney	■		■	■	
C65		■		■	■	
C66		■		■	■	
C73	Thyroid	■			■	
C67	Bladder	■		■		
C90	Multiple myeloma	■			■	
C92	Acute myeloid leukaemia	■		■		

Paper III

Paper III investigated survival among women diagnosed with breast, colorectal, and lung cancer separately, in terms of all-cause and site-specific cancer mortality.

3.5 Covariates

Other available variables that were important for describing the study sample, were potential confounders or effect modifiers, or of specific interest for subgroup analysis were considered covariates of interest. All sociodemographic, reproductive-related, and hormone-related variables were self-reported in the questionnaires. Clinicopathological covariates were obtained from linkage to the Cancer Registry of Norway.

Sociodemographic and family health covariates

Year of birth and dates for baseline questionnaire and follow-up questionnaire; as well as month of cancer diagnosis, death, and emigration were used to calculate age at baseline questionnaire, cancer diagnosis, death, and emigration; and, in addition for Paper II, age at follow-up questionnaire. Education was considered a proxy for SEP and was assessed as the number of years of schooling. Family history of breast cancer in a first degree relative was relevant for all analyses focused on women diagnosed with breast cancer. This included analysis where the event of interest was breast cancer in Papers I and II; and in the analysis of survival among women diagnosed with breast cancer in Paper III.

Reproductive- and hormone-related covariates

Several reproductive and hormone-related covariates were relevant for models related to breast cancer incidence in Paper I and II, ovarian and endometrial cancer incidence in Paper I, reproductive-related cancer subgroup incidence in Paper II, and mortality among women diagnosed with breast cancer in Paper III. This included age at menarche, oral contraceptive use, hormone replacement therapy use, parity, cumulative breastfeeding time, and age at menopause. Women were considered postmenopausal if they reported their age at menopause or when they were ≥ 53 years old, as defined in the Million Women Study (116).

Clinicopathological characteristics of the cancer tumour

Clinicopathological characteristics of cancer tumours were important for the analysis of mortality among women diagnosed breast, colorectal, and lung cancer in Paper III. Breast cancer cases were classified according to the TNM Classification of Malignant Tumours (I, II, III, IV, unknown) (117), while colorectal and lung cancer cases were classified according to the Surveillance, Epidemiology and End Results (SEER) programme summary staging (local, regional, distant, unknown) (118). Molecular subtype and hormone receptor status information for BC tumours included oestrogen receptor status (ER) (positive if ≥ 10% reactivity until January 2012, thereafter positive if ≥ 1 % due to treatment protocol changes in Norway), progesterone receptor status (PR) (positive if >= 10%), and human epidermal growth factor receptor 2 (HER2) (positive if >10% tumour cells stained by immunochemistry).

3.6 Statistical analysis

Time-to-event analysis was the statistical focus in Papers I-III. To best understand the nature of the associations between the HLI score/change and events of interest, the HLI score was modelled in several ways, including as a continuous variable, categorical variable, and with restricted cubic splines (RCS). Statistical hypotheses were tested two-sided, allowing a Type I error rate of 5%. All analyses were conducted in R Studio and R Versions between 2019 and 2023 – the latest being RStudio Version 2023.03.0 with R Version 4.2.3 (119).

3.6.1 Cox proportional hazard regression models

We used Cox proportional hazard regression (Cox regression) to model the time-to-event of interest and estimate hazard ratios (HR) with 95% confidence intervals (CI). In Papers I and II, age (years) was used as the underlying time metric. In Paper III, time under study was instead used as the time metric and age was introduced in the model as a covariate. Follow-up started at the time of the baseline questionnaire in Paper I, at the second questionnaire in Paper II, and at cancer diagnosis in paper III. Participants were considered at-risk until the event of interest occurred (Paper I and II: cancer incidence; Paper III: death) or they were censored due to emigration (Papers I-III), death (Papers I and II), or end of the study period (Paper I and II: December 2018; Paper III: December 2020), whichever occurred first. Figure 6 provides a summary of the relevant timepoints in the study design for the Cox regression models in Papers I-III.

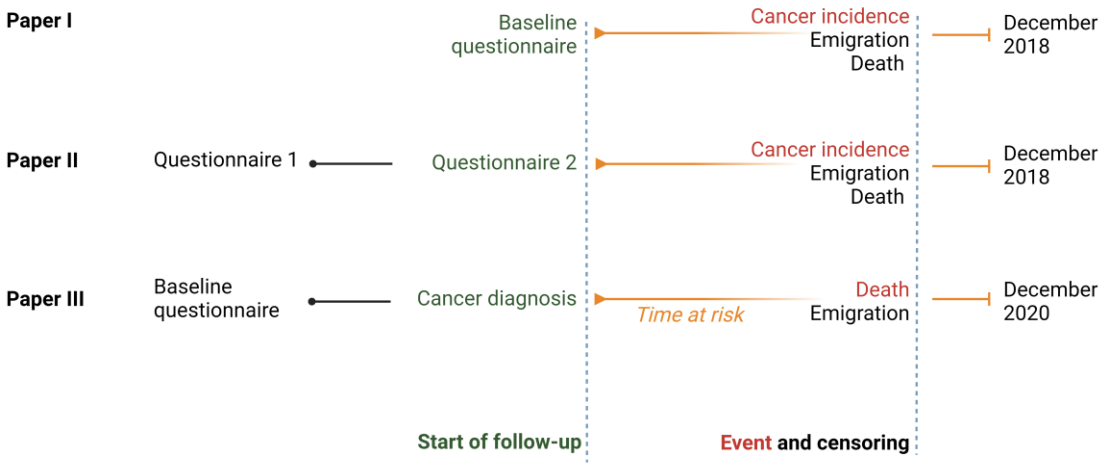


Figure 6. Relevant timepoints in the study designs for Papers I-III.

The proportional hazards assumption was tested by visual inspection of Kaplan-Meier survival plots and by assessing the correlation between Schoenfeld residuals and survival time. Evidence of correlation ($p < 0.05$) indicated the presence of time-dependent covariates, which were handled by including the covariate(s) in question as stratified terms in the Cox model (120).

3.6.1.1 Examining nonlinearity

The resulting estimate for continuous exposure terms in Cox regression – HLI score in Papers I and II, and HLI score change in Paper II – assumes linear risk differences across the exposure scale. To examine potential departures from linearity, the HLI score in Paper I and the HLI score change in Paper II was modelled using restricted cubic splines (RCS). RCS is described in greater detail in Section 5.1.7.2. We selected the number of knots by minimising the AIC from three-, four-, or five-knot exposure terms on a per-outcome basis and checked for overfitting. The placement of knots was determined by percentile, as recommended by F. E. Harrell (121). We used Harrell's package in R, *rms* (122), to model continuous predictors as RCS terms with Cox regression.

Nonlinearity in the associations was assessed in two ways. Firstly, we relied heavily on visual inspection of the resulting plots of the exposure modelled with RCS against the HR and 95% CIs. Secondly, we conducted a hypothesis test using the likelihood ratio test to compare the

goodness of fit between the model where the HLI score was modelled as a log-linear term (nested) and where the HLI score was modelled as a RCS term (full).

3.6.1.2 Competing risks

In Paper III, standard Cox models for site-specific mortality were forced to censor participants for other death events that may have presented competing risks, in that they disqualified the occurrence or changed the probability of the event of interest (123). So as to not assume the independence of competing events, we estimated subdistribution hazard ratios (SHR) and 95% CIs with the subdistribution hazards regression model proposed by Fine and Gray (124) for site-specific cancer models. We used the R package, *cmprsk*, developed by Gray (125) to perform the analysis.

3.6.1.3 Multivariable adjustment of confounding

Fully-adjusted models provided the main estimates of association in Papers I-III. Adjustment factors were chosen on a per-outcome basis, a priori. They were included in the multivariable models if they were considered to be associated with the exposure and a cause of the outcome based on previous literature. The quality of the instrument in measuring the potential confounder was also considered.

In Papers I-III, education (Papers I and II: continuous years; Paper III: ≤9, 10-12, 13-16, ≥17 years) and height (cm) were controlled for in every model. For models where the outcome was breast, endometrial, ovarian, or reproductive-related cancer incidence (Papers I and II), or the sample was women diagnosed with breast cancer (Paper III), the following variables were included in the multivariable analysis: age at menarche (continuous years), menopausal status at entry (pre-, postmenopausal), breastfeeding (continuous) cumulative months, parity (0, 1, ≥2), hormone replacement therapy use (never, former, current), and oral contraceptive use (never, ever). Specifically, breast cancer incidence outcome models and breast cancer case samples were additionally adjusted for family history of breast cancer in a first degree relative (yes/no). In Paper III, baseline period (1996-98, 2002-05, 2011-14) was modelled as a stratification variable in the Cox models such that separate hazard functions were estimated for each strata and HRs optimized for each strata were then fitted.

3.6.2 Sensitivity analysis

Sensitivity analyses of various forms were conducted in Papers I-III to test the robusticity of the main associations by identifying possible mechanisms of bias and to provide depth to the main associations. Since sensitivity analyses are inherently data-driven, a greater than ± 5% difference between main and comparison estimates was used as a rule-of-thumb for defining a meaningful difference, where:

$$\% \text{ difference} = \frac{HR_{main} - HR_{comparison}}{HR_{main}} \times 100$$

To identify if single lifestyle factors were largely driving the associations, we excluded each lifestyle factor, in turn, from the HLI to produce reduced-HLI scores consisting of four factors. In such reduced-HLI models, the excluded factor was introduced into the multivariable model as a covariate for adjustment. We then estimated the associations between each reduced-HLI score and the outcomes throughout Papers I-III. Since the full-HLI ranged from 0 to 20 and the reduced-HLI ranged from 0-16, the magnitude of HRs representing the risk difference for a 1-point increase were not comparable. In Paper II we addressed this by standardising the reduced-HLI scores to per 1-SD unit as the units from the main analyses were also 1-SD. In Paper III, we multiplied the reduced-HLI scores by a factor of 1.25 to standardise the units to the 1-point increase equivalent on the full-HLI.

Interaction terms were selected a-priori on the grounds of plausible hypothesis and previous findings. The presence of interaction was tested by comparing the fit of models with and without the interaction term using the likelihood ratio test. P-values < 0.05 indicated the model with the interaction term explained the variation in the data better than the model without and the term was kept in the final models. In Paper II, we tested interaction terms for age at Q1 and age at Q2 with HLI score change. We tested if a correlation was present between HLI score change and years between Q1 and Q2 using Pearson's correlation coefficient. In Paper III, we tested for interaction by the time between baseline questionnaire and cancer diagnosis (prediagnostic interval) since this ranged considerably and had a normal distribution.

To avoid bias in the estimates due to reverse causation, the first years two years of follow-up time were excluded in Papers I-III. In Paper III, we excluded women who were diagnosed with cancer less than two years after the questionnaire and women who were deceased less than five years after the baseline questionnaire.

As a crude validation of the exposure-related cancer groupings used as outcome measures in Paper II, we estimated associations for the incidence of all non-lifestyle-related cancers (i.e. all remaining cancers not considered lifestyle-related).

3.6.3 HLI score compared to single factors

As an extension to the analysis conducted in Papers I-III, the following was conducted for the purpose of contributing to the discussion on the use of the HLI in the thesis. To compare the HLI score with single lifestyle factors in terms of goodness-of-fit and parsimony (126), AIC values were compared between the two models. The AIC values were computed for cancer incidence models in all sites examined in Paper I using two sets of explanatory variables: i) baseline HLI score and ii) five single factors (physical activity level, BMI, smoking score (0-4), daily alcohol intake (g/day), and diet score (0-18)) simultaneously included in each model.

3.6.4 Handling of missing data

The fraction of missing information (FMI) for the HLI score was considerable for the samples in Papers II and III at 34% and 28%, respectively, and also not negligible in Paper I at 16%. The approach to missing data in Papers I-III was two-fold – analysis by complete-case data conducted under the missing completely at random (MCAR) assumption and multiple imputation by chained equations (MICE) data conducted under the missing at random assumption (MAR) (127). Complete-case analysis was performed by listwise deletion and was handled automatically by R.

MICE was conducted using the R package, *mice* (128). Fully conditional specification was employed, whereby each incomplete variable was modelled iteratively by a series of multivariable regression models. In Papers I and II, one MICE model was executed with cancer incidence specified as the outcome. In Paper III, six MICE models were executed such that analytical MICE datasets were generated separately for women diagnosed with breast, colorectal, and lung cancer by all-cause and site-specific mortality. Paper I generated 20 datasets with 10 iterations. Papers II and III generated 100 datasets with 10 iterations. Exposure variables, all covariates, binary outcome variable, and the Nelson Aalen cumulative hazard estimator for the event of interest were included in the MICE models.

Visual inspection of plots displaying mean HLI score in Papers I and III, and mean HLI score change in Paper II against iteration number for each imputed dataset was performed to assess

whether MICE models converged. Descriptive statistics for each variable that had missing values was compared between observed and imputed values to assess any large differences in central tendency. Regarding the estimation of associations using MICE datasets, parameter estimates and standard errors from each imputed dataset were averaged according to D. B. Rubin's rule to account for within and between imputation variance (129).

3.7 Ethical considerations and data management

All waves of NOWAC enrolment took place between 1991 and 2007, during which lawful consent was given by participants for the collection and storage of their information, as well as linkage to the Cancer Registry of Norway, Mammography Registry of Norway, and the Norwegian Cause of Death Registry. NOWAC received approval from the Regional Committee for Medical Research Ethics in Northern Norway (REK Nord) in 2008 (REK Nord 141/2008.) as ethical laws in Norway changed at the time. Papers I and II are based on the ethical approval obtained up to 2008. The study on which Paper III is based received separate approval from REK (REK Nord 487111) due to further evolved ethical laws. NOWAC participants have the right to view their own registered information as well as to correct any mistakes. The can also withdraw their consent at any time (130).

At study initiation, data handling was considered ethical under the Norwegian Data Protection Authority. Data obtained from NOWAC for Papers I-III were transferred from the central databases of the NOWAC management through a secure file sender. No information that could be used to personally identify participants or pseudo-anonymous information was provided. The data were considered "yellow" by the UiT The Arctic University of Norway's Guidelines for classification of information (131). As such, data was not shared publicly and was only handled by the first author of Papers I-III.

4 Results

4.1 Paper I – HLI and incidence

The aim of Paper I was to examine the associations between combined lifestyle factors, assessed through the HLI score, and the incidence of common cancer types among women in Norway. Paper I included 96 869 women in the analysis. There were 81 554 women with complete data available for descriptive presentation and complete-case analysis.

The mean age at baseline was 52 years across all HLI groups. The reported mean years of education was lowest in the 0-5 HLI group (11.6) and increased slightly with higher HLI groups (6-10 HLI: 11.9; 11-15 HLI: 12.3; 16-20 HLI: 12.9). By design, women in higher HLI groups reported, on average, higher physical activity level, lower BMI, less frequent current and former smoking behaviour, least alcohol intake, and with the lowest diet score.

The median follow-up time was 20 years and 15.2 postmenopausal years during which 3 397 postmenopausal breast, 1 213 colorectal, 1 006 lung, 807 postmenopausal endometrial, 425 postmenopausal ovarian, 284 pancreatic, and 268 kidney cancer cases occurred. We observed negative dose-response associations between 1-point HLI score increments and the incidence of postmenopausal breast (HR: 0.97, 95% CI: 0.96-0.98), colorectal (HR: 0.98, 95% CI: 0.96-1.00), lung (HR: 0.86, 95% CI: 0.84-0.87), postmenopausal endometrial (HR: 0.93, 95% CI: 0.91-0.95), pancreatic (HR: 0.92, 95% CI: 0.89-0.95), and kidney (0.94, 95% CI: 0.91-0.97) cancer incidence. No association was observed for postmenopausal ovarian cancer (HR: 0.99, 95% CI: 0.6-1.02).

We observed indications of a nonlinear association for postmenopausal breast cancer incidence, whereby HRs for women with higher baseline HLI scores were stronger than the linear model estimated. Nonlinearity was also present for lung cancer incidence, whereby the negative association was less pronounced among women with HLI scores greater than 15 compared to women with lower HLI scores. The associations for colorectal, postmenopausal endometrial, pancreatic, and kidney cancer did not demonstrate departures from linearity.

4.2 Paper II – HLI change and incidence

Paper II aimed to estimate the association of changes in combined lifestyle, as assessed by HLI score change between two timepoints, with lifestyle-related cancer incidence, including

separate examination of several exposure-related cancer subgroupings, breast cancer and colorectal cancer incidence. There were 66 233 women included in the analyses after handling of missing values by MICE and 44 404 women available for complete-cases analysis. At Q2, representing the start of follow-up, the mean age was 58.2 years old. The mean HLI score change was -0.2 (range: -11 – 14). When assessing the distribution of HLI score change in categories (≤ -3, -2, -1, 0, +1, +2, ≥ +3), the majority of women presented zero HLI score change (17.2%). The mean time between Q1 and Q2 was 7 years (range: 2 – 11).

After a median follow-up time of 14.2 years, there were 6 384 incident lifestyle-related cancer cases, within which there were 3 512 alcohol-related, 2 931 tobacco-related, 3 385 reproductive-related, 2 384 breast, and 839 colorectal cancer cases. A 1-SD increment in HLI score change was associated with lower incidence in all cancer groupings, specifically 7% lower incidence for lifestyle-related (HR: 0.93, 95% CI: 0.90-0.96), 4% lower incidence for alcohol-related (HR: 0.96, 95% CI: 0.91-0.99), 8% lower for tobacco-related (HR: 0.92, 95% CI: 0.88-0.96), 6% lower for obesity-related (HR: 0.94, 95% CI: 0.91-0.98), and 10% lower for reproductive-related (HR: 0.90, 95% CI: 0.84-0.98) cancers. We observed 4% lower breast cancer incidence for a 1-SD increment in HLI score change (HR: 0.96, 95% CI: 091-1.01), although our data was also compatible with no association. HLI score change was not associated with colorectal cancer incidence (HR: 0.98, 95% CI: 0.90-1.07).

We observed from categorical analysis of the associations that negative HLI score changes of 3 or more were associated with higher incidence of lifestyle-related cancer compared to no HLI score change ($HR_{\leq -3 \text{ v. } 0}$: 1.16, 95% CI: 1.05-1.27). Greater positive HLI score changes compared to no HLI score change demonstrated an increasingly protective trend. However, we observed that the data was compatible with no association for the greatest positive HLI score change category of 3-points or more compared to no HLI score change ($HR_{\geq +3 \text{ v. } 0}$: 0.93, 95% CI: 0.84-1.03). There were no indications of nonlinearity in the associations.

4.3 Paper III – Prediagnostic HLI and survival

The aim of Paper III was to estimate the associations between combined prediagnostic lifestyle factors and survival among women diagnosed with breast, colorectal, and lung cancer. Women diagnosed with lung cancer had a lower median HLI score (11) compared to women diagnosed with breast (13) and colorectal (13) cancers. The mean ages at the time prediagnostic lifestyle assessment for women diagnosed with breast, colorectal, and lung

cancers were 55.5, 59.0, and 58.1 years, respectively. The mean ages at the time of breast, colorectal, and lung cancer diagnoses were 63.4, 68.5, and 67.7 years, respectively. The distribution of stage at diagnosis differed across cancer sites. Breast cancer was most frequently diagnosed with TNM stage I (52%). Colorectal cancer was most frequently diagnosed with SEER regional stage (54%). Lung cancer was most frequently diagnosed with SEER distant stage (44%). Women diagnosed with breast cancer were most frequently postmenopausal (91%), ER positive (89%), PR positive (68%), and HER2 positive (87%).

There were 5 032 breast, 2 468 colorectal, and 1 594 lung cancer cases included in the analysis after MICE and 3 241 breast, 1 574 colorectal, and 1 005 lung cancer cases available for complete-case analysis. Women diagnosed with breast, colorectal, and lung cancer were followed-up for a median duration of 9.8, 7.1, and 5.9 years, respectively. Among breast cancer cases, there were 912 all-cause deaths of which 509 were BC deaths. Among colorectal cancer cases, there were 902 all-cause deaths of which 679 were CRC deaths. Among lung cancer cases, there were 1 094 all-cause deaths of which 961 were LC deaths. Women diagnosed with lung cancer had the lowest 5-year survival rate at 29%, followed by women diagnosed with colorectal cancer (67%) and breast cancer (90%).

The following main results are associations for every 1-point increment in prediagnostic HLI score. Among women diagnosed with breast cancer, there was a 6% lower risk of all-cause mortality (HR: 0.94, 95% CI: 0.92-0.97) and 3% lower risk of breast cancer mortality (HR: 0.97, 95% CI: 0.94-1.00), also compatible with no association. Among women diagnosed with colorectal cancer, there was a 3% lower risk of all-cause mortality (HR: 0.97, 95% CI: 0.95-1.00) and 2% lower risk of colorectal cancer mortality (HR: 0.98, 95% CI: 0.95-1.01), although both estimates were compatible with no association. The was no indication of association between prediagnostic HLI score and all-cause (HR: 1.00, 95% CI: 0.98-1.02) or lung cancer mortality (HR: 1.00, 95% CI: 0.98-1.03), among women diagnosed with lung cancer. Site-specific cancer mortality estimated with SHRs from competing risks analyses were within ± 5% of HRs.

4.4 Comparing HLI to single factor models

In the complementary analyses that were not part of Papers I-III, we compared cancer incidence models explained by i) baseline HLI score and ii) lifestyle factors as separate explanatory variables entered into the model simultaneously. The AIC values from models

with single lifestyle factor predictors were, for all cancer sites, lower compared to cancer incidence modelled with the HLI score (Table 3).

Table 3 Akaike Information Criterion (AIC) values for cancer incidence sites modelled with HLI score compared to single lifestyle factors

	HLI score	**Single factors**
Postmenopausal breast	72486.99	71169.28
Colorectal	25149.03	24643.93
Lung	20444.73	19038.27
Postmenopausal endometrial	17097.21	16517.23
Postmenopausal ovarian	8884.213	8719.047
Pancreatic	5757.353	5596.564
Kidney	5598.731	5451.772

5 Discussion

5.1 Discussion of methods

5.1.1 Study design

Papers I-III employed a prospective cohort design to address the aims. Contrasted with this observational design, the randomised control trial – an experimental design – is considered the gold standard for causal inference. However, in the context of the exposures and outcomes studied, where the number of observations required for adequate statistical power is large and there is a long and undefined latency period between lifestyle factor exposure and cancer occurrence, an experimental design would prove unfeasible. Further, there is unlikely to be a modern setting in which an experimental design could ethically test the impact of several lifestyle factors on health outcomes.

NOWAC is a large prospective cohort with repeated measurements and with virtual complete follow-up for cancer diagnosis, death, emigration from Norwegian national registries. The temporality of the prospective design empowers causal inference as the exposure measurement precedes the outcome (132) and there is a known time interval between exposure and cancer onset (133). However, the long and undefined latency period of cancers, and thus long subclinical periods, can challenge the assumed temporal sequence where hypothesized exposures coexist with the subclinical phase of the outcome or the subclinical outcome causes the exposure, known as reverse causation (133). In Papers I-III, potential reverse causation was addressed by excluding the first two years of follow-up and comparing estimates to the complete follow-up time. We did not observe any differences. In Paper III, the variation in prediagnostic interval due to study design was an important aspect to consider. Among women diagnosed with lung cancer within five years after the baseline questionnaire, a healthier overall prediagnostic lifestyle was associated with higher all-cause and lung cancer mortality. There were no associations detected among women with prediagnostic intervals longer than five years. We suspected that subclinical lung cancer may have resulted in unintentional weight loss and thus misclassification of women with pre-clinical lung cancer in a healthier HLI range.

The opportunistic nature of new enrolment waves and follow-up measurements in NOWAC was key for increasing its sample from the original ~ 60 000 women recruited in 1991/92 and for having access to repeated measurements. However, the introduced variation in the time

between baseline and follow-up measurements for observations in Paper II risked being strongly associated with HLI score change. After testing, we observed no correlation of the time between baseline to follow-up with HLI score change. Further, when examining HLI score change transformed to a rate, those with the shortest and longest baseline to follow-up times had unrealistic magnitudes of change. As such, we did not standardise HLI score change to time.

5.1.2 Information bias

Information bias is a systematic error that results from the placement of participants into incorrect exposure or outcome categories. Observation, classification, and measurement bias are all terms used to describe this occurrence of misclassification (134). The implications of misclassification as a threat to validity depends on the degree of misclassification, whether the misclassification is *nondifferential* or *differential* (132), the number of categories and its cut-offs, the measurement error distribution, and the distribution of the variable in the sample (135). Nondifferential misclassification occurs when the exposure is misclassified to the same degree and direction between observations who experience and do not experience the event. Differential misclassification occurs when the exposure is misclassified differently between observations who experience the event and do no experience the event in terms of degree and/or direction (132).

Information bias in the exposure measure, HLI score, must consider possible misclassification in the five component lifestyle factors – physical activity level, BMI, smoking habits, alcohol intake, and dietary habits. Universal to all measurements used, differential misclassification due to recall bias was likely not present due to the collection of measurements before cancer diagnosis.

5.1.2.1 Physical activity level

A validation study reported that the use of the NOPAQ scale was a valid tool for ranking individuals' physical activity level in NOWAC (Spearman's rank correlation coefficient: 0.36-0.46, $p < 0.001$) (113). On a population level, NOPAQ was capable of distinguishing physically from not physically active groups. Further, use of a simple measure may have minimised non-response compared to more detailed questions of high cognitive burden (14). The NOPAQ scale was categorised based on quantile cut-offs for the physical activity component of the HLI score, thus likely preserving the ranking of individuals. Assuming

nondifferential misclassification and no other systematic errors, error in the NOPAQ scale likely biased our estimates of association towards the null.

5.1.2.2 BMI

The underestimation of BMI through overreported height and underreported weight is common when using self-report instruments (136) and was found to be present in NOWAC (137). In a validation study of BMI in NOWAC, underreporting of weight occurred in all categories of BMI, although occurred most frequently among women who were overweight and more severely among women who were obese. Differences between self-reported and directly measured BMI were small in absolute terms and the ranking of individuals was adequate (137). Although nondifferential misclassification of BMI can be assumed, the categorisation of BMI into five categories and the misreporting of BMI being dependent on the BMI category, the estimates of association may have been biased towards or away from the null (138).

5.1.2.3 Smoking habits

Several features of smoking habits were of interest, including smoking status – never, former, and current smoker – as well as intensity for current smokers and timing of last exposure (time since smoking cessation) for former smokers. The questions on smoking in NOWAC's questionnaires have not been formally validated. Underreporting of smoking status with self-report instruments is well-documented (139). However, there can be considerable variation in the misclassification in smoking status groups across single studies (140). Nevertheless, our access to follow-up lifestyle information in Paper II was able to provide some indication. There were 2 426 (of 66 233 total) participants that had reported being ever smokers at baseline and never smokers at follow-up in the data cleaning for Paper II. This is nonsensical and likely represents misclassification of smoking status at follow-up. The majority (97%) reported being former smokers at baseline. Although some misclassification was present, its impact was less severe than if most of the misclassification had arisen from participants who reported current smoking at baseline. We attempted to correct this discrepancy in Paper II by replacing never smoking with former smoking at follow-up. Nevertheless, some participants may have been current smokers at follow-up. Moreover, differences in the style of questioning used across study waves and follow-up questionnaires and period effects may have resulted in varying degrees of misclassification across different types of questionnaires.

With this tendency to underreport smoking among current and former smokers, several smoking categories in the HLI, questionnaire differences, period effects, and assuming nondifferential misclassification, the estimates may have been bias towards or away from the null (135).

5.1.2.4 Alcohol intake and dietary habits

FFQs aim to measure average, habitual diet and often must sacrifice the precision attained from measuring short term diet with less precise information that captures long term diet (141). Due to a suspected long induction period for diet on cancer and cancer survival, use of the FFQ is probably appropriate compared to other feasible instruments. Alcohol intake in NOWAC was found to be substantially underreported in a comparison of food group measures collected through the FFQ and repeated 24-hour dietary recalls (142). However, the ranking of intakes was considered to be good. In terms of food groups included in the HLI, report of fruits and vegetables was lower and dairy was higher in the FFQ compared to 24-hour dietary recalls. There was no statistically significant difference for whole grains, meats, and processed meat. However, the study did not separate red meat from poultry (142). A test-retest study conducted on the NOWAC FFQ reported that the reproducibility of the FFQ responses was comparable to the reproducibility seen in other similar instruments (143).

Energy intake estimation is a known challenge with FFQs due to differing interpretations of food types and portions, as well as non-response (141). In NOWAC, missing frequencies were imputed with no intake and missing portion sizes were imputed with the smallest portion available on the question, which likely contributed to the lower energy intake reported in the FFQ compared to the 24-hour dietary recalls (142). In total, we expect the information bias present in the variables obtained from the FFQ section of the NOWAC questionnaire to have been nondifferential. However, due to the multiple forms of categorisation, questionnaire dissimilarities, and period effects, the estimates of association may have been biased towards or away from the null.

5.1.2.5 Outcomes

Misclassification of cancer diagnosis most likely did not occur due to the high accuracy of classification at the Cancer Registry of Norway and near complete follow-up (144). In addition, the incidences of cancer types in NOWAC are comparable to that of national figures (112, 145), suggesting successful data linkage to the Cancer Registry of Norway. The

Norwegian Cause of Death Registry has near complete coverage of the Norwegian population and it has been estimated that there is 98% coverage of all deaths (146). Classifying the cause of death is challenging, even with the best of registry coverage. Studies assessing the quality of cause of death registries ranked the Norwegian Cause of Death Registry as "medium" (147), "medium-high" (148), and in the "best" group but below other Nordic registries (149). It is unlikely that the misclassification of outcomes was dependent on the HLI score. As such, we can expect a reduction of precision in the estimates and likely, no bias from misclassification of outcomes (150).

5.1.3 Confounding

Confounding is a threat to causal inference and occurs when a third variable – the confounder – is non-causally or causally associated with the exposure and causally associated with the outcome, but does not lie as an intermediary on a causal pathway from the exposure to the outcome (132). Associations that are confounded can be either further or closer to the null compared to the true, causal association and can lead to spurious conclusions.

Potential confounders were selected a priori in Papers I-III based on current literature and consideration of factors that could fulfil the definition of a confounder. Further, inclusion of a potential confounder in the final statistical models took the quality of the confounder variable and whether it impacted the estimate of association into consideration. For example, pre-existing cardiovascular disease is likely associated with overall lifestyle before cancer diagnosis and mortality. In NOWAC, the prevalence of these diseases were self-reported and their validity has not been assessed. As a sensitivity analysis, cardiovascular diseases were adjusted for in the multivariable models and yielded estimates were no different from estimates prior to adjustment. Thus, cardiovascular diseases were not included in the final multivariable models.

Incomplete adjustment may have occurred and can be classified into residual and unmeasured confounding. There is the possibility of residual confounding from the covariates included in the final models due to crudeness of confounder categories or misclassification of the confounder (132). Most potential confounders were measured and modelled as continuous variables, which would minimise residual confounding. Nevertheless, confounders may be nonlinearly associated with the outcome, resulting in incomplete adjustment (151). Misclassification of confounders is possible. For example, a study assessing the long term

maternal recall of the duration of breastfeeding among elderly US women found there to be substantial recall bias (152). If a similar recall bias were present in NOWAC, there would have been inadequate adjustment for cumulative breastfeeding months in breast cancer and reproductive-related cancer models. However, as also mentioned in Section 5.1.2, the prospective design increases the plausibility that the misclassification was only nondifferential.

Due to the inclusion of physical activity level, BMI, smoking habits, alcohol intake, and dietary habits in the HLI score, these single lifestyle factors were not considered potential confounders. However, the relationships between the lifestyle factors are complex and it is possible that single lifestyle factors could have served as proxies for an unmeasured confounder. For example, physical activity level may be an indicator for mental wellbeing, where mental wellbeing is a potential confounder in the association between prediagnostic HLI score and lung cancer survival (153). In addition, some covariates may have served as suboptimal proxies for the theoretical confounder. For example, years of education was treated as a proxy for SEP. However, the SEP concept, in terms of how it relates to health outcomes, probably differs across cultures (154) and there are age-period-cohort effects in such variables (155). The construct of SEP is thus considerably more complex than years of education and there was likely residual confounding present.

It is likely that there were several unmeasured factors that are known causes of the outcomes and associated with the exposure. It is known that some cancer types are heritable, such as breast, colorectal, bladder, pancreatic, and skin cancer (10). However, information on family history of cancers other than breast cancer was not available in NOWAC. Cancer treatment is a major determinant of cancer survival (3). Although available from the Cancer Registry of Norway, cancer treatment was not included in the analysis due to the inconsistency of cancer treatments and protocols across the study period and a high fraction of missing values. We suspected that adjustment for cancer stage served as a proxy for treatment given standard protocols for cancer treatment in Norway (156). Although, unwarranted variation – i.e. disparities in healthcare utilisation that cannot be accounted for by patient needs or preferences (157) – is likely present, and partially due to known inequitable access to treatments in Norway (158).

In a simulation study, Fewell et al. (159) observed that unmeasured and residual confounding had a lower impact on estimates when they were more correlated with adjusted confounders.

While we cannot know the extent of residual confounding, it is likely that the unmeasured confounders were partially correlated with adjusted confounders. Taken together, bias from the true causal estimate due to incomplete adjustment of confounding depends on the culmination of the individual biases introduced by all residual and unmeasured confounders. Due to the varying types of misclassification and differing potential for bias towards and away from the null, it is difficult to speculate on the overall over- or underestimation of associations due to confounding bias in Papers I-III.

5.1.4 Selection bias

Selection bias occurs when the study sample differs systematically from the target population such that the association between the exposure and outcome is different in the sample population compared to the entire source population (133). The temporality of the prospective cohort designs in Papers I-III – where study participants were enrolled and exposure was measured before the outcome event – should largely diminish the differential selection of participants according to event status.

Another source of selection bias in the estimates due to differential selection can be considered confounding in prospective cohort studies (133). It is often the case that the sample and target population have different distributions of the exposure, confounders (measured and unmeasured), and events, which infringes upon the *representativeness* of the sample, and possibly its estimates, to the source population. Education level in NOWAC, which is correlated with the HLI score, is slightly higher compared to the national average (145). As such, there may have been underrepresentation of lower HLI scorers in NOWAC. This would have impacted the accuracy of the descriptive statistics and selection bias would have arisen for estimates of absolute risk since they are dependent on accurately estimating the distribution of characteristics in the sample. However, we only estimated HRs, which is a relative measure of risk. With complete adjustment and consideration of effect modification, differences in the distribution of exposure characteristics between the sample population and source population, selection bias would not have been present (133, 160). Although, as discussed in Section 5.1.3, unadjusted confounding was likely present.

Differential loss to follow-up, informative censoring, or attrition bias refer to the same form of selection bias that can afflict prospective cohort studies that use time-to-event analysis. It occurs when individuals who are right-censored, especially due to mortality, have different

risks of experiencing the event compared to those who remain until the study end (132). In causal inference, this can be interpreted as there being common cause of the attrition and event. Due to the shared risk factors for cancer and mortality, it is possible that differential loss to follow-up biased the estimates towards null. Competing risk analysis, discussed in 5.1.7.2, was used in Paper III in order to address informative censoring bias.

5.1.5 Summary of internal validity

The estimated associations were vulnerable to information bias, confounding, and selection bias. Information bias was present in the data for each lifestyle factor included in the HLI and likely represented nondifferential misclassification. However, there were many instances of categorisation and non-random misclassification. The models were not fully adjusted as there was likely residual and unmeasured confounding. There was likely a presence of selection bias during enrolment due to unmeasured or residual confounding, which hindered the establishment of comparable exposure groups through statistical analysis. As the impact of epidemiological biases are complex, it is not possible to deduce the extent to which the estimates of association may have been under- or overestimated (138).

5.1.6 Representativeness

NOWAC is likely representative of Norwegian women in the source population (age 30-70). A study was conducted in 2007, assessing the representativeness of NOWAC (145). The random sampling of Norwegian women through the national registry with near complete coverage increased the likelihood of those invited being representative of Norwegian women at the time, thus minimising sampling bias. The response rates were approximately 60% from age groups 30-34 to 55-59 years, but was 45% for the eldest age group, 65-70 years. There were no differences in oral contraceptive use, parity, and years of education between responders and non-responders. As stated previously, responders reported more years of education compared to the source population, although there was not a large difference. The study also compared incidence rates of cancer between NOWAC women and national figures provided by the Cancer Registry of Norway for the period 1991-99, finding them to be nearly identical.

Representativeness is not a prerequisite for the generalisability of estimates, although it can be helpful (160). Generalisation of the estimates to other populations of women and time should be approached with caution. The lack of interpretability of the HLI score restricts the

inferences that can be drawn for various HLI scorers as the same score can represent several combinations of lifestyle factors. See 2) in Section 5.2.6.1 for more on these challenges. However, due to the similar distribution of several exposure and outcome related characteristics between NOWAC women and the source population and the ability to adjust for several potential confounders, it is likely that the estimates from the Papers I-III reflect those of the source population.

5.1.7 Chosen statistical methods

5.1.7.1 Considerations related to Cox regression

Choice of time scale

Cox regression was used in Papers I-III to estimate associations and requires a continuous time scale to model rates of events. The underlying time scale was age with left truncation on the entry age in Papers I and II. Time under study was the underlying time scale in Paper III, with multivariable adjustment of entry age. There is a lack of consensus on the optimal time scale to use in Cox regression (161, 162, 163). However, entry age was handled in both instances, which, according to an empirical example published by Chalise et al. (163), is most crucial when defining a time scale. Further, they observed minimal differences in regression coefficients between the two methods used in Papers I-III, but suggested that using time under study with adjustment of entry age as it tends to perform equally or better compared to the age scale with left truncation model (163). This motivated our use of time under study as the underlying time scale in Paper III.

Proportional hazards assumption

The PH assumption was upheld in all models in Paper I and Paper II according to visual inspection of Kaplan-Meier survival plots and by assessing the correlation between Schoenfeld residuals and survival time. In Paper III, the test of Schoenfeld residuals revealed that the PH assumption was violated by stage at diagnosis. We performed stratification on stage at diagnosis, which allowed different baseline hazards to be estimated within each stratum but estimation of the same regression coefficient (164). Further testing revealed that the non-proportionality had been corrected.

Competing risks

In Paper III, death from specific causes (i.e. breast, colorectal, and lung cancer) were events of interest. As such, death from other causes was treated as a censoring event in site-specific cancer mortality models. However, this may have violated the assumption of noninformative censoring – that instantaneous risk for the event occurring is the same for participants who are still being followed as for participants who have been censored – and biased the results (165). As there is no method to test the assumption that censoring is noninformative (166), we treated informative censoring events as competing risks and applied the commonly used Fine and Gray methods (124) in a parallel analysis. Here, the cumulative incidence function estimated the marginal probabilities for each competing event and then modelled with SH regression to estimate SH ratios (SHR) and 95% CIs.

While SHRs do not assume that censoring is noninformative, they are not necessarily unbiased. As P. Allison (167) explained, in keeping observations that experience the competing event in the risk set, they can bias the estimate of association for a predictor down for the event of interest by artificially lowering the probability of the event of interest occurring. This bias may be more severe compared to cause-specific HRs. However, as several authors note, SHRs may be more useful in prediction settings, whereas cause-specific HRs could be better suited to explanatory/aetiological/causal research questions (165, 167, 168, 169). As the aim of Paper III was to estimate the rate of the occurrence of death given the prediagnostic HLI score, the research question of interest was explanatory. We therefore made inferences based on site-specific cancer mortality HRs. We observed that SHRs were nearly identical to the HRs estimated from site-specific cancer mortality hazard models likely because the associations between the exposure and event of interest, and between the exposure and competing events, were both weak.

5.1.7.2 Restricted cubic splines

In any regression model where the predictor is continuous, the default regression coefficient predicts a linear association to the outcome. Specifically, a unit increase in the predictor is associated with the same incremental risk difference of the outcome. While the coefficients of continuous predictors are easy to interpret and communicate, nonlinear associations are common. Thus, predicting linearly makes a strong assumption about the true association and can lead to poor explanations of the outcome (121).

Breaking continuous variables into categories is a common approach for addressing potential nonlinearity. However, there are several unfavourable properties of categorisation, including, but not limited to, a loss of statistical power, assuming the association is constant throughout the entire interval of the category, assuming an even distribution of the sample throughout each category, the arbitrary nature of chosen cut-points, and temptation to use cut-points that minimize the p-value (121).

Splines, also known as piecewise polynomials, offer a well-powered solution to allow and detect departures from linearity in a continuous exposure while maintaining their continuous nature (121, 170). They are a series of polynomials that are connected across intervals of the continuous predictor, and are thus insensitive to category cut-offs. We applied a class of splines, called restricted cubic splines (RCS) to examine nonlinearity in the associations between HLI score and cancer incidence (Paper I), HLI score change and cancer incidence (Paper II), and the interaction of prediagnostic interval with HLI score and cancer survival (Paper III). RCS, like other piecewise polynomial models, splits the continuous predictor into a series of intervals bound by *knots*. Cubic functions are fitted between knots such that the first and second derivative (slope and rate of change of the slope) are continuous at the knots, resulting in a smooth curve across all values of the predictor (121). RCS are considered *restricted* as they force linear functions to be fitted before the first knot and after the last knot as cubic splines can be poorly behaved at the tails (171).

RCS are highly flexible, whereby the number and position of knots can be specified. With this flexibility, there is a risk of over- and underfitting (121). Knot position has been found to be minimally important, thus prompting the general recommendation to place knots at the percentiles, ensuring equal distributions of observations within each interval (171). Conversely, it is the number of knots that is crucial for the models. Harrell (121) recommends four or five knots for sample sizes where $n > 100$ and, in a data-driven approach, choosing the number of knots that minimises the AIC value. We combined both approaches, starting with four knots and assessed overfitting by visual inspection with the substantive knowledge that there should not be sharp inflections across the HLI score. We compared AIC values of models with three, four, and five knots, and considered the use of the model that minimised the AIC value. Where overfitting was suspected, the number of knots was decreased.

The modelling of the HLI score or HLI score change with RCS was not only advantageous for the aforementioned reasons, but it enabled elegant graphical presentation of risk

differences and uncertainty across the HLI scale. In terms of the presentation of estimates, HRs and 95% CIs can be extracted for specific values of the predictor when modelling with RCS, so it would have been possible to retain the power and validity of using a continuous HLI score while presenting concrete risks (at least for complete-case estimates). However, we chose to present categorical estimates of HLI score or HLI score change for two reasons. First, we were unable to model RCS with MICE data and so presented categorical estimates to make use of the larger and, potentially, more valid sample of MICE (discussed in Section 5.1.8). Second, the presentation of categorical estimates is ubiquitous and it provided easily comprehensible estimates comparing risk among those healthiest and least healthy according to the HLI.

5.1.8 Missing values

Steps were taken in the design of the questionnaires to minimise the amount of missing data in NOWAC – for example, careful attention to the phrasing and formatting of the questionnaire. Further, some missing values have been imputed in NOWAC under certain assumptions (i.e. missing food frequencies were imputed with zero intake). However, missing data is inevitable with self-reported data in population-based studies and can produce biased estimates depending on the method for handling missing data, the proportion of missing data, assumed mechanism of missingness (172). Little and Rubin have provided a framework for classifying three broad mechanisms of missingness (173). Missing data can be: i) missing completely at random (MCAR), whereby the probability of missing is independent of the observed or unobserved data; ii) missing at random (MAR), whereby the probability of missing is independent of the unobserved data but dependent on the observed data; and iii) missing not at random (MNAR), whereby the probability of missing is dependent on the unobserved data and the observed data.

The exclusion of observations that have a missing value for any of the variables included in the statistical model – known as *complete-case analysis* – is considered a default technique and is commonly performed in epidemiological studies. In situations of MCAR, complete-case analysis is unbiased. However, where the FMI is higher, there is an automatic loss of precision in the estimates due to the reduced sample size and biased estimates if under MAR (129). Of note, there are some non-MCAR mechanisms that cannot be neatly categorised into MCAR/MAR/MNAR under which complete-case analysis is considered a valid approach – namely when the missingness is independent of the outcome and dependent on the

explanatory variables in the analytical model of interest (174). In prospective cohort studies, it is plausible that missingness is independent of the outcome since the measurement of predictors precede the outcome. However, complete-case analysis can produce biased estimates when missingness in the explanatory variables is caused by other factors that independently affect the outcome. Since knowing the mechanism of missingness would require us to know the missing values, it is impossible to ascertain this (175). As such, we considered other methods to handle missing values as a parallel approach.

Single imputation – replacing missing values with, for example, the median, mean, or mode of observed values – is somewhat common as it enables observations to be retained. However, this technique can underestimate the standard errors of the estimate and produce biased estimates (176). Multiple imputation by chained equations (MICE) is an increasingly used approach to missingness and is valid under MAR. MICE employs a series of multivariable regression models that take into account the associations between observed data variables to create multiple predictions for each missing value (177). It retains the sample size as single imputation would, but considers uncertainty in the imputed values, thus not inflating precision of the estimates. Further, the MICE model is meant to be compatible with the analytical model (i.e. cancer incidence regressed on HLI score) such that it appreciates incompletely observed associations (178). It has been recommended that when complete-case and MICE analysis are plausibly valid, MICE is preferrable due to its greater efficiency (175).

MICE was performed in Papers I-III due to substantial FMI under the MAR assumption. There is no consensus on the optimal number of datasets, *m*, that should be generated to reduce sampling variability from MICE (176). While as low as 3-5 datasets have been shown to be adequate, it has also been recommended that a minimum of 20 datasets be generated for each MICE model (176). As a rule of thumb, White et al. (179) recommend that *m* be at least the FMI percentage. However, they encourage $m > 100$ if possible. We generated 20 datasets in Paper I. In Paper II and III we had access to greater computational power and generated 100 datasets. All MICE models performed 10 iterations. Non-convergence of MICE models with 10 iterations has never been observed in simulation models according to White et al. (179), except in rare cases. We checked for convergence as recommended by van Buuren and Groothuis-Oudshoorn (128), observing no abnormalities.

Choosing the set of MICE predictors was not a simple process due to the multiple levels of variable transformation, particularly for the HLI score. A missing value on physical activity

level would render the HLI score missing, while all other components were observed. We considered it too crude to impute directly on the HLI score and preferred imputation on score components. The highest FMI was observed for the physical activity score (~10%) in Papers I-III. When the number of components is small relative to the sample size, as in Papers I-III, including the single components of sum scores as predictors has been shown to produce unbiased results (180).

There were several measures taken to prevent missing values from being replaced with implausible values. For example, imputing on higher level variables given constraints by design applied in Paper III (i.e. observed current smokers could not be imputed with a smoking score of 2 to 4 as higher scores were reserved for never and former smokers) or by passive imputation for transformed variables (i.e. the feedback loop produced when imputing on height, weight, and BMI) (181). Paper II required that we consider the implausibility of ever smokers at baseline being imputed with never smoker scores at follow-up. By included baseline component scores and component change scores as predictors, the chances of an ever smoker at baseline (smoking score 0-3) being imputed as having a positive smoking score change that would represent never smoking at follow-up (smoking score 4) was virtually zero as we ensured nonsensical follow-up smoking status was corrected in the observed data. After MICE, we checked all MICE datasets for implausible sequences of baseline to follow-up information. We subsequently computed follow-up HLI scores within each MICE dataset. A wide range of predictors were included in the MICE models to increase the likelihood that missing values were predicted in a way that would correct for MAR.

Use of MICE required high attention to detail and the number of specifications made it clear that, while flexible and powerful, MICE could also be a source of bias if specified incorrectly. When we tested a number of MICE specifications in preliminary analysis (i.e. comparing 20 to 100 datasets in Paper II; comparing 5 to 10 iterations in Paper III), there was virtually no variation in the estimates suggesting the bias from misspecification may be small. However, this cannot be known for certain given the many specifications that were not tested and the ultimate inability to observe missing data.

Further, White et al. (175) highlighted non-MCAR mechanisms in which MICE analysis produced biased estimates, while complete-case analysis was unbiased. We encountered barriers to using MICE data for some types of analysis, such as the modelling of RCS. As

such, the parallel approach of MI and complete-case data were important for the validity of the results given that the missingness mechanism was unknown.

The estimated HRs were similar between complete-case and MICE analysis, which could suggest that the missingness mechanism at work was MCAR. However, due to the substantial FMI, the presence of MAR is likely. It is possible that MICE was incorrectly specified and produced similarly biased estimates to complete-case. Still, MAR may not have extensively affected the data or may have only been related to missingness in the HLI score and covariates, rather than cancer incidence or survival.

5.1.9 Defining the HLI

The goal of the HLI was to define a simple multifactor exposure measure that placed NOWAC participants on a scale from least healthy to healthiest lifestyle, based on current scientific evidence for causes of cancer. We did not restrict the definition of a healthy lifestyle to a single public health body. Rather, we considered many sources of lifestyle recommendations and summaries of the scientific evidence, such as findings reported in the WCRF/AICR 2018 CUP (23), IARC Monographs (31, 32), IARC 2020 World Cancer Report (10), and recommendations from the Norwegian Cancer Society (182). Several lifestyle factors emerged as preventive, including being physically active, maintaining a healthy body weight, avoiding smoking, avoiding alcohol intake, consuming a varied diet consisting of high intake of fruits, vegetables, and wholegrains, sleep, and mental health. However, NOWAC only collected information on the first five factors. Most HLIs constructed in studies on other cohorts have typically included these five factors (183, 184, 185), which was considered an advantage for comparing findings.

We chose to operationalise increasingly healthy behaviour with a granular scale that allocated five ordinal categories to each component such that partially exposure could be detected. Several HLIs have allocated binary scores to each component (i.e. meeting the recommendation, 1; or not, 0). However, the dichotomisation of components results in a loss of information, power, and has limitations for understanding risk variation across a more detailed exposure spectrum (170, 183). Current evidence acknowledges the dose-response association of increasing exposure to low levels of physical activity (10), BMI (normal weight through to obese) (23), duration and intensity of smoking (31), and alcohol intake (10, 23)

with the higher risk of many cancer sites. We wanted the HLI score to reflect the risk gradient suggested by current scientific evidence.

The design of the HLI is challenged by the competing acts of defining risk by i) exposure level while simultaneously acknowledging that exposures are differentially related to risk across cancer sites, ii) cut-offs for score categories may not accompany risk differences for some cancers or at all, and iii) the risk differences for single factors may not be linear. It was also essential that the HLI operated within the confines of the extent and type of information available from the NOWAC questionnaires and the distribution of lifestyle factors among NOWAC women so as to not have underpowered score categories. For example, NOWAC participants who were categorically underweight (BMI < 18.5 kg/m^2) were assigned to the healthiest BMI category along with those considered normal weight due to the low proportion of underweight. A J-shaped association between continuous BMI and cancer mortality has been suggested based on large cohort studies, whereby the underweight population has higher mortality compared to the normal weight population (186, 187). Whether this J-shaped association transfers to cancer risk is less defined and has been found to be vary across cancer sites in a large meta-analysis (188). If the HLI acknowledged potentially higher risk for underweight compared to normal weight, assigning underweight to a lower score (2, 1, 0) would also make strong assumptions about risk. We excluded those considered underweight in a sensitivity analysis to assess if the scoring of underweight participants biased the results to the null, observing no differences in estimates.

There were several options for creating a five-category gradient for the smoking component. "Pack-years" – product of the number of packs of cigarettes smoked per day and the number of years the person has smoked – is frequently used to study cumulative lifetime exposure to smoking (34, 62). While this is somewhat simple to compute, it only takes into account intensity and duration, not the time since quitting for former smokers. As it is well-known that there is a reduction in lung and upper aerodigestive cancer risks after smoking cessation and that this risk continues to decrease with longer durations of cessation (51, 189), we chose to recognize this directly in the score.

The amount of average daily alcohol intake was scored in order of highest, presumed to confer the greatest cancer risk, to lowest intake, in order to reflect the dose-response association to cancer risk (10). Due to overall low alcohol intake reported in NOWAC, establishing cut-offs based on recommendations would have resulted in a heavily right-

skewed distribution and few participants in the higher intake levels. Categories based on percentile cut-offs were considered. However, the cut-offs were nearly indistinguishable from an absolute perspective and thus may not have represented risk differences. We employed cut-offs based on grams of daily intake used in a previous EPIC study on a HLI and breast cancer incidence (190), which resulted in categories with a more even distribution of participants compared to cut-offs reflecting recommendations, and categories more reflective of real differences in alcohol intake compare to percentiles.

Consensus of dose-response associations for dietary habits is sparse. We included food groups where there was strong evidence, according to the WCRF/AICR 2018 CUP, for increased or decreased cancer risk for women. The diet score was designed such that high scorers would be defined by a diverse diet that was high in fruits, vegetables, wholegrains, and dairy products, while low in red- and processed meat. However, it is unknown whether this HLI diet score actually represents the healthiest diet pattern with respect to cancer risk, as exemplified by the number of dietary quality indices in existence (191).

We considered weighting the HLI components to address the issue of the different strengths of association for lifestyle factors. However, this would have limited the analysis to very few outcomes as it would be logical to develop weights based on the risk for single cancer sites and types. Further, in a guide to the development of health measurement scales, Streiner et al. (192) argued that weighting of scale components seldom results in different ranking compared to equally-weighted scales. Consistent with this, Jiao et al. (185) observed that their weighted HLI did not perform better compared to their equally-weighted HLI with respect to pancreatic cancer risk despite the known, strong effects of smoking and BMI. More importantly, use of a weighted HLI would undermine the aim of the study – to assess if a simple measure of overall lifestyle is associated with cancer-related outcomes. To maintain simplicity, the HLI was designed to weight the five components equally; however, at the expense of possible attenuated estimates.

5.1.9.1 HLI compared to single factor models

According to the AIC values comparing cancer incidence modelled with the HLI score to single lifestyle factors. In NOWAC, the modelling of single factors as explanatory variables fit the data better than the HLI score. There could be two main reasons for this:

1) Categorisation of continuous lifestyle factors and assigning ordinal scores reduced statistical power
2) Single factor models appreciated that lifestyle factors were differently associated with the outcome, which the HLI averaged

5.2 Discussion of main results

5.2.1 Paper I – Lifestyle and incidence

A healthier overall lifestyle, as measured by the HLI score, during adulthood was protective against the occurrence of cancer in several common sites among women in Norway. As expected, the associations varied across cancer site and were strongest for cancers sites that are known to be strongly associated with smoking, including cancers of the lung and pancreas. There were clear associations for postmenopausal breast, postmenopausal endometrial, and kidney cancer incidence. We observed a weak, protective association for colorectal cancer incidence that was also compatible with no association. Postmenopausal ovarian cancer incidence was not explained by overall lifestyle. Associations modelled with restricted cubic splines (RCS) offered a novel approach to observing multifactor exposure measure relationships.

5.2.1.1 Breast cancer

The observed negative association between the HLI score and the incidence of postmenopausal breast cancer is largely consistent with the literature. The results were similar to those in two previous studies also using HLI score exposures with 0-20 score ranges. They observed 3% and 4% lower postmenopausal breast cancer incidence for every 1-point increment in a Canadian (193) and US cohort (184), respectively. Other publications that have examined the association between a multifactor exposure and breast cancer incidence, where the unit of measurement was challenging to compare, observed that a healthier overall lifestyle was associated with lower postmenopausal breast cancer incidence (190, 194, 195, 196, 197) and lower overall breast cancer incidence (198).

We were unable to directly compare our results to a smaller Norwegian study (n = 17 145) that investigated combined unfavourable lifestyle factors and postmenopausal breast cancer incidence as they estimated associations in subgroups of hormone replacement therapy use

(103). We did not perform hormone replacement therapy use subgroup analysis as it was outside the scope of Paper I and we assumed that it had confounding rather than interactive effects in the association. However, it has been shown that hormone replacement therapy has differential effects across categories of BMI (199) and this should be appreciated in the future.

When allowing for nonlinearity in the associations, there appeared to be greater risk reductions for an increase in HLI score for women with high HLI scores (>13) compared to women with lower HLI scores at baseline. There were no other publications examining nonlinearity in the association to offer comparison. Further, it is not realistic to interpret the nature of dose-response from percentile categories in the exposure as presented by other studies using similarly granular HLIs (184, 193). However, if the magnitude of risk differences is greater among women who are already healthy, this would suggest that establishing overall healthy lifestyles early in life is crucial for all women.

Breast cancer is currently the most common cancer in Norway and incidence trends show a steady increase, with no signs of plateau (12). Given that the increasing trend is largely attributable to the age group 60-79 according to national figures (12) and over 60% of NOWAC women were diagnosed with breast cancer in this age range, adherence to healthy lifestyles during adulthood are likely to be beneficial on a population level.

5.2.1.2 Colorectal cancer

We observed a protective association between HLI score and colorectal cancer incidence, with risk reductions only apparent for women with the highest HLI scores. While the linear association was weak and compatible with no association, RCS detected strong reduced risk for women with an HLI score over 17 compared to an HLI score of 5.

Other prospective cohort studies that employed similarly constructed multifactor exposures observed comparable linear associations in the EPIC cohort (183), two cohorts from the United States (US) consisting of health professionals (200), and a cohort from Denmark (201). In a sample of Norwegians invited for colorectal cancer screening (n = 6315), a more favourable lifestyle defined by a HLI was associated with lower risk of screen-detected advanced colorectal neoplasia (104). Although a comparison of estimates is made challenging due to differences in HLI construction and underlying lifestyle patterns, the association in Paper I could be interpreted as weaker compared to the screened sample in Knudsen et al

(104). Another large cohort from the US that employed a score measuring adherence to ACS recommendations observed stronger risk reductions compared to our results (198). A systematic review and meta-analysis on the WCRF/AICR score and cancer outcomes reported a 14% lower colorectal cancer risk for every 1-point increment on the 0-7 scale based on 10 studies (102). Four of the seven points in the WCRF/AICR score are distributed to dietary factors. If the relative importance of diet in the association is greatest compared to the remaining lifestyle factors, a multifactor exposure that is weighted heavily towards diet would be more strongly associated with colorectal cancer risk compared to an equally weighted score.

There were no indications that certain lifestyle factors were particularly responsible for the association with colorectal cancer incidence. Rather, the results from the sensitivity analyses supported the importance of the clustering and possible synergy of multiple healthy lifestyle factors to reduce risk (94). Associations between single lifestyle factors – including physical activity (48), BMI (27), smoking (202), alcohol, and dietary food groups and nutrients (203) – and colorectal cancer incidence have not been detected as strong in studies conducted in NOWAC. However, overall dietary patterns have not been investigated with respect to cancer in NOWAC. Greater adherence to the Mediterranean Diet Score, Healthy Eating Index (2005 or 2010), and Dietary Approaches to Stop Hypertension (DASH) have been associated with lower colorectal cancer risk in several cohort studies (204). It is possible that the NOWAC HLI diet score did not adequately capture the mentioned dietary patterns due to the selection of food groups and energy adjustment in addition to information bias present. Nevertheless, research should also search beyond lifestyle factors for explaining colorectal cancer risk as incidence is high in Norway and the association for overall lifestyle, as estimated by the HLI score, was weak, only conferring risk reductions for women with the healthiest overall lifestyles.

5.2.1.3 Lung cancer

The HLI score was strongly, negatively associated with lung cancer incidence, with a plateau in incidence reduction among the highest HLI scorers. The sensitivity analysis clarified that this association was explained by the smoking component. In agreement with this finding, Kabat et al. (198) observed that the ACS score, which did not include a smoking component, was not associated with lung cancer incidence. The relative importance of smoking in the HLI score association is consistent with the well-established dominance of smoking in lung cancer

risk (19). Although findings from other studies have suggested that some dietary factors, such as fruits and vegetables, and physical activity are associated with reduced lung cancer risk, while red meat, processed meat, and alcohol intake are associated with increased lung cancer risk, these studies have tended to show weak evidence for an association and there is a strong possibility of being confounded by smoking (23).

Lung cancer is the second most diagnosed cancer among women in Norway. The impact of smoking prevalence on cancer incidence is clear in Norway, demonstrated both by analysis of risk (as in Paper I) and the parallel time trends in smoking exposure and lung cancer incidence (12, 205). The high incidence and poor survival outcomes for lung cancer is concerning and points to the importance of reducing smoking in the population.

5.2.1.4 Postmenopausal endometrial cancer

Paper I supported a protective role of overall healthy lifestyle in lowering the risk of postmenopausal endometrial cancer, which is consistent with previous studies (101, 193, 198, 206). Previous HLI studies observed that a 1-point increment in HLI score was associated with 5% and 6% lower postmenopausal endometrial cancer incidence in a Canadian (193) and US cohort (206), respectively, when employing a 0-20 point HLI exposure similar to ours.

Examining the association after the exclusion of BMI from the HLI was indicative of both its role as a major contributor to the protective association for the HLI score and as a strong risk factor for postmenopausal endometrial cancer incidence. In the meta-analysis performed by the WCFR 2018 CUP, they reported 50% higher risk of endometrial cancer for every 5 BMI units (23). It is well-established that factors which are suspected to increase the exposure to oestrogen increase the risk of endometrial cancer (10). Since postmenopausal women derive most of their circulating oestrogen from oestrogen synthesis in adipose cells, it is sensible that a higher BMI would be associated with higher risk (10).

There was no indication of nonlinearity in the associations, supporting a consistent dose-response association between higher HLI score and postmenopausal endometrial cancer incidence. It has been reported that the positive association between BMI and endometrial cancer risk is nonlinear, whereby risk increases more steeply with higher BMI and the strengthening of risk occurs at around 40 kg/m^2 (23). However, as the BMI component categorised all participants with $kg/m^2 \geq 30$ together, this nonlinearity in the BMI association would not have been detectable.

Endometrial cancer is the most common gynaecological cancer in Norway (12) and in most high income countries (7). In total, BMI especially, but also other lifestyle factors combined appear to be associated with a reduced risk of postmenopausal endometrial cancer.

5.2.1.5 Postmenopausal ovarian cancer

There were no indications that the HLI score was associated with postmenopausal ovarian cancer incidence. This is consistent with findings from several other studies, who also observed null associations between HLIs of similar construct and postmenopausal ovarian cancer incidence in prospective cohorts from the US (184), Canada (193), and France (196). Additionally, prospective cohort studies that used the WCRF/AICR score (97) or the ACS score (198) were in agreement with our results by their null associations.

The removal of single components from the HLI did not elucidate any antagonistic behaviour or single lifestyle factors that may have opposed potentially protective behaviour of the remaining lifestyle factors. According to the WCFR 2018 CUP, BMI is the only lifestyle factor with any strong evidence for an association with ovarian cancer risk (23). However, it has been demonstrated that, in the NOWAC study, BMI is not associated with ovarian cancer incidence (27). BMI has been found to only be associated with ovarian cancer risk among users of hormone replacement therapy in a large pooled analysis (207). It is possible that we (and Da Silva et al. (27)) did not detect an association because of effect modification by hormone replacement therapy use (207). Subgroup analysis by hormone replacement therapy use was outside the scope of Paper I. However, this should be investigated in future studies to better elucidate the association between BMI, and potentially other risk factors, and ovarian cancer risk.

5.2.1.6 Pancreatic cancer

There was a strong, negative association between the HLI score and pancreatic cancer incidence, which appeared to be linear. A study conducted in the EPIC cohort observed a similar strength of association between a comparable HLI (0-20 score range) and pancreatic cancer incidence (208). Consistent with our study, they also observed that smoking was responsible for the HLI score association. Our results were also consistent with prospective cohort studies employing scores that did not include smoking, including the WCRF/AICR score (102) and the ACS score (198).

In the literature, smoking is the most well-established risk factor for pancreatic cancer (10). Our results support the dominance of smoking, but did not reveal any influence of other lifestyle factors. Body fatness has been identified as an additional important risk factor and ranked as having strong and convincing evidence for increasing the risk of pancreatic cancer by the WCFR 2018 CUP (23). As discussed in the IARC World Cancer Report 2020, the impact of obesity on pancreatic cancer risk is complex and may be confounded by type II diabetes mellitus (10). Both have been repeatedly associated with higher pancreatic cancer risk and are correlated. Moreover, type II diabetes mellitus can emerge as sequelae to early stage pancreatic cancer and can cause obesity (209). We tested whether comorbidities, including type II diabetes mellitus, impacted the estimates in preliminary analyses and detected no effect. However, BMI may not have been an adequate tool for measuring body fatness (210, 211), particularly when detecting an association that may be weak. In support of this, Naudin et al. (208) observed, in the EPIC study, that an association remained after the exclusion of smoking when waist-hip ratio was used instead of BMI as the measure of body fatness in the HLI. As a digestive organ, we would suspect that alcohol intake and diet should be risk factors for pancreatic cancer. However, our findings are consistent with the limited evidence reported for dietary factors and alcohol intake as risk factors for pancreatic cancer (23).

The annual incidence of pancreatic cancer has been stable between 2012 and 2021 among women in Norway (12). However, prognosis is extremely poor and incidence is fairly common with a cumulative risk of 1.2% for developing pancreatic cancer by age 80 among females in Norway. Population-wide healthy living, especially avoiding smoking and maintaining a healthy weight, should be a priority to reduce the burden of pancreatic cancer.

5.2.1.7 Kidney cancer

There was a strong, negative, and linear association between the HLI score and kidney cancer incidence. Results from a recently published study on a prospective cohort from the Netherlands support our observations, reporting a similar protective association when using a comparable 0-20 point HLI (212). Protective associations were also observed between greater concordance with recommendations as operationalised by the WCRF/AICR score (97) and the ACS score (198).

When BMI was removed from the HLI, the association attenuated, but remained protective. Our results support the findings from the WCRF/AICR 2018 CUP, who reported strong and convincing evidence that body fatness increases the risk of kidney cancer (23). They also listed the higher intake of alcoholic drinks as having probable evidence for increasing risk. Further, smoking has been identified as an important risk factor for kidney cancer risk (31). The remaining association after the removal of BMI from the HLI is indicative of the important, but perhaps smaller role of the remaining lifestyle factors.

5.2.2 Paper II – lifestyle change and incidence

The findings from Paper II indicated that overall lifestyle change, measured as the difference in HLI score between two time points, among Norwegian adult women was associated with the risk of lifestyle-related cancers. Specifically, greater positive HLI score change was associated with lower alcohol-, tobacco-, obesity-, reproductive-, and overall lifestyle-related cancer incidence in a dose-response manner. We observed that lifestyle worsening was more strongly associated with lifestyle-related cancers than lifestyle improvement when they were compared to stable lifestyles. Among women who experienced negative HLI score changes ≥ 3, lifestyle-related cancer incidence was 16% higher compared to those with no HLI score change. In contrast, lifestyle-related cancer incidence was 7% lower for women who experienced positive lifestyle changes ≥ 3 compared to those with no HLI score change, although also compatible with no association. When HLI score change was modelled with RCS, the greater effect of lifestyle worsening compared to lifestyle improvement was observable. However, the visualisation of the RCS clearly demonstrated the large uncertainty of the estimates. Through this uncertainty, a linear association was compatible with the data and clearly visible from inspection of the plots (Paper II, Fig. 2).

It is possible that the protective effect of overall lifestyle improvement on cancer risk is smaller and thus more challenging to detect compared to overall lifestyle worsening. Several studies on weight change have observed that weight gain is more strongly and precisely associated with higher breast (44, 45, 46) and obesity-related cancer risk (27) compared to weight loss. In addition, increased alcohol intake over a period of five years has been associated with higher breast cancer risk, while no association was observed for decreased alcohol intake (49). Alcohol cessation has been associated with lower risk for aerodigestive cancers, but not tested for alcohol increase (50, 51). However, the proportion of our sample that ceased drinking alcohol may have been low and thus not captured by the HLI score

change. The cumulation of these differential effects of improving and worsening may have been detected and reflected in the results of HLI score change.

Two studies investigating overall lifestyle change and cancer incidence were identified in the literature search. Our findings are largely consistent with those in Botteri et al. (105), who observed that overall lifestyle improvement compared to consistently overall unhealthy lifestyle was associated with lower lifestyle-related cancer incidence in a cohort of women from Sweden. They did not include analysis of women who worsened overall lifestyle compared to consistent overall healthy lifestyle. As such, we cannot compare our observation of the relative weak association for lifestyle improvement compared to worsening.

We did not observe that HLI score change was associated with breast cancer incidence, although there may have been a slight protective trend for positive HLI score changes. In a Swedish cohort, Botteri et al. (105), observed that women who had unhealthy lifestyles at baseline and improved to healthy lifestyles at follow-up, compared to women who had unhealthy lifestyle both at baseline and follow-up, had a lower risk of breast cancer. However, it is challenging to compare results due to major differences in categorisation of the HLI.

There were no indications from our results that HLI score change was associated with colorectal cancer incidence, nor were any trends identified. In contrast, Botteri et al. (99) recently observed that, in the EPIC cohort, there was a protective trend between the magnitude of positive HLI score change and colorectal cancer incidence among women. Nevertheless, their findings for women were also compatible with no association. The average time between baseline and follow-up in EPIC was similar compared to in NOWAC. The difference between EPIC and NOWAC estimates could be attributable to many sources, such as small differences in the construction of the HLI, smaller sample size in NOWAC compared to EPIC, and follow-up time which was shorter in EPIC compared to NOWAC. As mentioned in Section 5.2.1.2, associations between lifestyle factors measured at baseline or at multiple timepoints and colorectal cancer incidence have not been convincingly detected in NOWAC. In comparison with EPIC, it is therefore unsurprising that Paper II did not detect associations for lifestyle changes.

Based on previous knowledge of the protective association between adult baseline healthy lifestyle and cancer incidence, healthy lifestyle behaviours are encouraged by public health bodies. Our use of the exposure-related cancer groupings defined by IARC was purposeful as

it enabled us to study the association between overall lifestyle change and the incidence of cancer in many sites in a sweeping manner and it lent itself to a broader assessment of site-specific cancers that may share aetiological mechanisms. The potentially weak or slow protective effect of lifestyle improvement and potentially stronger deleterious effect of lifestyle worsening in adulthood highlights the importance of healthy lifestyle maintenance throughout adulthood to lower cancer risk. However, our results should not be used to negate individuals' and public health efforts to improve lifestyle among those who, for example, have unhealthy lifestyles. As we observed from Paper I, overall lifestyle, measured at one timepoint during adulthood, is an important risk factor for many of the most common cancers affecting Norwegian women, regardless of any changes in lifestyle that may have occurred before or after baseline. Further, the body of evidence on overall lifestyle changes and cancer risk is scarce and, among the studies that do exist, suggest that lifestyle improvements reduce cancer risk (99, 105).

5.2.2.1 Time window of change

Lifestyle changes in the calendar period

The observed changes in lifestyle behaviours are inseparable from cultural, political, and economic trends occurring during the calendar period 1996 to 2014 affecting women between the ages of 40 to 76, which represented the earliest baseline and latest follow-up measurement in Paper II. During this time, tobacco restrictions were enforced (213), which saw declines in the prevalence of smoking in the Norwegian population (205). Further, a report combining the results from repeated surveys reported that alcohol intake doubled between 1973 to 2004 among women age 15 and over (214). In Paper II, NOWAC women registered an increase in the HLI smoking score and decrease on the average HLI alcohol score. NOWAC women tended to experience a weight increase, which could be attributable to universally observed increase in weight during adult years and to the trend of increasing prevalence of obesity (215). Physical activity in NOWAC was reported to have increased on average; however, the standard deviation was large. The interpretability of change from one subjective physical activity reporting to the subsequent is challenging as we cannot assert with much confidence that the assessed change reflects true change (48). The dietary patterns of Norwegian women have likely evolved between 1996 to 2014, given changes in food availabilities and in tradition. There was, on average, no change in HLI diet score in NOWAC, although with a high standard deviation. This could have been attributed to simultaneous increases and

decreases in the reported intake across food groups with changes in food availability. However, such subgroup analysis was not undertaken to explore trends in individual food groups.

Lifestyle factor trends are likely connected to wider exposures, also fluid in time, that impact cancer risk. This may result in different patterns of behavioural change that are represented by the same numeric HLI score change according to different ages, periods, and cohorts. In our sensitivity analysis, we did not observe that the associations were different among women recruited early compared to late in the sampling period, which did not indicate period effects. However, the interaction of age, period, and cohort could have been investigated much deeper as it is variability of the baseline year together with the length of time between baseline and follow-up, and age that could modify the effect of lifestyle behaviours on cancer incidence. It would be necessary to investigate age-period-cohort effects in single lifestyle factors first, rather than in combined form, as the factors may have independent trends.

Variable time between baseline and follow-up

Paper II assessed lifestyle change by evaluating HLI scores at two timepoints that were on average 7 years apart, where the youngest women were age 40 at baseline and the eldest were 76 at follow-up. There was substantial variability in the distance between baseline and follow-up measurements, with a range of 2 to 11 years. It was thus important to consider the impact of the study design on the results. Firstly, we were unable to define when in the time window between baseline and follow-up the lifestyle change occurred. Changes occurring earlier in the time window could be expected to have a greater impact on cancer risk as the duration of the exposure to the new exposure level has time to take effect. For example, it takes approximately 5 to 9 years after smoking cessation for a detectable decrease in lung cancer risk (39). Given the prospective study design, we suspect that differences in the timing of lifestyle change in the time window of change occurred nondifferentially and likely attenuated our estimates.

We hypothesized that women with longer intervals between measurements would register as having changed more than women with shorter intervals since change could be expected to be a function of time. However, we did not observe that HLI score change was correlated with time between the two measurements. Furthermore, in background analysis not presented in Paper II, we observed that the associations between HLI score change rate and the outcomes

were similar to those modelled with absolute HLI score change. There are several possible explanations that may have acted concurrently: i) women generally engage in similar lifestyle behaviours in middle-late adulthood, ii) lifestyle modification occurs in a punctuated pattern rather than gradually as a function of time, iii) lifestyle behaviours are constantly shifting bidirectionally on the HLI. Unfortunately, none of these explanations can be explored with the available data. However, the discussion is relevant in the interpretation of the results and for planning future studies that can better approximate true exposure change with repeated measurements.

Unintentional change

Assessing HLI score change intended to measure overall lifestyle modifications preceding the latent period of cancer, not lifestyle modifications as a result of latent cancer. Ensuring that the exposure precedes the outcome is key criteria of causal inference (216) and is a temporal aspect that must be considered when building statistical models to estimate associations. Changes in physical activity, body weight, smoking habits, alcohol intake, and dietary habits can occur for a variety of reasons. Women may experience a decrease in the amount of physical activity due to increasing demands in other areas of life, reducing their time and capacity to engage in physical activity (217). Decreased physical activity levels can also occur as a result of reduced energy due to undiagnosed cancer or other morbidities. It is challenging to predict in which direction the estimates would be biased due to possible bidirectional change in several lifestyle factors in response to latent cancer as well as latent change being captured as improved or worsened depending on the lifestyle factor. Latent change in physical activity would typically manifest as worsened physical activity level, which would inflate estimates by strengthening the association between lower physical activity level and higher cancer risk. In contrast, latent weight change is often manifested as weight loss (218), which represents improved BMI with the HLI, thus attenuating estimates. In an attempt to eliminate the presence of lifestyle changes due to latent cancer in the sample, we excluded the first two years of follow-up and did not find any differences in estimates. However, this may not have been adequate given the long and undefined induction and latency periods of most adult cancers.

Probability distribution

It is plausible that some of the lifestyle change captured by HLI score change among the lowest and highest scorers was a reflection of the design of the HLI and regression to the mean. Regression to the mean is the phenomenon of the probability distribution, whereby natural variations in repeated data can be misleadingly interpreted as actual changes (219). In the context of HLI score change, extreme HLI scorers at baseline will be more likely to have less extreme HLI scores at follow-up. The effect of this tendency was heightened by the maximum and minimum constraints defined by the HLI. According to overall lifestyle as defined by the HLI, those with the lowest score were not permitted to worsen their lifestyle while those with the highest score were not permitted to improve their lifestyle. Outlining design constraints and regression to the mean, Figure 7 displays the distribution of HLI score change within baseline HLI groups. The random error that is responsible for fluctuations between repeated measurements and regression to the mean did not represent real change in the exposure and thus presented as noise in the models. The estimates may have experienced a degree of attenuation due to this phenomenon.

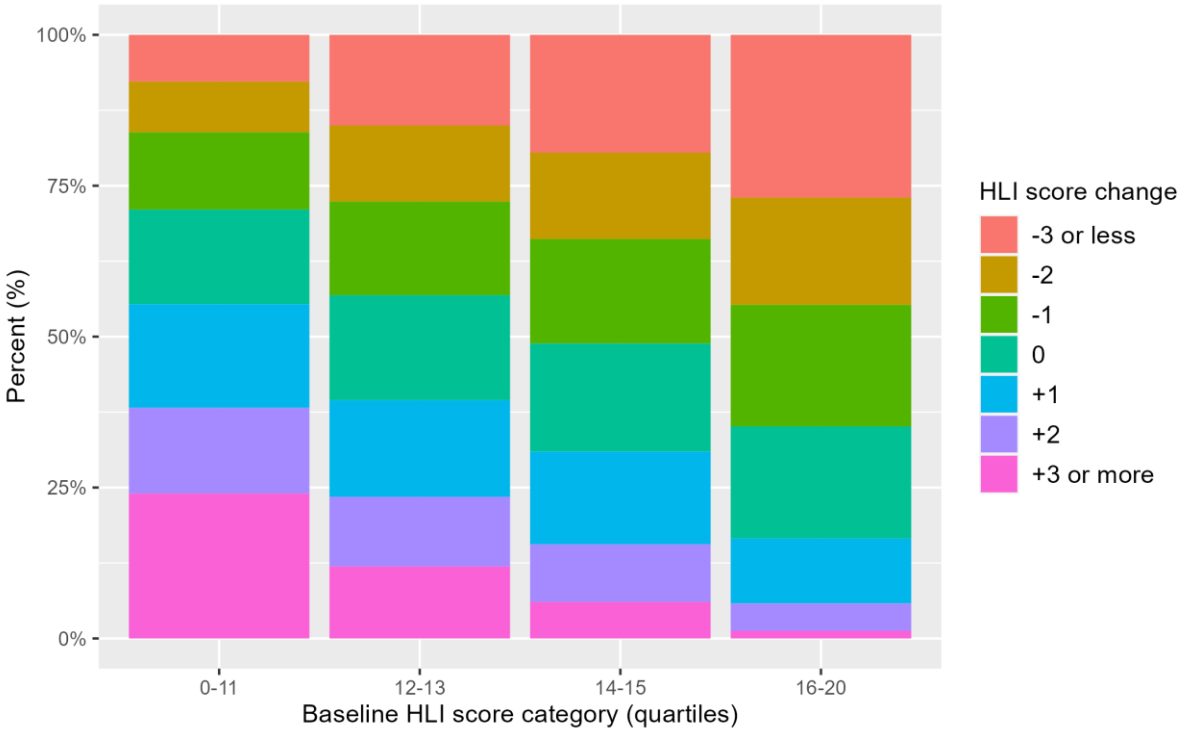


Figure 7. Distribution of HLI score change according to baseline HLI score categories.

5.2.3 Paper III – prediagnostic lifestyle and survival

The associations between prediagnostic HLI score and survival varied between breast, colorectal, and lung cancer survivors, in order of most protective to null. Among women diagnosed with breast and colorectal cancer, the associations were more protective in terms of all-cause mortality compared to breast and colorectal cancer mortality. We suspect that shared risk factors with prominent cancer comorbidities (220) that are also major causes of death in Norway (i.e. cardiovascular disease, chronic obstructive pulmonary disease, and, indirectly, type II diabetes mellitus) (11) were responsible for the stronger all-cause associations. It has been reported that the poorer the overall survival is for a cancer type, the smaller the effect of comorbidity on mortality (54). In accordance with this and in line with national figures (12), we observed that the five-year survival was highest for breast (90%), followed by colorectal (67%) and lung (29%) cancers.

There were indications that a higher prediagnostic HLI score was protective for breast and colorectal cancer mortality. This was also stronger among women diagnosed with breast compared to colorectal cancer. The cancer condition itself likely has higher relative explanatory importance for survival among lung compared to breast cancer survivors. It is thus reasonable that the impact of prediagnostic lifestyle factors would be most discernible in survival after breast cancer or where prognosis is already relatively good (221).

5.2.3.1 Breast cancer survival

The findings from Paper III indicated that a higher prediagnostic HLI score was associated with improved survival among women diagnosed with breast cancer. This association was especially clear for all-cause mortality, and present, although less strong for breast cancer mortality. To our knowledge, investigation of overall prediagnostic lifestyle and survival of breast cancer in a Norwegian cohort (103) is the only publication to examine this association. The findings from Paper III were consistent with this study's findings in terms of all-cause mortality. Unlike the findings from Paper III, they did not observe an association for breast cancer mortality. However, their sample size was small (n = 573) compared to ours (n = 5032), thus possibly being underpowered to detect an association.

We observed that the smoking component of the HLI was responsible for driving the protective association between the HLI score and survival. No protective association persisted for the remaining lifestyle factors combined. Our findings likely reflect the overwhelming

pool of evidence indicating the strong impact of smoking on higher mortality. As summarised in a meta-analysis, current compared to never smoking has been consistently associated with poorer survival among women diagnosed with breast cancer (62). We did not detect an effect from the BMI component combined with the remaining lifestyle factors despite consistent findings of poorer survival with higher prediagnostic body fatness (59). Findings for physical activity level (68, 75, 222, 223) and alcohol have been inconclusive (77, 79, 80, 81, 82, 83).

5.2.3.2 Colorectal cancer survival

There were indications that a higher prediagnostic HLI score was associated with improved survival among women diagnosed with colorectal cancer. Although, the evidence was weak. Our findings are consistent in terms of direction, but are weaker, compared to two studies employing the WCRF/AICR score in the EPIC cohort (71) and a US cohort (108); and one study employing an equally-weighted HLI based on the same five lifestyle factors as the NOWAC HLI in another US cohort (63). The WCRF/AICR score emphasizes diet and the stronger, protective associations observed in studies using the score may be due to prediagnostic diet being a stronger determinant of survival among colorectal cancer survivors. However, this does not explain the stronger association observed in the equally-weighted HLI score. Further, there is little consistent evidence for the protective effect of specific dietary factors or patterns in the literature (66, 84, 86, 87).

There were slight attenuations observed when smoking was excluded from the HLI, which is largely consistent with previous studies investigating single prediagnostic lifestyle factors and survival of colorectal cancer. Several studies have observed that smoking is associated with poorer prognosis for colorectal cancer survivors (63, 65, 74). However, results are inconsistent or weak for other single prediagnostic lifestyle factors. A systematic review reported that greater prediagnostic body fatness was only associated with poorer colorectal cancer prognosis among men, not women (60). According to single studies, the evidence for prediagnostic physical activity (63, 64, 66, 69, 70, 71, 72, 74, 76), alcohol (63, 66, 71, 74, 78), and diet (66, 84, 86, 87) is inconsistent.

5.2.3.3 Lung cancer survival

There was no association observed between prediagnostic HLI score and survival after lung cancer. Given the poor prognosis of lung cancer, we suspected that any influence from pre-diagnostic lifestyle was negligible and therefore, difficult to detect. However, there were

indications of a protective association for women diagnosed with early stage lung cancer only. This again supports the hypothesis that the relative explanatory power of prediagnostic lifestyle is only detectable when other negative prognostic factors are limited (221). To my knowledge, this was the first study to investigate this topic using a multifactor exposure. Studies on single prediagnostic lifestyle factors have pointed to body fatness (61), smoking (67), and some dietary factors (67) (224).

5.2.3.4 The importance of secondary prevention

The findings from Paper III highlight the importance of early detection of cancer as it suggests that the protective impact of prediagnostic healthy lifestyle for overall survival is greater when cancer is diagnosed at an earlier stage. Norway has had a mammography screening program in place since 1996 (225), which has contributed to the high rates of breast cancer survival. Of note, a colorectal cancer screening programme was recently initiated in 2022 (226). If it proves effective, there will be fewer with subclinical colorectal cancer diagnosed at a late stage.

In an ideal world, primary prevention would eliminate the occurrence of cancer. However, the dramatic modification of lifestyle behaviours and exposure to other environmental factors not only takes time, but the elimination of all carcinogenic factors is not feasible. Paper III has provided results suggesting that early detection, which is most often achieved through screening today, may be a facilitating factor that will allow the benefits of primary prevention to be realised in the event of cancer diagnosis.

5.2.4 Biological mechanisms

5.2.4.1 Cancer aetiology

Biological mechanisms that can explain the effect of lifestyle factors on cancer induction have been extensively investigated and several hypotheses have been shown to be plausible. They are especially clear for smoking, but are not completely understood for all of the lifestyle factors investigated and the across cancer sites and types. There is no single proposed biological mechanism that links physical activity, body fatness, smoking, alcohol intake, and dietary habits to cancer. Rather, each lifestyle factor may activate multiple plausible mechanisms that could have varying effects across different cancer sites, which are unified by in their presentation of the hallmarks of cancer (1).

Physical inactivity and greater adiposity have, independent of each other, been shown to increase the risk of metabolic dysfunction (i.e. hyperinsulinemia, insulin resistance, increased insulin-like growth factor I, increased fasting glucose) which are associated with tumour development in the breast and colorectum (22). Among postmenopausal women, it has been observed that hyperinsulinemia increases circulating oestrogen and testosterone (227). The evidence is weaker for premenopausal women. It is suspected that higher oestrogen levels promote tumour development in several oestrogen-related sites, such as the breast, ovary, and endometrium (227). Greater adiposity has also been shown to promote chronic inflammation from molecular and observational studies, which may cause tumour development (22). Clinical studies have shown that regular physical activity reduces pro-inflammatory biomarkers and observational studies have found that sedentary behaviour correlates with pro-inflammatory biomarkers (22). A common challenge with the aetiology of physical activity is separating the direct effect of physical activity on tumour development from the effect mediated through adiposity (10, 22). However, the connectedness of these two risk factors as behaviours and aetiology supports the use of a multifactor exposure measure.

The IARC Monographs programme has identified over 70 carcinogens in tobacco smoke that have demonstrated sufficient evidence for carcinogenicity in laboratory animals or humans (31). There are clear molecular links between smoking and DNA-adduct formation (chemicals bound to DNA that can cause a DNA mutation resulting in the upregulation of oncogenes), particularly in the lung, mouth, and bladder (228). The carcinogenic effect of alcohol intake is known to have a synergistic effect with smoking. It is hypothesised that ethanol can function as a solvent through which other carcinogens, such as those provided by smoking exposure, can penetrate the tissues of the upper aerodigestive tract (229). Independent of smoking, there is evidence that acetaldehyde – the primary metabolite of ethanol – can form DNA-adducts in human cells in vitro and in animal studies (229). Ethanol may also promote oxidative stress, which has been shown to damage DNA and is especially relevant for tumour development in the liver (229, 230). Alcohol intake has also been observed to increase chronic inflammation, which promotes tumour development in many sites (230). Heavy alcohol intake may also cause nutritional deficiencies by impairing nutrient absorption and altering metabolism (229).

The hypothesised mechanisms linking diet to cancer are too broad to explore in depth as the number of dietary exposures are numerous and correlated, as are the pathways to diverse cancer sites. Overall, dietary habits may contribute to tumour development through impacting the microbiota, which exerts effects on metabolism and immunity (10). The molecules and

metabolites of food processing have been linked to chronic inflammation (10, 231). In experimental studies, oxidative stress has been shown to be induced by high-glycaemic foods, which can activate inflammatory genes (231).

5.2.4.2 Survival after cancer diagnosis

Proposed biological mechanisms linking lifestyle factors to cancer-specific survival are scarce. While some publications suggest that risk factors for cancer can indirectly promote host death through the acceleration of tumour progression as an extension of tumour initiation (54, 60, 77, 223, 232), it is metastasis that is most commonly responsible for cancer death (233). Although related to tumour progression, metastasis is a distinct biological process, and its molecular mechanisms remain poorly understood (233). We found weak evidence for an association between overall prediagnostic lifestyle and cancer-specific survival. Our findings therefore do not provide robusticity for the hypothesis that unhealthy prediagnostic lifestyles trigger the biological mechanisms that accelerate metastasis and thus cause death.

5.2.5 Lifestyle factors in the aetiological exposure time window

The aetiological exposure time window refers to when in time the exposure was relevant for inducing the disease or event (14). As such, it is important that the timing of the measured exposure is situated in the aetiological exposure time window for a valid interpretation of results. The current thesis and most epidemiological studies rely on a "summary measure", which is a "simplifying assumption" that a single exposure measure measured at one timepoint is an indicator for risk (133) and thus covers the aetiological exposure time window (14). Papers I and III employ this study design and, like most observational studies, acknowledge that changes in exposure before or after measurement would have attenuated the estimates. In a study on physical activity level change in NOWAC, Borch et al. (68) observed that reduced physical activity level from pre- to postdiagnosis was associated with higher mortality among NOWAC women diagnosed with breast cancer. Such changes in lifestyle would have attenuated the association between overall prediagnostic lifestyle and survival of breast cancer in Paper III. Further, Paper II observed that lifestyle changes do occur among Norwegian women and that there is an association between lifestyle change and cancer risk. As such, the associations between baseline lifestyle and incidences of several cancers in Paper I would have been attenuated.

When associations are said to be attenuated in explanatory models such as these, it implies that there was underestimation of a causal association. This is true if the time of measurement captured some point in the aetiological time window. However, if the measurement was not in the aetiological exposure time window at all, the observed association would not be compatible with causation. Instead, the exposure would be a "mismeasured version of the true exposure" (133). For example, if higher body fatness during only childhood (unmeasured) was a true cause of cancer, and body fatness during adulthood (measured) was only correlated with childhood body fatness, results from a cohort study could infer that higher adult body fatness causes cancer. In the example, childhood body fatness would be an unmeasured confounder forcing a false causal association between adult body fatness and cancer risk. It is also reasonable to assume that past lifestyle behaviours are correlated with future lifestyle behaviours. Indeed, longitudinal analysis of repeated measurements must take correlations between measurements of individuals into account with covariance structures (234).

This being is a worst-case scenario, it is unlikely that most HLI score measurements were not situated somewhere in the aetiological exposure time window for cancer incidence given our knowledge on biological mechanisms. Although the HLI score(s) was measured during adulthood, the lifestyle factors that comprise the HLI are not sudden exposures. Lifestyle factors probably have long aetiological time windows for cancer where duration, in addition to intensity, of exposure plays a role such that a longitudinal and life course perspective would be relevant. Indeed, the assessment of lifestyle change in Paper II can be interpreted as a measure of duration of exposure according to intensity and was found to be associated with cancer risk. In a review of the role of physical activity in breast cancer aetiology, Friedenreich et al. (235) reported that lifetime physical activity was most strongly associated with breast cancer risk compared to physical activity at various timepoints in life. However, in terms of breast cancer survival, more recent physical activity levels before diagnosis could be a stronger predictor compared to physical activity levels earlier in adulthood (236). It is also accepted that longer smoking duration is associated with the higher risk of cancer in several sites independent of smoking intensity (31). Lifestyle factors extending into childhood may also be important. For example, childhood nutrition status is a risk factor for adult height and age at menarche, which are risk factors for cancer (23)).

In terms of cancer survival, our observation that the associations between the prediagnostic HLI score and mortality were markedly modified by the length of the prediagnostic interval highlighted the importance of investigating the timing of the exposure. Among breast cancer

survivors, we observed the strength of the protective associations increased with longer prediagnostic intervals, suggesting that lifestyle long before diagnosis has the greatest overlap with the true relevant exposure window (133). We tested if it was in fact age at questionnaire, which was correlated with prediagnostic interval, that modified the associations. However, models were not improved by an age at questionnaire interaction term, suggesting that it was the number of years before diagnosis, independent of age, that was important.

The importance of exposure accumulation of lifestyle factors over time and the changes in intensity of exposure are relevant for cancer risk and likely survival. This and the high potential for recall bias in retrospective study instruments should motivate the uptake of longitudinal study designs with frequent and regular measurements. In the context of cancer risk, and most likely other chronic diseases and mortality beyond the scope of this thesis, this should ideally take place throughout the life course.

5.2.6 What can we know from the HLI?

In explanatory epidemiological studies, such as Papers I-III, providing a sound explanation for the results is dependent on many considerations. Perhaps the most basic and essential is considering the degree to which the explanatory variable reflects the intended study exposure. Given the number of studies using the HLI as an exposure measure in explanatory models for cancer-related outcomes, it is important to reflect upon the HLI's surrogate relationship to its several component exposures and the true exposure it aims to measure in light of the results.

Every exposure used in epidemiology is a proxy measure for a conceptual "true exposure" that is itself not directly measurable (14). The researcher must therefore operationalise the exposure, ideally in the closest way possible, while balancing feasibility. For example, BMI measures kg/m^2 to operationalise body fatness. It is an imperfect measure as it does not account for the proportion and distribution of lean compared to fat mass, among other factors (237). However, BMI is simple to obtain and relevant for many health outcomes on a population level (238).

What is the true exposure of the HLI when it is comprised of several lifestyle factors that are diverse and dynamic in their carcinogenic processes and impacts? This is challenging to define as there is no single or set of proposed biological mechanisms that unify the five lifestyle factors with respect to cancer risk or survival. If we revisit the purpose of constructing the HLI in this thesis, it was to provide a simple measure that represents a

gradient of overall healthy lifestyle as defined by the inclusion of selected evidence-based risk factors. Throughout Papers I-III, we use the term "combined" lifestyle factors; as do others (183, 184). Several studies that use an HLI score to explain cancer risk maintain that the HLI is a "joint" measure of overall lifestyle (90, 184, 239, 240). Both terms denote an appreciation of potential interactive effects between lifestyle factors. In this sense, the HLI can be regarded as a generalised measure of simultaneous adherence to healthy lifestyle behaviours.

Would it be valid to infer that a high HLI score is the same as overall healthy lifestyle? As we have defined the HLI in the context of cancer prevention, this is generally true as it would take at least partial adherence to lifestyle recommendations for cancer prevention in order to score highly. Papers I-III observed associations between the HLI score, or HLI score change, and cancer-related outcomes. From a statistical inference perspective, the HLI score explained the variation in many outcomes better than chance. If the HLI score is a measure of overall healthy lifestyle, does this also mean that overall healthy lifestyle explained the variation in outcomes? As I will discuss below, the apparent simplicity of the HLI score in its construction has limitations due to lost information, the dynamic interplay and independent effects of the HLI's component exposures, as well as the relative importance of single components for different cancers in terms of incidence and survival. However, this trait that limits the interpretation of results may also be its main strength in explanatory models.

5.2.6.1 Interpretation of an increment

As Figure 3 in 3.3.1 describes, a 1-point increment on the HLI can represent many exposure differences. For example, BMI of 32 kg/m^2 compared to 28 kg/m^2 or physical activity level in the first compared to second quintile are each represented by 1-point. There are two reasons for why the 1-point increment loses its meaning as it moves from being a descriptive feature of an individual to an explanatory increment in the sample population.

1)

Unintentional weighting of the lifestyle factors would have occurred due to the unequal risk associated with each contributing lifestyle factor. This pattern of unintentional weighting would be unique to the outcome. For example, in Paper I, a 1-point HLI increment was associated with 14% lower lung cancer incidence. Clearly, inferring that women with BMI between 27.0 kg/m^2 and 29.9 kg/m^2 have 14% lower incidence of lung cancer compared to women with BMI $\geq$ 30.0 kg/m^2 would be absurd given our a priori knowledge on the minimal

effect BMI has on lung cancer risk and the null association observed after smoking was removed from the HLI. Correspondingly, 14% lower incidence among current smokers of < 15 cigarettes per day compared to current smokers of ≥ 15 cigarettes per day is probably an underestimated strength of association due to the attenuating effect of the other lifestyle factors. As such, if attempting to interpret the HLI score risk differences on a per-component basis, the impact of weakly associated factors would become overestimated while the impact of strongly associated factors would be underestimated.

Although lung cancer is an extreme example due to the dominance of smoking, the interpretations of a 1-point increment with respect to the risk of other outcomes are not immune to the ambiguity that comes with unintentional weighting. Even in a hypothetical setting where all five lifestyle factors were equally associated with the outcome, I would be sceptical of attributing the risk difference of a 1-point HLI increment to any of the specific components. Firstly, the interpretation would ignore possible synergy that the HLI captures as a multifactor exposure. Secondly, in the case of overall lifestyle change, it is impossible for participants to improve their smoking component score from less than 4 (former or current smoker) to 4 (never smoker). This type of restriction complicates the flexible interpretation of the HLI increment for risk.

2)

The HLI score can represent a wide variation of lifestyle patterns, which as paragraph 1) eluded to, would be differentially associated with the outcome. For a HLI score of 20, the highest scores were achieved in all five lifestyle components and the lifestyle pattern here is relatively well-defined, although may be different across study samples and populations if there is large variation in these “high achievers”. If we consider a HLI score of 19, the highest scores were achieved in four of five lifestyle components; and only three of four points were achieved for the remaining lifestyle component. Problematically, the lifestyle component to which this missing point belongs becomes lost in the score, the distribution of the 3-point factor may not be equal across all five factors in the sample, and the distribution of the 3-point factor may differ between populations. If we consider a score of 10, the combinations of points attributable to lifestyle components become overwhelming to consider. As such, risk differences associated with a 1-point HLI score increment would be fixed to the lifestyle patterns of the sample population which prevents the ability to relate risk estimates back to lifestyle factors and to risk estimates observed in other studies. This also limits the

generalisability of the results. The study design in Paper II highlights this dilemma as a HLI score change equal to zero is defined as no lifestyle change when the underlying lifestyle patterns at baseline and follow-up could be theoretically be very distinct and represent much lifestyle change.

Considering paragraphs 1) and 2) together, the HR associated with a 1-point increment on the HLI represents the risk difference of a healthier compared to a less healthy lifestyle defined by patterns of physical activity level, BMI, smoking habits, alcohol intake, and dietary habits that are based on the variation in the sample where there is unintentional weighting across the lifestyle factors depending on the outcome. While a weighted HLI would remedy some of the issues with unintentional weighting, it would not remedy the interpretability of an increment. The ambiguity of the 1-point increment in relation to the true exposures of its components renders any other increment, such as per 1-SD or quartiles just as challenging to interpret. Maintaining a stable lifestyle, or a HLI score change of zero, is analogous to the variation of lifestyle patterns that could comprise the same numeric score – also ambiguous. Importantly, this ambiguity impedes the end goal of explanatory models, which is causal inference. As such, interpretations of increments associated with risk differences should be cautious as they cannot provide much more insight than risk associated with average overall healthier lifestyle or average greater adherence to lifestyle recommendations for cancer prevention.

5.2.6.2 Addressing synergy and clustering

The use of HLIs in explanatory models is commonly motivated by a brief mention of its suggested superior ability to appreciate synergy (206, 239, 241) and clustering (Paper I) (184, 193, 206) of lifestyle factors compared to the modelling of single lifestyle factors. It seems the brevity of scientific articles have not allowed the space for a discussion of how synergy and clustering is addressed or understood when employing a multifactor exposure like the HLI. I will offer some reflections in the following sections.

Approach to clustering

Clustering is a characteristic of the exposure and describes the concurrent exposure to several factors. Given the recommendations targeting multiple lifestyle factors for cancer prevention, it would be of interest to compare the risk associated with meeting all recommendations to meeting fewer or none. However, it would be important that the lifestyle exposures composing each cluster are known. The HLIs ability to assess the risk associated with the

clustering of lifestyle factors has been presented as an advantage over the assessment of single factors. There is evidence that healthy lifestyle behaviours often cluster in Western populations (92, 93, 94). Although in NOWAC, clustering is probably less dichotomous (242). In the context of the HLI used in this thesis, clustering could only be captured by individuals who have extreme HLI scores. The pattern of lifestyle factor-clustering for those scoring centrally on the HLI is entirely unknown and is where the majority of NOWAC women score. As such, comparing the risk associated with the women that score the highest to lowest, as done in Papers I-III with the HLI score modelled in categories, may be the only approach to understanding the effect of clustering and where synergy may be suggested.

A cleaner approach to addressing clustering would be to construct HLIs with binary-scored components as used in several studies where the score corresponds to the number of lifestyle factor recommendations met (100, 183, 185, 198, 239). These explanatory studies compare cancer risk associated with belonging to more favourable lifestyle factors to fewer. The binary HLIs face similar challenges of identifying the exposure level among the central scorers, although to a lesser degree than more graded HLIs. Latent class analysis would be a superior approach in the pursuit to identify causal associations between lifestyle patterns and cancer outcomes.

Approach to synergy

Synergy can occur in some cases of clustering, but not necessarily. Synergy, synonymous with *interaction*, refers to the concept that the effect of an exposure on an outcome depends on the presence of a second exposure. For example, smoking and alcohol intake interact to increase the risk of upper aerodigestive cancers (31); and high body fatness and alcohol intake has been reported to increase the risk of liver cancer (243). It is reasonable to assume, given our current understanding of biological mechanisms and studies estimating the statistical interaction between two lifestyle factors, that there is a biological synergistic effect between combinations and subsets of physical activity, body fatness, smoking, alcohol intake, and diet on cancer risk and other health outcomes.

Investigating biological interaction typically involves the assessment of statistical interaction by the inclusion of an interaction term in a multivariable model and/or assessing the association of interest in subgroups across the hypothesized interacting variable. The approach of the HLI represents none of these. Rather it takes on an approach resembling that

of dietary quality scores, whereby components of the multifactor exposure are summed. As far as I understand, there is no method to test for interaction between components using an additive score and their purpose in dietary epidemiology is not to assess interaction (244). Further, due to the limitations described in Section 5.2.6.1, it would not be possible to identify the source of the interaction (i.e. which two or more lifestyle factors are involved). Indeed, HLI scores can represent diverse lifestyle patterns that interact and are related to risk uniquely. To glean specific results on interaction, use of other methods, such as latent class analysis would be preferred over the HLI. With latent class analysis, specific lifestyle patterns that are present in the sample population can be compared, rather than an average increase or decrease in all components. Although, if the point is that the HLI can simply appreciate synergy without testing for its presence, this advantage would probably be overshadowed by the inability to make meaningful interpretations of the risk differences.

Despite these limitations, there are cases where meaningful interpretation of the cause of risk differences is less pertinent, and the HLI could provide a crude indication of synergy and generate hypotheses for more in-depth research.

Colorectal cancer incidence

As discussed in Section 5.2.1.2, Paper I observed that the HLI score was associated with lower colorectal cancer incidence when, previously in NOWAC associations for single lifestyle factors were not detected. As such, it is possible that the HLI captured a synergistic effect between the component lifestyle factors. This could be worth exploring further given the high incidence of colorectal cancer in Norway.

Lung cancer incidence

Multifactor lifestyle exposures and impact of lung cancer have seldom been explored. I suspect this to be because few see the benefit of investigating risk for a cancer site with a known dominant risk factor and have prioritised the use of multifactor lifestyle exposures for multifactorial cancers. Indeed, it has been estimated that smoking was responsible for 80% of lung cancer cases among women (23), which would seem to leave little room for the impact of other lifestyle factors after also considering the strong role of occupational and environmental exposures (31). However, examination of the HLI score – after removing smoking – to explain lung cancer risk is possibly a more informative setting since it could allow for associations to be detected on the basis of synergy that would not otherwise be

detectable with the isolation of single lifestyle factors. In the case of NOWAC, there was no association between the HLI score and lung cancer incidence when smoking was excluded from the HLI in Paper I. This may not be the case in other populations and samples where internal validity may be higher and lifestyle patterning is different.

5.2.6.3 Public health implications

Use of the HLI has been encouraged in the literature for its potential public health implications as it can relate several lifestyle behaviours to public health goals (90, 93, 94, 245) – i.e. reducing the incidence of cancer (10). As discussed, the HLI score is likely a poor exposure measure for explanatory models and unimpressively examines the clustering and synergy of lifestyle factors. The central problem of relating risk differences back to the HLI component exposures highlights the inability of HLI models to implicate specific lifestyle modifications. Public health recommendations require findings that point to the health impact of actionable modification, which the HLI does not provide. Proportional attributable fractions, which estimate the number of events avoided given the strength of effect and frequency of exposure (246), have been estimated in some HLI studies (183, 196, 247). However, this does not remedy the problem of exposure interpretability and actionability.

Despite these issues, HLIs can be useful in public health for some explanatory contexts. For example, they can be used to test if adherence to new recommendations is differently associated with outcomes compared to old recommendations when there is a change in official recommendations. The HLI can also be used for surveillance as descriptive epidemiology. For example, it can provide a crude measure of overall lifestyle before and after implementation of a public health intervention as an indicator of progress. More generally, the HLI could be used to survey population exposure to lifestyle factors over time. This can help identify subgroups of the population that most require attention (93).

In terms of clinical implications, predictive contexts for the HLI may also be useful. The HLI's predictive capacity can be evaluated to create a basic risk assessment tool for cancer and other health outcomes (248). This could be helpful in dialogue with a family physician and for identifying at-risk patients.

5.2.6.4 Added benefit of the HLI

Despite the possible synergy that was captured by the HLI score in explaining the variation of colorectal cancer incidence in Paper I, several reasons have been outlined above for why the HLI may not be an ideal approach to studying overall lifestyle from a scientific and public health perspective. Further, the benefits of its approach to synergy and clustering of lifestyle factors likely did not outweigh it limitations. This was tested empirically, as described in Section 3.6.3, finding that the modelling of single lifestyle factors was superior to the HLI score in explaining the variation in cancer incidence for all cancer sites investigated in Paper I (results presented in Section 4.4). It is almost certain that the modelling of HLIs that were weighted to the specific outcomes investigated would have performed better than unweighted HLIs. However, as discussed earlier, this would come with its own challenges of interpretability and could not be considered a simple multifactor measure. Use of HLIs as exposure measures in future studies should be considered carefully, taking into account the aim of the research question and limitations that may be encountered when interpreting results.

5.2.7 Healthier living in the population

Despite the limitations of the HLI score as a multifactor exposure measure, the results generally support lifestyle recommendations for cancer prevention for Norwegian women listed by WCFR/AIRC (249), WHO (250), ACS (89), and the Norwegian Cancer Society (182). This is not a novel finding and should not be expected to be novel given the volume of study findings, meta-analyses, and panel judgements that form the foundation for these recommendations. While there is much that is yet to be known about the true association between lifestyle behaviours and cancer risk and survival, particularly for the impact of dietary factors, there is much we do know that has yet to be implemented (251, 252).

How do we lead healthier lives? The public health bodies mentioned have delivered strong public health messaging. Yet, we continue to see low levels of physical activity (253, 254), trends of increasing obesity (252, 255), and high alcohol intake (214, 256) globally and in Norway. There are clearly barriers to populations adopting healthier lifestyles. There is heavy emphasis from public health bodies on individual responsibility for adhering to a healthy lifestyle as behaviours are regularly framed as choices made by individuals. However, healthy living among individuals, who form the population, is inseparable from culture and the socio-political systems that shape it (252). It is therefore logical that obstacles to healthy lifestyles

are a function of the wider systems. For example, despite our knowledge about the negative health effects of alcohol intake, there is an active blindness towards its health risks outside of the scientific sphere and there lacks impetus – cultural or political – to reduce its prevalence (257). Pressure is also being placed on economic policy from public health bodies, as exemplified by smoking bans and taxes on cigarettes, alcohol, and sugar in Norway. Indeed, the prevalence of smoking has declined since 1998 (258).

Yet, the disparities in healthy living across the SEP spectrum suggest that there is an unequal distribution of access to healthy behaviours (254). In Norway, women with the lowest education levels disproportionately bare low physical activity levels (259), excessive alcohol intake habits (214), tobacco smoking (205), and obesity (260). In a qualitative study, a sample of young Norwegian women identified high costs of healthy foods and sports participation, lack of time, as well as clear education on nutrition and preparation of healthy meals as barriers for weight management (217). These psychological, time, and financial-related obstacles appear to be somewhat universal in industrialised societies (261).

This leads me to a contention with a specific lifestyle factor recommendation – maintaining a healthy weight. It is firstly arguable whether body weight is a lifestyle factor as it is not a behaviour and not actionable in the sense that exercising more, smoking less, drinking less, and eating healthier are. Secondly, obesity represents a disease status itself and can be the sequelae of other underlying disease (209). Further, it is not only a product of the energy balance as determined by the mentioned actionable factors, but is also result of the complex interplay of factors, many of which are not modifiable by the individual (262). Statistics Norway reports that those with lower education level tend to have more obesity compared to people with higher education (259). There is substantial social shame in Western society associated with overweight and obesity, as the cultures, which are influenced by public health messaging, place responsibility on the individual for poor lifestyle choices. I fear that the focus on body weight, and its inclusion in HLIs, overshadows more sinister problems at the root of obesity that should be at the frontier of public health, medicine, and policy.

It would seem that for adherence to healthy lifestyle behaviours to be sustainable by individuals, they must be compatible with the wider societal framework. Exploring this further represents an entire field of study on its own. However, knowledge from other fields are necessary to translate epidemiological findings into sustainable and ethical change; and to

address the barriers that more disadvantaged segments of the population disproportionally face.

6 Conclusion

This thesis, as a collection of three scientific research papers, aimed to investigate a HLI, representing a simple multifactor exposure measure for a gradient of healthy lifestyle behaviours, and its association to cancer-related health outcomes among women in Norway.

There were several conclusions:

1) Overall healthier lifestyle was associated with the lower risk of cancer in several frequent sites among Norwegian women. Our findings are generally consistent with the current evidence on risk factors for cancer and support recommendations for cancer prevention.
    a. Higher HLI score was associated with a lower incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, pancreatic, and kidney cancer; but was not associated with postmenopausal ovarian cancer risk.
    b. Smoking was responsible for the association between HLI score and lung cancer risk, and HLI score and pancreatic cancer risk.
2) Overall lifestyle changes in adulthood were associated with the risk of cancer in several sites combined. Our findings support adherence to a healthy lifestyle, avoidance of lifestyle worsening, and lifestyle improvement.
    a. Greater positive magnitude of HLI score change was associated with a lower incidence of lifestyle-related cancers.
    b. The strength of association between lifestyle worsening, as defined by negative HLI score change, compared to stable lifestyle and the incidence of lifestyle-related cancer may be stronger than that for lifestyle improvement, as defined by positive HLI score change.
3) Overall prediagnostic lifestyle was not strongly associated with survival among women diagnosed with breast, colorectal, or lung cancer. More research on the prognostic effect of lifestyle factors on cancer survival is required.
    a. Prediagnostic HLI score was associated with lower all-cause and breast cancer mortality among women diagnosed with breast cancer; however smoking was responsible for the associations.
    b. Prediagnostic HLI score may be associated with lower all-cause and colorectal cancer mortality among women diagnosed with colorectal cancer. Smoking may have been responsible for the associations.

c. Prediagnostic HLI score was not associated with lower all-cause and lung cancer mortality among women diagnosed with lung cancer

4) Due to the limitations of the HLI score as an exposure measure of overall lifestyle and its dependence on the distribution of factors in the source sample, it may not be reasonable to extrapolate these results to other populations of women.
5) Generalised multifactor lifestyle exposures are not particularly useful in explanatory models as they cannot provide knowledge on the strength of interaction between potentially aetiologically dependent factors. Further, their resulting measures of risk cannot implicate actionable exposures that are required by public health to implement interventions.

7 Future perspectives

The undertaking of this PhD project has offered some insights on promising directions and lessons learned in the field of lifestyle risk factors and cancer outcomes. Some suggestions have already been presented in earlier sections of this thesis and are summarised here. Most of these future perspectives can also apply to chronic, lifestyle-related disease outcomes other than cancer.

1. The use of multifactor exposure measures should be carefully considered with respect to the research question before use in explanatory models.
2. A healthy lifestyle index can:
    a. Provide a measure of overall healthy lifestyle for population surveillance and for measuring the effectiveness of lifestyle interventions
    b. Provide a convenient summary measure and be used as a potential risk assessment tool. However, the HLI's predictive ability will need to be tested in predictive models.
3. The clustering of lifestyle factors is relevant for the risk of health outcomes and should be investigated using exposure measures with a high degree of interpretability, such as latent class analysis. Real lifestyle patterns that exist in the population should be considered across different birth cohorts to appreciate age-period-cohort fluidity in behaviours.
4. To better address causal associations in observational studies, obtaining longitudinal data with frequent and regular repeated measurements should a priority when designing cohort studies. The knowledge potential in accessing a greater number of repeated measurements would improve the internal validity of studies assessing most slow-building exposures on health outcomes. Measurements throughout the life course are important as relevant time windows of exposure and latent periods can be extensive. The leveraging of technology should be embraced to enable less costly, but frequent data collection.
5. Dietary scores are only as powerful as their raw data. Reducing measurement error in data collection of diet is critical for better understanding causal associations between dietary exposures and health outcomes as well as for providing adequate adjustment as a confounder in models with other explanatory variables of interest. Technology has potential to assist in this area as well.

6. There is a need for studies that are designed to assess pre-defined exposures, particularly with respect to exposure change (i.e. alcohol cessation).
7. It is a public health, societal, and political obligation to provide equitable access to healthy lifestyle behaviours. Collaboration between academia, policy, and economy will be necessary to effectively move populations towards healthier living.

Works cited

1. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discovery. 2022;12(1):31-46.
2. Miller ME. Cancer. New York: Momentum Press; 2018.
3. Gospodarowicz M, O'Sullivan B. Prognostic factors in cancer. Semin Surg Oncol. 2003;21(1):13-8.
4. Adami HO, Nyrén O. Enigmas, priorities and opportunities in cancer epidemiology. European journal of epidemiology. 2016;31(12):1161-71.
5. Kattner P, Strobel H, Khoshnevis N, Grunert M, Bartholomae S, Pruss M, et al. Compare and contrast: pediatric cancer versus adult malignancies. Cancer and Metastasis Reviews. 2019;38(4):673-82.
6. Kocarnik JM, Compton K, Dean FE, Fu W, Gaw BL, Harvey JD, et al. Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life Years for 29 Cancer Groups From 2010 to 2019: A Systematic Analysis for the Global Burden of Disease Study 2019. JAMA Oncol. 2022;8(3):420-44.
7. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. 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.
8. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.
9. Global Burden of Disease Cancer Collaboration. Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life Years for 29 Cancer Groups From 2010 to 2019: A Systematic Analysis for the Global Burden of Disease Study 2019. JAMA Oncology. 2022;8(3):420-44.
10. International Agency for Research on Cancer (IARC). World Cancer Report: cancer research for cancer prevention. Lyon: International Agency for Research on Cancer (IARC); 2020.
11. Norwegian Cause of Death Register. Figures from the Norwegian Cause of Death Register 2021. 2022 [Available from: https://www.fhi.no/hn/helseregistre-og-registre/dodsarsaksregisteret/tall-fra-dodsarsaksregisteret-2021/]
12. Cancer Registry of Norway. Cancer in Norway 2021 - Cancer incidence, mortality, survival and prevalence in Norway. Oslo: Cancer Registry of Norway; 2022.
13. Statistics Norway. 07459: Population, by sex and one-year age groups (M) 1986 - 2023. 2023 [Available from: https://www.ssb.no/en/statbank/table/07459/]
14. White E, Armstrong BK, Saracci R. Principles of Exposure Measurement in Epidemiology: Collecting, evaluating and improving measures of disease risk factors: OUP Oxford; 2008.
15. Rothman KJ. Causes. Am J Epidemiol. 1976;104(6):587-92.
16. Anand P, Kunnumakkara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, et al. Cancer is a preventable disease that requires major lifestyle changes. Pharmaceutical research. 2008;25(9):2097-116.
17. Brown KF, Rumgay H, Dunlop C, Ryan M, Quartly F, Cox A, et al. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. Br J Cancer. 2018;118(8):1130-41.
18. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. Eur J Cancer. 2001;37 Suppl 8:S4-66.
19. Danaei G, Vander Hoorn S, Lopez AD, Murray CJ, Ezzati M. Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. Lancet (London, England). 2005;366(9499):1784-93.
20. Parkin DM, Boyd L, Walker LC. 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. Br J Cancer. 2011;105 Suppl 2(Suppl 2):S77-81.
21. Brennan P, Davey-Smith G. Identifying Novel Causes of Cancers to Enhance Cancer Prevention: New Strategies Are Needed. JNCI: Journal of the National Cancer Institute. 2021:djab204.

22. Friedenreich CM, Ryder-Burbidge C, McNeil J. Physical activity, obesity and sedentary behavior in cancer etiology: epidemiologic evidence and biologic mechanisms. Molecular Oncology. 2021;15(3):790-800.
23. World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR). Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. 2018.
24. McTiernan A, Friedenreich CM, Katzmarzyk PT, Powell KE, Macko R, Buchner D, et al. Physical Activity in Cancer Prevention and Survival: A Systematic Review. Med Sci Sports Exerc. 2019;51(6):1252-61.
25. World Health Organization (WHO). Fact sheet: obesity and overweight.
26. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness and Cancer — Viewpoint of the IARC Working Group. New England Journal of Medicine. 2016;375(8):794-8.
27. da Silva M, Weiderpass E, Licaj I, Lissner L, Rylander C. Excess body weight, weight gain and obesity-related cancer risk in women in Norway: the Norwegian Women and Cancer study. Br J Cancer. 2018;119(5):646-56.
28. Parra-Soto S, Cowley ES, Rezende LFM, Ferreccio C, Mathers JC, Pell JP, et al. Associations of six adiposity-related markers with incidence and mortality from 24 cancers—findings from the UK Biobank prospective cohort study. BMC medicine. 2021;19(1):7.
29. Recalde M, Davila-Batista V, Díaz Y, Leitzmann M, Romieu I, Freisling H, et al. Body mass index and waist circumference in relation to the risk of 26 types of cancer: a prospective cohort study of 3.5 million adults in Spain. BMC medicine. 2021;19(1):10.
30. National Center for Chronic Disease P, Health Promotion Office on S, Health. Reports of the Surgeon General. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2014.
31. International Agency for Research on Cancer (IARC). IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Lyon: IARC; 2012.
32. International Agency for Research on Cancer (IARC). List of classifications by cancer sites with sufficient or limited evidence in humans, IARC Monographs Volumes 1–132a. World Health Organization (WHO); 2022.
33. International Agency for Research on Cancer (IARC). Human Cancer: Known Causes and Prevention by Organ Site. 2022 [updated 01/09/2021. Available from: https://monographs.iarc.who.int/human_cancer_known_causes_and_prevention_organ_site/]
34. Macacu A, Autier P, Boniol M, Boyle P. Active and passive smoking and risk of breast cancer: a meta-analysis. Breast Cancer Res Treat. 2015;154(2):213-24.
35. World Cancer Research Fund International / American Institute for Cancer Research (WCRF/AICR). Wholegrains, vegetables and fruit and the risk of cancer. 2018.
36. Rumgay H, Shield K, Charvat H, Ferrari P, Sornpaisarn B, Obot I, et al. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. Lancet Oncol. 2021;22(8):1071-80.
37. Schwingshackl L, Schwedhelm C, Galbete C, Hoffmann G. Adherence to Mediterranean Diet and Risk of Cancer: An Updated Systematic Review and Meta-Analysis. Nutrients. 2017;9(10):1063.
38. Schwingshackl L, Bogensberger B, Hoffmann G. Diet Quality as Assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and Health Outcomes: An Updated Systematic Review and Meta-Analysis of Cohort Studies. Journal of the Academy of Nutrition and Dietetics. 2017;118(1):74-100.e11.
39. International Agency for Research on Cancer (IARC). IARC Handbooks of Cancer Prevention - Tobacco Control: Reversal of Risk after Quitting Smoking. Lyon2007.
40. Bosetti C, Gallus S, Peto R, Negri E, Talamini R, Tavani A, et al. Tobacco smoking, smoking cessation, and cumulative risk of upper aerodigestive tract cancers. Am J Epidemiol. 2008;167(4):468-73.
41. Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. BMJ. 2000;321(7257):323.
42. Jha P, Peto R. Global Health: Global Effects of Smoking, of Quitting, and of Taxing Tobacco. New England Journal of Medicine. 2014;370(1):60-8.

43. Keum N, Greenwood DC, Lee DH, Kim R, Aune D, Ju W, et al. Adult weight gain and adiposity-related cancers: a dose-response meta-analysis of prospective observational studies. Journal of the National Cancer Institute. 2015;107(2).
44. Radimer KL, Ballard-Barbash R, Miller JS, Fay MP, Schatzkin A, Troiano R, et al. Weight change and the risk of late-onset breast cancer in the original Framingham cohort. Nutrition and cancer. 2004;49(1):7-13.
45. Teras LR, Goodman M, Patel AV, Diver WR, Flanders WD, Feigelson HS. Weight loss and postmenopausal breast cancer in a prospective cohort of overweight and obese US women. Cancer causes & control : CCC. 2011;22(4):573-9.
46. Emaus MJ, van Gils CH, Bakker MF, Bisschop CN, Monninkhof EM, Bueno-de-Mesquita HB, et al. Weight change in middle adulthood and breast cancer risk in the EPIC-PANACEA study. International journal of cancer. 2014;135(12):2887-99.
47. Robsahm TE, Heir T, Sandvik L, Prestgaard E, Tretli S, Erikssen JE, et al. Changes in midlife fitness, body mass index, and smoking influence cancer incidence and mortality: A prospective cohort study in men. Cancer medicine. 2019.
48. Oyeyemi SO, Braaten T, Licaj I, Lund E, Benjaminsen Borch K. Physical activity patterns and the risk of colorectal cancer in the Norwegian Women and Cancer study: a population-based prospective study. BMC cancer. 2018;18(1):1216.
49. Dam MK, Hvidtfeldt UA, Tjønneland A, Overvad K, Grønbæk M, Tolstrup JS. Five year change in alcohol intake and risk of breast cancer and coronary heart disease among postmenopausal women: prospective cohort study. BMJ. 2016;353:i2314.
50. Ahmad Kiadaliri A, Jarl J, Gavriilidis G, Gerdtham UG. Alcohol drinking cessation and the risk of laryngeal and pharyngeal cancers: a systematic review and meta-analysis. PLoS One. 2013;8(3):e58158.
51. Marron M, Boffetta P, Zhang ZF, Zaridze D, Wünsch-Filho V, Winn DM, et al. Cessation of alcohol drinking, tobacco smoking and the reversal of head and neck cancer risk. Int J Epidemiol. 2010;39(1):182-96.
52. Anderson KN, Schwab RB, Martinez ME. Reproductive risk factors and breast cancer subtypes: a review of the literature. Breast Cancer Research and Treatment. 2014;144(1):1-10.
53. Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971-2011: a population-based study. Lancet (London, England). 2015;385(9974):1206-18.
54. Søgaard M, Thomsen RW, Bossen KS, Sørensen HT, Nørgaard M. The impact of comorbidity on cancer survival: a review. Clinical epidemiology. 2013;5(Suppl 1):3-29.
55. International Agency for Research on Cancer (IARC). Social Inequalities and Cancer. Lyon: IARC; 1997.
56. Braaten T, Weiderpass E, Lund E. Socioeconomic differences in cancer survival: the Norwegian Women and Cancer Study. BMC Public Health. 2009;9:178.
57. van Zutphen M, Winkels RM, van Duijnhoven FJ, van Harten-Gerritsen SA, Kok DE, van Duijvendijk P, et al. An increase in physical activity after colorectal cancer surgery is associated with improved recovery of physical functioning: a prospective cohort study. BMC cancer. 2017;17(1):74.
58. American Cancer Society (editor). The American Cancer Society's principles of oncology : prevention to survivorship. Society AC, editor. Hoboken, New Jersey ; Chichester, West Sussex, England :: John Wiley & Sons, Inc; 2018.
59. Chan DSM, Vieira AR, Aune D, Bandera EV, Greenwood DC, McTiernan A, et al. Body mass index and survival in women with breast cancer-systematic literature review and meta-analysis of 82 follow-up studies. Ann Oncol. 2014;25(10):1901-14.
60. Parkin E, O'Reilly DA, Sherlock DJ, Manoharan P, Renehan AG. Excess adiposity and survival in patients with colorectal cancer: a systematic review. Obesity Reviews. 2014;15(5):434-51.
61. Gupta A, Majumder K, Arora N, Mayo HG, Singh PP, Beg MS, et al. Premorbid body mass index and mortality in patients with lung cancer: A systematic review and meta-analysis. Lung Cancer. 2016;102:49-59.
62. Berube S, Lemieux J, Moore L, Maunsell E, Brisson J. Smoking at time of diagnosis and breast cancer-specific survival: new findings and systematic review with meta-analysis. Breast Cancer Research. 2014;16(2).

63. Pelser C, Arem H, Pfeiffer RM, Elena JW, Alfano CM, Hollenbeck AR, et al. Prediagnostic lifestyle factors and survival after colon and rectal cancer diagnosis in the National Institutes of Health (NIH)-AARP Diet and Health Study. Cancer. 2014;120(10):1540-7.
64. Boyle T, Fritschi L, Platell C, Heyworth J. Lifestyle factors associated with survival after colorectal cancer diagnosis. Br J Cancer. 2013;109(3):814-22.
65. Zhu Y, Yang SR, Wang PP, Savas S, Wish T, Zhao J, et al. Influence of pre-diagnostic cigarette smoking on colorectal cancer survival: overall and by tumour molecular phenotype. Br J Cancer. 2014;110(5):1359-66.
66. Oyeyemi SO, Braaten T, Skeie G, Borch KB. Competing mortality risks analysis of prediagnostic lifestyle and dietary factors in colorectal cancer survival: the Norwegian Women and Cancer Study. BMJ Open Gastroenterol. 2019;6(1):e000338.
67. Wu QJ, Yang G, Zheng W, Li HL, Gao J, Wang J, et al. Pre-diagnostic cruciferous vegetables intake and lung cancer survival among Chinese women. Scientific reports. 2015;5:10306.
68. Borch KB, Braaten T, Lund E, Weiderpass E. Physical activity before and after breast cancer diagnosis and survival - the Norwegian women and cancer cohort study. BMC cancer. 2015;15:967.
69. Harriss DJ, Cable NT, George K, Reilly T, Renehan AG, Haboubi N. Physical activity before and after diagnosis of colorectal cancer: disease risk, clinical outcomes, response pathways and biomarkers. Sports Med. 2007;37(11):947-60.
70. Kuiper JG, Phipps AI, Neuhouser ML, Chlebowski RT, Thomson CA, Irwin ML, et al. Recreational physical activity, body mass index, and survival in women with colorectal cancer. Cancer causes & control : CCC. 2012;23(12):1939-48.
71. Romaguera D, Ward H, Wark PA, Vergnaud AC, Peeters PH, van Gils CH, et al. Pre-diagnostic concordance with the WCRF/AICR guidelines and survival in European colorectal cancer patients: a cohort study. BMC medicine. 2015;13:107.
72. Walter V, Jansen L, Knebel P, Chang-Claude J, Hoffmeister M, Brenner H. Physical activity and survival of colorectal cancer patients: Population-based study from Germany. International journal of cancer. 2017;140(9):1985-97.
73. Yang JJ, Yu D, White E, Lee DH, Blot W, Robien K, et al. Prediagnosis Leisure-Time Physical Activity and Lung Cancer Survival: A Pooled Analysis of 11 Cohorts. JNCI Cancer Spectrum. 2022;6(2):pkac009.
74. Jayasekara H, English DR, Haydon A, Hodge AM, Lynch BM, Rosty C, et al. Associations of alcohol intake, smoking, physical activity and obesity with survival following colorectal cancer diagnosis by stage, anatomic site and tumor molecular subtype. International journal of cancer. 2018;142(2):238-50.
75. Irwin ML, Smith AW, McTiernan A, Ballard-Barbash R, Cronin K, Gilliland FD, et al. Influence of pre- and postdiagnosis physical activity on mortality in breast cancer survivors: the health, eating, activity, and lifestyle study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(24):3958-64.
76. Meyerhardt JA, Giovannucci EL, Holmes MD, Chan AT, Chan JA, Colditz GA, et al. Physical activity and survival after colorectal cancer diagnosis. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2006;24(22):3527-34.
77. Allemani C, Berrino F, Krogh V, Sieri S, Pupa SM, Tagliabue E, et al. Do pre-diagnostic drinking habits influence breast cancer survival? Tumori. 2011;97(2):142-8.
78. Walter V, Jansen L, Ulrich A, Roth W, Bläker H, Chang-Claude J, et al. Alcohol consumption and survival of colorectal cancer patients: a population-based study from Germany. Am J Clin Nutr. 2016;103(6):1497-506.
79. Reding KW, Daling JR, Doody DR, O'Brien CA, Porter PL, Malone KE. Effect of prediagnostic alcohol consumption on survival after breast cancer in young women. Cancer Epidemiol Biomarkers Prev. 2008;17(8):1988-96.
80. Barnett GC, Shah M, Redman K, Easton DF, Ponder BAJ, Pharoah PDP. Risk Factors for the Incidence of Breast Cancer: Do They Affect Survival From the Disease? Journal of Clinical Oncology. 2008;26(20):3310-6.
81. Holmes MD, Stampfer MJ, Colditz GA, Rosner B, Hunter DJ, Willett WC. Dietary factors and the survival of women with breast carcinoma. Cancer. 1999;86(5):826-35.

82. Franceschi S, Dal Maso L, Zucchetto A, Talamini R. Alcohol consumption and survival after breast cancer. Cancer Epidemiol Biomarkers Prev. 2009;18(3):1011-2; author reply 2-3.
83. Holm M, Olsen A, Christensen J, Kroman NT, Bidstrup PE, Johansen C, et al. Pre-diagnostic alcohol consumption and breast cancer recurrence and mortality: results from a prospective cohort with a wide range of variation in alcohol intake. International journal of cancer. 2013;132(3):686-94.
84. Baker JR, Umesh S, Jenab M, Schomburg L, Tjønneland A, Olsen A, et al. Prediagnostic Blood Selenium Status and Mortality among Patients with Colorectal Cancer in Western European Populations. Biomedicines. 2021;9(11).
85. Kyrø C, Zamora-Ros R, Scalbert A, Tjønneland A, Dossus L, Johansen C, et al. Pre-diagnostic polyphenol intake and breast cancer survival: the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Breast Cancer Res Treat. 2015;154(2):389-401.
86. Ward HA, Norat T, Overvad K, Dahm CC, Bueno-de-Mesquita HB, Jenab M, et al. Pre-diagnostic meat and fibre intakes in relation to colorectal cancer survival in the European Prospective Investigation into Cancer and Nutrition. Br J Nutr. 2016;116(2):316-25.
87. Zhu Y, Zhao J, Vallis J, Shi F, Woodrow JR, Kong Y, et al. Prediagnostic consumption of vitamin D, calcium and dairy products and colorectal cancer survival: results from the Newfoundland Colorectal Cancer Registry Cohort Study. Br J Nutr. 2021:1-10.
88. Chellappan S. Smoking Cessation after Cancer Diagnosis and Enhanced Therapy Response: Mechanisms and Significance. Curr Oncol. 2022;29(12):9956-69.
89. Rock CL, Thomson CA, Sullivan KR, Howe CL, Kushi LH, Caan BJ, et al. American Cancer Society nutrition and physical activity guideline for cancer survivors. CA: A Cancer Journal for Clinicians. 2022;72(3):230-62.
90. Friedenreich CM, McTiernan A. Combining Variables for Cancer Risk Estimation: Is the Sum Better than the Parts? Cancer prevention research (Philadelphia, Pa). 2018;11(6):313-6.
91. Ocké MC, Bueno-de-Mesquita HB, Feskens EJ, Kromhout D, Menotti A, Blackburn H. Adherence to the European Code Against Cancer in relation to long-term cancer mortality: intercohort comparisons from the Seven Countries Study. Nutrition and cancer. 1998;30(1):14-20.
92. Poortinga W. The prevalence and clustering of four major lifestyle risk factors in an English adult population. Prev Med. 2007;44(2):124-8.
93. Berrigan D, Dodd K, Troiano RP, Krebs-Smith SM, Barbash RB. Patterns of health behavior in U.S. adults. Prev Med. 2003;36(5):615-23.
94. Pronk NP, Anderson LH, Crain AL, Martinson BC, O'Connor PJ, Sherwood NE, et al. Meeting recommendations for multiple healthy lifestyle factors. Prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. American journal of preventive medicine. 2004;27(2 Suppl):25-33.
95. Meng L, Maskarinec G, Lee J, Kolonel LN. Lifestyle Factors and Chronic Diseases: Application of a Composite Risk Index. Preventive Medicine. 1999;29(4):296-304.
96. Shams-White MM, Brockton NT, Mitrou P, Romaguera D, Brown S, Bender A, et al. Operationalizing the 2018 World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR) Cancer Prevention Recommendations: A Standardized Scoring System. Nutrients. 2019;11(7).
97. Romaguera D, Vergnaud AC, Peeters PH, van Gils CH, Chan DS, Ferrari P, et al. Is concordance with World Cancer Research Fund/American Institute for Cancer Research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. Am J Clin Nutr. 2012;96(1):150-63.
98. McKenzie F, Biessy C, Ferrari P, Freisling H, Rinaldi S, Chajes V, et al. Healthy Lifestyle and Risk of Cancer in the European Prospective Investigation Into Cancer and Nutrition Cohort Study. Medicine. 2016;95(16):e2850.
99. Botteri E, Peveri G, Berstad P, Bagnardi V, Chen SLF, Sandanger TM, et al. Changes in Lifestyle and Risk of Colorectal Cancer in the European Prospective Investigation Into Cancer and Nutrition. Am J Gastroenterol. 2022.
100. Buckland G, Travier N, Huerta JM, Bueno-de-Mesquita HB, Siersema PD, Skeie G, et al. Healthy lifestyle index and risk of gastric adenocarcinoma in the EPIC cohort study. International journal of cancer. 2015;137(3):598-606.

101. Zhang Y-B, Pan X-F, Chen J, Cao A, Zhang Y-G, Xia L, et al. Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. British journal of cancer. 2020;122(7):1085-93.
102. Solans M, Chan DSM, Mitrou P, Norat T, Romaguera D. A systematic review and meta-analysis of the 2007 WCRF/AICR score in relation to cancer-related health outcomes. Ann Oncol. 2020;31(3):352-68.
103. Lofterød T, Frydenberg H, Flote V, Eggen AE, McTiernan A, Mortensen ES, et al. Exploring the effects of lifestyle on breast cancer risk, age at diagnosis, and survival: the EBBA-Life study. Breast cancer research and treatment. 2020;182(1):215-27.
104. Knudsen MD, de Lange T, Botteri E, Nguyen DH, Evensen H, Steen CB, et al. Favorable lifestyle before diagnosis associated with lower risk of screen-detected advanced colorectal neoplasia. World journal of gastroenterology. 2016;22(27):6276-86.
105. Botteri E, Berstad P, Sandin S, Weiderpass E. Lifestyle changes and risk of cancer: experience from the Swedish women's lifestyle and health cohort study. Acta Oncol. 2021:1-8.
106. Li Q, Lesseur C, Neugut AI, Santella RM, Parada H, Jr., Teitelbaum S, et al. The associations of healthy lifestyle index with breast cancer incidence and mortality in a population-based study. Breast Cancer. 2022;29(6):957-66.
107. Parada H, Jr., Sun X, Tse CK, Olshan AF, Troester MA. Lifestyle Patterns and Survival Following Breast Cancer in the Carolina Breast Cancer Study. Epidemiology. 2019;30(1):83-92.
108. Song R, Petimar J, Wang M, Tabung FK, Song M, Liu L, et al. Adherence to the World Cancer Research Fund/American Institute for Cancer Research Cancer Prevention Recommendations and Colorectal Cancer Survival. Cancer Epidemiol Biomarkers Prev. 2021;30(10):1816-25.
109. Van Blarigan EL, Fuchs CS, Niedzwiecki D, Zhang S, Saltz LB, Mayer RJ, et al. Association of Survival With Adherence to the American Cancer Society Nutrition and Physical Activity Guidelines for Cancer Survivors After Colon Cancer Diagnosis: The CALGB 89803/Alliance Trial. JAMA Oncol. 2018;4(6):783-90.
110. Hansen JM, Nagle CM, Ibiebele TI, Grant PT, Obermair A, Friedlander ML, et al. A healthy lifestyle and survival among women with ovarian cancer. International journal of cancer. 2020;147(12):3361-9.
111. Karavasiloglou N, Pestoni G, Wanner M, Faeh D, Rohrmann S. Healthy lifestyle is inversely associated with mortality in cancer survivors: Results from the Third National Health and Nutrition Examination Survey (NHANES III). PLOS ONE. 2019;14(6):e0218048.
112. Lund E, Dumeaux V, Braaten T, Hjartåker A, Engeset D, Skeie G, et al. Cohort Profile: The Norwegian Women and Cancer Study—NOWAC—Kvinner og kreft. International Journal of Epidemiology. 2007;37(1):36-41.
113. Borch KB, Ekelund U, Brage S, Lund E. Criterion validity of a 10-category scale for ranking physical activity in Norwegian women. The international journal of behavioral nutrition and physical activity. 2012;9:2.
114. The Norwegian Food Composition Database 2021 [Internet]. 2021. Available from: https://www.matvaretabellen.no/.
115. International Agency for Research on Cancer (IARC). Human Cancer: Known Causes and Prevention by Organ Site. In: IARC Monographs & IARC Handbooks Working Group, editor. 2022.
116. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet (London, England). 2003;362(9382):419-27.
117. Union for International Cancer Control. TNM Classification of Malignant Tumours, 8th edition. James D. Brierley, Mary K. Gospodarowicz, Wittekind C, editors. West Sussex: Wiley Blackwell; 2016.
118. Ruhl J.L., Callaghan C., Schussler N. (eds). Summary Stage 2018: Codes and Coding Instructions. Bethesda, MD: National Cancer Institute; 2021.
119. Posit team. RStudio: Integrated Development Environment for R. Boston, MA: Posit Software, PBC; 2023.
120. Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text, Third Edition. New York, NY: New York, NY: Springer New York; 2011.
121. Harrell FE. Regression modeling strategies. 2 ed. Cham, Switzerland: Springer International Publishing; 2015 2015/8/15. 582 p.

122. Harrell FE. Regression Modeling Strategies. 2023.
123. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. Statistics in Medicine. 1999;18(6):695-706.
124. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999;94(446):496-509.
125. Gray B. cmprsk: Subdistribution Analysis of Competing Risks. R package version 2.2-11 ed2022.
126. Cavanaugh JE, Neath AA. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. WIREs Computational Statistics. 2019;11(3):e1460.
127. Rubin DB. Multiple Imputation for Nonresponse in Surveys. United States of America: John Wiley & Sons, Inc.; 1987.
128. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011;45(3):1 - 67.
129. Rubin DB. 1.5. Multiple Imputation. Multiple Imputation for Nonresponse in Surveys. United States of America: John Wiley & Sons, Inc.; 1987. p. 15-8.
130. UiT The Artic University of Norway. Kvinner og Kreft [The Norwegian Women and Cancer Study]. 2023 [Available from: https://uit.no/research/nowac]
131. UiT The Artic University of Norway. Information security and data protection at UiT. UiT. 2023 [Available from: https://uit.no/om/informationsecurity#innhold_683826]
132. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 4 ed. Sudbury: Sudbury: Jones & Bartlett Learning, LLC; 2018.
133. Rothman KJ. Modern epidemiology. 3rd ed. Greenland S, Lash TL, Buehler JW, Cahill J, Glymour MM, Willett W, editors. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2008.
134. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet (London, England). 2002;359(9302):248-52.
135. Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. Am J Epidemiol. 1991;134(10):1233-44.
136. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. Obes Rev. 2007;8(4):307-26.
137. Skeie G, Mode N, Henningsen M, Borch KB. Validity of self-reported body mass index among middle-aged participants in the Norwegian Women and Cancer study. Clinical epidemiology. 2015;7:313-23.
138. Flegal KM, Brownie C, Haas JD. The effects of exposure misclassification on estimates of relative risk. Am J Epidemiol. 1986;123(4):736-50.
139. Ogden MW, Morgan WT, Heavner DL, Davis RA, Steichen TJ. National incidence of smoking and misclassification among the US married female population. Journal of Clinical Epidemiology. 1997;50(3):253-63.
140. Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. Am J Public Health. 1994;84(7):1086-93.
141. Willett W. Nutritional epidemiology. Oxford: Oxford University Press; 2013.
142. Hjartåker A, Andersen LF, Lund E. Comparison of diet measures from a food-frequency questionnaire with measures from repeated 24-hour dietary recalls. The Norwegian Women and Cancer Study. Public health nutrition. 2007;10(10):1094-103.
143. Parr CL, Veierod MB, Laake P, Lund E, Hjartaker A. Test-retest reproducibility of a food frequency questionnaire (FFQ) and estimated effects on disease risk in the Norwegian Women and Cancer Study (NOWAC). Nutr J. 2006;5:4.
144. Larsen IK, Småstuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur J Cancer. 2009;45(7):1218-31.

145. Lund E, Kumle M, Braaten T, Hjartaker A, Bakken K, Eggen E, et al. External validity in a population-based national prospective study--the Norwegian Women and Cancer Study (NOWAC). Cancer causes & control : CCC. 2003;14(10):1001-8.
146. Pedersen AG, Ellingsen CL. Data quality in the Causes of Death Registry. Tidsskr Nor Laegeforen. 2015;135(8):768-70.
147. Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull World Health Organ. 2005;83(3):171-7.
148. Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon FC, Rao C, et al. Civil registration systems and vital statistics: successes and missed opportunities. Lancet (London, England). 2007;370(9599):1653-63.
149. Phillips DE, Lozano R, Naghavi M, Atkinson C, Gonzalez-Medina D, Mikkelsen L, et al. A composite metric for assessing data on mortality and causes of death: the vital statistics performance index. Popul Health Metr. 2014;12:14.
150. Hall GC, Lanes S, Bollaerts K, Zhou X, Ferreira G, Gini R. Outcome misclassification: Impact, usual practice in pharmacoepidemiology database studies and an online aid to correct biased estimates of risk ratio or cumulative incidence. Pharmacoepidemiology and Drug Safety. 2020;29(11):1450-5.
151. Groenwold RHH, Klungel OH, Altman DG, van der Graaf Y, Hoes AW, Moons KGM, et al. Adjustment for continuous confounders: an example of how to prevent residual confounding. CMAJ. 2013;185(5):401-6.
152. Promislow JHE, Gladen BC, Sandler DP. Maternal Recall of Breastfeeding Duration by Elderly Women. Am J Epidemiol. 2005;161(3):289-96.
153. Avancini A, Sartori G, Gkountakos A, Casali M, Trestini I, Tregnago D, et al. Physical Activity and Exercise in Lung Cancer Care: Will Promises Be Fulfilled? Oncologist. 2020;25(3):e555-e69.
154. Lindberg MH, Chen G, Olsen JA, Abelsen B. Explaining subjective social status in two countries: The relative importance of education, occupation, income and childhood circumstances. SSM - Population Health. 2021;15:100864.
155. Delaruelle K, Buffel V, Bracke P. Educational expansion and the education gradient in health: A hierarchical age-period-cohort analysis. Soc Sci Med. 2015;145:79-88.
156. Norwegian Directorate of Health. Pakkeforløp for kreft – Generell informasjon for alle pakkeforløpene for kreft [Cancer patient pathways - general information for all CPPs for cancer]. 2023 [Available from: https://www.helsedirektoratet.no/nasjonale-forlop/generell-informasjon-for-alle-pakkeforlopene-for-kreft]
157. Olsen F, Jacobsen BK, Heuch I, Tveit KM, Balteskard L. Equitable access to cancer patient pathways in Norway – a national registry-based study. BMC Health Services Research. 2021;21(1):1272.
158. Fiva JH, Hægeland T, Rønning M, Syse A. Access to treatment and educational inequalities in cancer survival Statistics Norway (SSB); 2013.
159. Fewell Z, Davey Smith G, Sterne JAC. The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study. Am J Epidemiol. 2007;166(6):646-55.
160. Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. International Journal of Epidemiology. 2013;42(4):1012-4.
161. Thiébaut ACM, Bénichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. Statistics in Medicine. 2004;23(24):3803-20.
162. Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. Am J Epidemiol. 1997;145(1):72-80.
163. Chalise P, Chicken E, McGee D. Time Scales in Epidemiological Analysis: An Empirical Comparison. International Journal of Statistics and Probability. 2016;5(3).
164. Kleinbaum DG, Klein M. The Stratified Cox Procedure. In: M. Gail, K. Krickeberg, J. Samet, A. Tsiatis, W. Wong, editors. Survival Analysis: A Self-Learning Text. 3rd ed. New York, NY: Springer New York; 2011. p. 173-210.
165. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016;133(6):601-9.

166. Tsiatis A. A nonidentifiability aspect of the problem of competing risks. Proc Natl Acad Sci U S A. 1975;72(1):20-2.
167. Allison P. For Causal Analysis of Competing Risks, Don't Use Fine & Gray's Subdistribution Method. Statistical Horizons. 2018 [Available from: https://statisticalhorizons.com/for-causal-analysis-of-competing-risks/]
168. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. International Journal of Epidemiology. 2012;41(3):861-70.
169. Lau B, Cole SR, Gange SJ. Competing Risk Regression Models for Epidemiologic Data. Am J Epidemiol. 2009;170(2):244-56.
170. Greenland S. Avoiding power loss associated with categorization and ordinal scores in dose-response and trend analysis. Epidemiology. 1995;6(4):450-4.
171. Stone CJ, Koo C-Y. Additive splines in statistics. Proceedings of the American Statistical Association Original pagination is p. 1985;45:48.
172. Rubin DB. Inference and Missing Data. Biometrika. 1976;63(3):581-92.
173. Little RJA, Rubin DB. Statistical Analysis with Missing Data., 3rd Edition: Wiley; 2019.
174. Bartlett JW, Harel O, Carpenter JR. Asymptotically Unbiased Estimation of Exposure Odds Ratios in Complete Records Logistic Regression. Am J Epidemiol. 2015;182(8):730-6.
175. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Statistics in Medicine. 2010;29(28):2920-31.
176. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj. 2009;338:b2393.
177. Rubin DB. Multiple imputation for nonresponse in surveys. Hoboken: WILEY; 2009.
178. Von Hippel PT. HOW TO IMPUTE INTERACTIONS, SQUARES, AND OTHER TRANSFORMED VARIABLES. Sociological Methodology. 2009;39(1):265-91.
179. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377-99.
180. van Buuren S. Item Imputation Without Specifying Scale Structure. Methodology. 2010;6(1):31-6.
181. van Buuren S. Flexible Imputation of Missing Data. Boca Raton, Florida: Taylor & Francis Group; 2018. Available from: https://stefvanbuuren.name/fimd/.
182. Norwegian Cancer Society. Prevention [forebygging]. 2023 [Available from: From a practical standpoint, this used an appreciable amount of time that could have been perhaps better afforded to other important aspects of the studies.]
183. Aleksandrova K, Pischon T, Jenab M, Bueno-de-Mesquita HB, Fedirko V, Norat T, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC medicine. 2014;12:168.
184. Arthur R, Wassertheil-Smoller S, Manson JE, Luo J, Snetselaar L, Hastert T, et al. The Combined Association of Modifiable Risk Factors with Breast Cancer Risk in the Women's Health Initiative. Cancer prevention research (Philadelphia, Pa). 2018;11(6):317-26.
185. Jiao L, Mitrou PN, Reedy J, Graubard BI, Hollenbeck AR, Schatzkin A, et al. A combined healthy lifestyle score and risk of pancreatic cancer in a large cohort study. Arch Intern Med. 2009;169(8):764-70.
186. Bhaskaran K, Dos-Santos-Silva I, Leon DA, Douglas IJ, Smeeth L. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. Lancet Diabetes Endocrinol. 2018;6(12):944-53.
187. The Global BMI Mortality Collaboration. Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. Lancet (London, England). 2016;388(10046):776-86.
188. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. Lancet (London, England). 2014;384(9945):755-65.
189. Wang QL, Xie SH, Li WT, Lagergren J. Smoking Cessation and Risk of Esophageal Cancer by Histological Type: Systematic Review and Meta-analysis. Journal of the National Cancer Institute. 2017;109(12).

190. McKenzie F, Ferrari P, Freisling H, Chajes V, Rinaldi S, de Batlle J, et al. Healthy lifestyle and risk of breast cancer among postmenopausal women in the European Prospective Investigation into Cancer and Nutrition cohort study. International journal of cancer. 2015;136(11):2640-8.
191. Wang P, Song M, Eliassen AH, Wang M, Fung TT, Clinton SK, et al. Optimal dietary patterns for prevention of chronic disease. Nature Medicine. 2023;29(3):719-28.
192. Streiner DL, Norman GR, Cairney J. Health Measurement Scales : A Practical Guide to Their Development and Use. Oxford: Oxford: Oxford University Press; 2014.
193. Arthur R, Kirsh VA, Kreiger N, Rohan T. A healthy lifestyle index and its association with risk of breast, endometrial, and ovarian cancer among Canadian women. Cancer causes & control : CCC. 2018;29(6):485-93.
194. McKenzie F, Ellison-Loschmann L, Jeffreys M, Firestone R, Pearce N, Romieu I. Healthy lifestyle and risk of breast cancer for indigenous and non-indigenous women in New Zealand: a case control study. BMC cancer. 2014;14:12.
195. Sánchez-Zamorano LM, Flores-Luna L, Angeles-Llerenas A, Romieu I, Lazcano-Ponce E, Miranda-Hernández H, et al. Healthy lifestyle on the risk of breast cancer. Cancer Epidemiol Biomarkers Prev. 2011;20(5):912-22.
196. Dartois L, Fagherazzi G, Boutron-Ruault MC, Mesrine S, Clavel-Chapelon F. Association between five lifestyle habits and cancer risk: results from the E3N cohort. Cancer prevention research (Philadelphia, Pa). 2014;7(5):516-25.
197. Romaguera D, Gracia-Lavedan E, Molinuevo A, de Batlle J, Mendez M, Moreno V, et al. Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case-control study. International journal of cancer. 2017;141(1):83-93.
198. Kabat GC, Matthews CE, Kamensky V, Hollenbeck AR, Rohan TE. Adherence to cancer prevention guidelines and cancer incidence, cancer mortality, and total mortality: a prospective cohort study. Am J Clin Nutr. 2015;101(3):558-69.
199. Collaborative Group on Hormonal Factors in Breast Cancer. Type and timing of menopausal hormone therapy and breast cancer risk: individual participant meta-analysis of the worldwide epidemiological evidence. The Lancet. 2019;394(10204):1159-68.
200. Wang K, Ma W, Wu K, Ogino S, Chan AT, Giovannucci EL, et al. Healthy lifestyle, endoscopic screening, and colorectal cancer incidence and mortality in the United States: A nationwide cohort study. PLoS Med. 2021;18(2):e1003522.
201. Kirkegaard H, Johnsen NF, Christensen J, Frederiksen K, Overvad K, Tjønneland A. Association of adherence to lifestyle recommendations and risk of colorectal cancer: a prospective Danish cohort study. BMJ. 2010;341:c5504.
202. Gram IT, Braaten T, Lund E, Le Marchand L, Weiderpass E. Cigarette smoking and risk of colorectal cancer among Norwegian women. Cancer causes & control : CCC. 2009;20(6):895-903.
203. Oyeyemi SO, Braaten T, Botteri E, Berstad P, Borch KB. Exploring geographical differences in the incidence of colorectal cancer in the Norwegian Women and Cancer Study: a population-based prospective study. Clinical epidemiology. 2019;11:669-82.
204. Steck SE, Murphy EA. Dietary patterns and cancer risk. Nature Reviews Cancer. 2020;20(2):125-38.
205. Norwegian Institute of Public Health. Smoking and snus use in Norway. 2016 [Available from: https://www.fhi.no/en/op/hin/lifestyle/royking-og-snusbruk-i-noreg/]
206. Arthur R, Wassertheil-Smoller S, Rohan TE, Brasky TM, Crane TE, Felix AS, et al. Associations of a Healthy Lifestyle Index With the Risks of Endometrial and Ovarian Cancer Among Women in the Women's Health Initiative Study. Am J Epidemiol. 2019;188(2):261-73.
207. Collaborative Group on Epidemiological Studies of Ovarian Cancer. Ovarian cancer and body size: individual participant meta-analysis including 25,157 women with ovarian cancer from 47 epidemiological studies. PLoS Med. 2012;9(4):e1001200.
208. Naudin S, Viallon V, Hashim D, Freisling H, Jenab M, Weiderpass E, et al. Healthy lifestyle and the risk of pancreatic cancer in the EPIC study. European journal of epidemiology. 2020;35(10):975-86.
209. Malone JI, Hansen BC. Does obesity cause type 2 diabetes mellitus (T2DM)? Or is it the opposite? Pediatr Diabetes. 2019;20(1):5-9.

210. Bracci PM. Obesity and pancreatic cancer: overview of epidemiologic evidence and biologic mechanisms. Mol Carcinog. 2012;51(1):53-63.
211. Barberio AM, Alareeki A, Viner B, Pader J, Vena JE, Arora P, et al. Central body fatness is a stronger predictor of cancer risk than overall body size. Nature Communications. 2019;10(1):383.
212. Meer R, van de Pol J, van den Brandt PA, Schouten LJ. The association of healthy lifestyle index score and the risk of renal cell cancer in the Netherlands cohort study. BMC cancer. 2023;23(1):156.
213. Norwegian Directorate of Health. Statistikk og historikk om røyking, snus og e-sigaretter [Statistics and history for smoking, snus, and e-cigarettes]. Oslo: Norwegian Directorate of Health. 2019 [updated 18 January 2022. Available from: https://www.helsedirektoratet.no/tema/tobakk-royk-og-snus/statistikk-om-royking-bruk-av-snus-og-e-sigaretter]
214. Vedøy TF, Skretting A. Alcohol use among women [Bruk av alkohol blant kvinner]. Oslo: The Norwegian Institute for Alcohol and Drug Research (SIRUS); 2009.
215. Norwegian Institute of Public Health. Overweight and obesity in Norway. 2017 [Available from: https://www.fhi.no/en/op/hin/health-disease/overweight-and-obesity-in-norway---/]
216. Hill AB. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine. 1965;58(5):295-300.
217. Sand AS, Emaus N, Lian OS. Motivation and obstacles for weight management among young women - a qualitative study with a public health focus - the Tromsø study: Fit Futures. BMC Public Health. 2017;17(1):417.
218. Perera LAM, Chopra A, Shaw AL. Approach to Patients with Unintentional Weight Loss. Med Clin North Am. 2021;105(1):175-86.
219. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. International Journal of Epidemiology. 2005;34(1):215-20.
220. Freisling H, Viallon V, Lennon H, Bagnardi V, Ricci C, Butterworth AS, et al. Lifestyle factors and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. BMC medicine. 2020;18(1):5.
221. Majed B, Moreau T, Senouci K, Salmon RJ, Fourquet A, Asselain B. Is obesity an independent prognosis factor in woman breast cancer? Breast Cancer Research and Treatment. 2008;111(2):329-42.
222. Abrahamson PE, Gammon MD, Lund MJ, Britton JA, Marshall SW, Flagg EW, et al. Recreational physical activity and survival among young women with breast cancer. Cancer. 2006;107(8):1777-85.
223. Schmidt ME, Chang-Claude J, Vrieling A, Seibold P, Heinz J, Obi N, et al. Association of pre-diagnosis physical activity with recurrence and mortality among women with breast cancer. International journal of cancer. 2013;133(6):1431-40.
224. Yu D, Takata Y, Smith-Warner SA, Blot W, Sawada N, White E, et al. Prediagnostic Calcium Intake and Lung Cancer Survival: A Pooled Analysis of 12 Cohort Studies. Cancer Epidemiol Biomarkers Prev. 2017;26(7):1060-70.
225. Lynge E, Braaten T, Njor SH, Olsen AH, Kumle M, Waaseth M, et al. Mammography activity in Norway 1983 to 2008. Acta Oncologica. 2011;50(7):1062-7.
226. Norwegian Directorate of Health. Screeningprogram for kreft [Cancer screening programs]. 2022 [Available from: https://www.helsedirektoratet.no/tema/kreft/screeningprogram-for-kreftf#:~:text=Fra%202022%20er%20det%20nasjonale,Kreftregisterets%20nettsider%20(kreftregisteret.no)]
227. de Boer MC, Wörner EA, Verlaan D, van Leeuwen PAM. The Mechanisms and Effects of Physical Activity on Breast Cancer. Clin Breast Cancer. 2017;17(4):272-8.
228. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer. 2003;3(10):733-44.
229. Boffetta P, Hashibe M. Alcohol and cancer. Lancet Oncol. 2006;7(2):149-56.
230. Rumgay H, Murphy N, Ferrari P, Soerjomataram I. Alcohol and Cancer: Epidemiology and Biological Mechanisms. Nutrients. 2021;13(9).
231. Furman D, Campisi J, Verdin E, Carrera-Bastos P, Targ S, Franceschi C, et al. Chronic inflammation in the etiology of disease across the life span. Nature Medicine. 2019;25(12):1822-32.

232. Cucina A, Dinicola S, Coluccia P, Proietti S, D'Anselmi F, Pasqualato A, et al. Nicotine stimulates proliferation and inhibits apoptosis in colon cancer cell lines through activation of survival pathways. Journal of Surgical Research. 2012;178(1):233-41.
233. Fares J, Fares MY, Khachfe HH, Salhab HA, Fares Y. Molecular principles of metastasis: a hallmark of cancer revisited. Signal Transduction and Targeted Therapy. 2020;5(1):28.
234. Twisk JWR. Applied Mixed Model Analysis: A Practical Guide. 2 ed. Cambridge: Cambridge University Press; 2019.
235. Friedenreich CM. The role of physical activity in breast cancer etiology. Semin Oncol. 2010;37(3):297-302.
236. Keegan TH, Milne RL, Andrulis IL, Chang ET, Sangaramoorthy M, Phillips KA, et al. Past recreational physical activity, body size, and all-cause mortality following breast cancer diagnosis: results from the Breast Cancer Family Registry. Breast Cancer Res Treat. 2010;123(2):531-42.
237. Bandera EV, Fay SH, Giovannucci E, Leitzmann MF, Marklew R, McTiernan A, et al. The use and interpretation of anthropometric measures in cancer epidemiology: A perspective from the world cancer research fund international continuous update project. International journal of cancer. 2016;139(11):2391-7.
238. Larsson SC, Spyrou N, Mantzoros CS. Body fatness associations with cancer: evidence from recent epidemiological studies and future directions. Metabolism. 2022;137:155326.
239. Hang J, Cai B, Xue P, Wang L, Hu H, Zhou Y, et al. The Joint Effects of Lifestyle Factors and Comorbidities on the Risk of Colorectal Cancer: A Large Chinese Retrospective Case-Control Study. PloS one. 2015;10(12):e0143696.
240. Zhang Q-L, Zhao L-G, Li H-L, Gao J, Yang G, Wang J, et al. The joint effects of major lifestyle factors on colorectal cancer risk among Chinese men: A prospective cohort study. International journal of cancer. 2018;142(6):1093-101.
241. Erben V, Carr PR, Holleczek B, Stegmaier C, Hoffmeister M, Brenner H. Strong associations of a healthy lifestyle with all stages of colorectal carcinogenesis: Results from a large cohort of participants of screening colonoscopy. International journal of cancer. 2019;144(9):2135-43.
242. Engeset D, Alsaker E, Ciampi A, Lund E. Dietary patterns and lifestyle factors in the Norwegian EPIC cohort: the Norwegian Women and Cancer (NOWAC) study. Eur J Clin Nutr. 2005;59(5):675-84.
243. Loomba R, Yang H-I, Su J, Brenner D, Barrett-Connor E, Iloeje U, et al. Synergism Between Obesity and Alcohol in Increasing the Risk of Hepatocellular Carcinoma: A Prospective Cohort Study. Am J Epidemiol. 2013;177(4):333-42.
244. Ocke MC. Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. The Proceedings of the Nutrition Society. 2013;72(2):191-9.
245. Zhang Q-L, Zhao L-G, Zhang W, Li H-L, Gao J, Han L-H, et al. Combined Impact of Known Lifestyle Factors on Total and Cause-Specific Mortality among Chinese Men: A Prospective Cohort Study. Scientific reports. 2017;7(1):5293.
246. Whiteman DC, Wilson LF. The fractions of cancer attributable to modifiable factors: A global review. Cancer Epidemiol. 2016;44:203-21.
247. Carr PR, Weigl K, Jansen L, Walter V, Erben V, Chang-Claude J, et al. Healthy Lifestyle Factors Associated With Lower Risk of Colorectal Cancer Irrespective of Genetic Risk. Gastroenterology. 2018;155(6):1805-15.e5.
248. Lucini D, Zanuso S, Blair S, Pagani M. A simple healthy lifestyle index as a proxy of wellness: a proof of concept. Acta Diabetologica. 2015;52(1):81-9.
249. World Cancer Research Fund International / American Institute for Cancer Research (WCRF/AICR). Cancer prevention recommendations. 2023 [Available from: https://www.wcrf.org/diet-activity-and-cancer/cancer-prevention-recommendations/]
250. World Health Organization (WHO). Cancer. 2018 [Available from: https://www.who.int/news-room/fact-sheets/detail/cancer]
251. Emmons KM, Colditz GA. Realizing the Potential of Cancer Prevention - The Role of Implementation Science. The New England journal of medicine. 2017;376(10):986-90.
252. Aranceta Bartrina J. Public health and the prevention of obesity: failure or success? Nutr Hosp. 2013;28 Suppl 5:128-37.

253. World Health Organization (WHO). Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva: World Health Organization (WHO); 2018.
254. Statistics Norway. Health, care and social relations, survey on living conditions. 2020 [Available from: https://www.ssb.no/en/helse/helseforhold-og-levevaner/statistikk/helseforhold-levekarsundersokelsen]
255. Midthjell K, Lee CM, Langhammer A, Krokstad S, Holmen TL, Hveem K, et al. Trends in overweight and obesity over 22 years in a large adult population: the HUNT Study, Norway. Clin Obes. 2013;3(1-2):12-20.
256. Slade T, Chapman C, Swift W, Keyes K, Tonks Z, Teesson M. Birth cohort trends in the global epidemiology of alcohol use and alcohol-related harms in men and women: systematic review and metaregression. BMJ Open. 2016;6(10):e011827.
257. Pennay A, Lubman DI, Frei M. Alcohol: prevention, policy and primary care responses. Aust Fam Physician. 2014;43(6):356-61.
258. Norwegian Directorate of Health. Tobacco Control in Norway. 2021 [Available from: https://www.helsedirektoratet.no/english/tobacco-control-in-norway]
259. Statistics Norway. 08250: Lifestyle habits (per cent), by contents, living habit, year, age and level of education. 2023 [Available from: https://www.ssb.no/en/statbank/table/08250/]
260. Meyer HE, Tverdal A. Development of body weight in the Norwegian population. Prostaglandins Leukot Essent Fatty Acids. 2005;73(1):3-7.
261. Subramaniam M, Devi F, AshaRani PV, Zhang Y, Wang P, Jeyagurunathan A, et al. Barriers and facilitators for adopting a healthy lifestyle in a multi-ethnic population: A qualitative study. PLOS ONE. 2022;17(11):e0277106.
262. Wright SM, Aronne LJ. Causes of obesity. Abdominal Radiology. 2012;37(5):730-2.

Figures 1 and 6 created by BioRender.com

Paper I

Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Combined Lifestyle Behaviors and the Incidence of Common Cancer Types in the Norwegian Women and Cancer Study (NOWAC)

Sairah LF Chen [1]
Tonje Braaten [1]
Kristin B Borch [1]
Pietro Ferrari [2]
Torkjel M Sandanger [1]
Therese H Nøst [1]

[1]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway; [2]Nutrition and Metabolism Branch, International Agency for Research on Cancer, World Health Organization, Lyon, France

Introduction: Only a small number of studies have examined the impact of combined lifestyle behaviors on cancer incidence, and never in a Norwegian population.

Purpose: To examine linear and nonlinear associations of combined lifestyle factors, assessed through a healthy lifestyle index (HLI), with the incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, postmenopausal ovarian, pancreatic, and kidney cancer among women in Norway.

Methods: This prospective study included 96,869 women enrolled in the Norwegian Women and Cancer (NOWAC) cohort. Baseline information on lifestyle factors was collected between 1996 and 2004. The HLI was constructed from five lifestyle factors: physical activity level, body mass index, smoking, alcohol consumption, and diet. Each factor contributed 0 to 4 points to the HLI score, which ranged from 0 to 20, with higher scores representing a healthier lifestyle. Multiple imputation was used to handle missing data. Cox proportional hazard regression models were used to estimate hazard ratios (HR) and 95% confidence intervals (CI). Restricted cubic splines were used to examine nonlinearity in the associations.

Results: The HRs for a one-point increment on the HLI score were 0.97 (95% CI: 0.96–0.98) for postmenopausal breast cancer, 0.98 (0.96–1.00) for colorectal cancer, 0.86 (0.84–0.87) for lung cancer, 0.93 (0.91–0.95) for postmenopausal endometrial cancer, 0.99 (0.96–1.02) for postmenopausal ovarian cancer, 0.92 (0.89–0.95) for pancreatic cancer, and 0.94 (0.91–0.97) for kidney cancer. Nonlinearity was observed for the inverse associations between HLI score and the incidence of lung cancer and postmenopausal breast cancer.

Conclusion: Based on our results, healthier lifestyle, as assessed by the HLI score, was associated with lower incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, pancreatic, and kidney cancer among women, although the magnitude and linearity varied. Adoption of healthier lifestyle behaviors should be a public health priority to reduce the cancer burden among Norwegian women.

Keywords: healthy lifestyle index, cancer prevention, prospective study, composite score

Introduction

Cancer is the second-leading cause of death worldwide,[1] with estimated 19.3 million new cancer cases and 10.0 million cancer deaths in 2020.[2] The latest corresponding numbers in Norway were 34,190 and 11,049 in 2018.[3] Women accounted for 46% of the new cases in Norway, where breast cancer remains the most common, followed by colorectal cancer and lung cancer.[3] According to the latest report from the Cancer Registry of Norway, age-standardized incidence rates

Correspondence: Sairah LF Chen
Department of Community Medicine, UiT The Arctic University of Norway, Hansine Hansens veg 18, Tromsø, 9019, Norway
Tel +47 77644786
Email sairah.chen@uit.no

Received: 12 April 2021
Accepted: 22 June 2021
Published: 16 August 2021

for most cancers are increasing, as are the number of incident cases.[3] This trend suggests that the already substantial cancer burden in Norway will continue to grow, placing increasing pressure on the healthcare system in the form of screening and treatment.

It has been estimated by the World Health Organization that 30% to 50% of all cancer cases are related to modifiable factors.[4] Lifestyle factors – namely physical inactivity, overweight and obesity, smoking habits, alcohol consumption, and diet – have been repeatedly identified as cancer risk factors.[5,6] There is substantial evidence that a large proportion of cancers can be prevented through the adoption of a healthier lifestyle, providing an optimistic avenue for decreasing the future cancer burden. While early diagnosis and treatment for the most common cancers have been improving, they remain a challenge, highlighting the importance of preventive strategies.

Epidemiological studies typically aim to isolate the relationship between single lifestyle factors and cancer risk. While these analyses are critical for identifying novel risk factors and the strength of associations in different populations,[7–9] they cannot assess the combined impact of several healthy or non-healthy behaviors. An alternate approach is to assess the effects of a combination of lifestyle factors on cancer risk. This concept is increasingly being used to explore the cancer-preventing benefits of an overall healthy lifestyle. Several studies have employed additive exposure scores – either based strictly on recommendations from public health bodies, or based on a combination of recommendations, current scientific knowledge, and sample-specific attributes – referred to as a healthy lifestyle index (HLI). These studies have, for the most part, observed linear risk decreases with increasing increments in their indices.[10] To our knowledge, nonlinearity in associations between a HLI and cancer incidence has only been explored in one study on lymphoma incidence.[11]

The present study aims to examine the linear and non-linear associations of combined lifestyle factors, assessed through a score on an a priori-defined HLI, with the incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, postmenopausal ovarian, pancreatic, and kidney cancer among women in Norway. To our knowledge, this is the first prospective study to explore associations between a combined measure of lifestyle, including diet, and the incidence of cancer in a Norwegian population.

Materials and Methods

Study Population and Data Collection

A detailed description of the Norwegian Women and Cancer Study (NOWAC) has been presented elsewhere.[12] Briefly, the NOWAC study is a nationwide, prospective cohort study, consisting of approximately 172,000 adult female participants. Women invited to participate in the NOWAC study were randomly sampled from the Norwegian Central Person Register between 1991 and 2007 in multiple sub-cohorts. Those who agreed to participate completed a first self-administered questionnaire. Second questionnaires were sent to all sub-cohorts, except for those enrolled from 2005 to 2007 (n = 42,671), and approximately 70% of participants responded. All questionnaires collected information on socio-demographic characteristics, reproductive and hormonal factors, self-reported health, physical activity, height, weight, smoking habits, dietary habits, and family history of breast cancer. Questionnaires consisted of either 4 or 8 pages depending on the sub-cohort, with the 8-page questionnaire containing a detailed food frequency questionnaire (FFQ).

The first completed 8-page questionnaire was used as the baseline for the present study. Therefore, sub-cohorts that did not complete an 8-page questionnaire as their first or second questionnaire were excluded (n = 71,210), leaving a total of 101,316 women available for this analysis. This number included first questionnaires for sub-cohorts enrolled from 1996 to 1997 (response rate: 57%) and 2003 to 2004 (response rate: 48%), and the second questionnaires, administered from 1998 to 1999, for sub-cohorts enrolled from 1991 to 1992. Thus, year at baseline ranged from 1996 to 2004. Women with prevalent cancer, those who died or emigrated before baseline, and those with extreme energy intakes (<2100 or >15,000 KJ/day) were excluded, leaving 96,869 cancer-free participants in the final study sample [see Figure S1 in Supplementary File].

Exposure Assessment and Construction

To capture overall lifestyle in one measure, relevant lifestyle factors were combined into a HLI, which was a priori based on public health recommendations for cancer prevention from the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR)[6] and current scientific knowledge. Thus, the HLI used for this analysis consisted of five modifiable lifestyle factors – physical activity level, body fatness, smoking status, alcohol consumption, and diet – and each was assigned a score

ranging from 0 to 4, with higher scores indicating a healthier lifestyle. Physical activity level was reported by participants on a 10-point scale ranging from not active to very active, where participants were asked to consider the entirety of activity at work, outside work, at home, exercise, and other forms of physical activity. Since this measure could not be categorized by cancer guidelines or other measures of the dose of physical activity, physical activity level was scored by quintile (highest quintile = 4 through to lowest quintile = 0). Body fatness was assessed by self-reported height (centimeters) and weight (kilograms) to calculated body mass index (BMI, <23 = 4, 23 to <25 = 3, 25 to <27 = 2, 27 to <30= 1, ≥30 = 0), smoking status was scored considering intensity and time since cessation (never smoker = 4, former smoker >10 years since cessation = 3, former smoker ≤10 years since cessation = 2, smoker <15 cigarettes/day = 1, current smoker ≥15 cigarettes/day = 0), and alcohol consumption was recorded in grams/day (none = 4, >0 to <5 = 3, 5 to <10 = 2, 10 to <20 = 1, >20 = 0). To quantify diet, a diet-specific score was generated, which ranged from 0 to 18, with 18 considered the healthiest diet. Six food groups were included in this score: whole grains, fruit, vegetables, dairy, red meat, and processed meat. Using the Norwegian Weight and Measurement Table,[13] grams of intake per day were estimated for each food group based on the frequency and portions reported in the FFQ. Each food group was adjusted for energy intake, by dividing grams of intake by daily energy intake, in millijoules (MJ). The energy-adjusted food groups were categorized into quartiles and scored from 0 (lowest quartile) to 3 (highest quartile). Red and processed meat were scored in reverse order. The 18-point diet score was then divided into quintiles for inclusion in the HLI (highest quintile = 4 through to lowest quintile = 0) (Table 1).

The scores for each lifestyle factor were summed to obtain the HLI score, which ranged from 0 to 20, with a score of 20 being considered the healthiest. Physical activity level,[14] BMI,[15] and the FFQ[16,17] have been validated in the NOWAC study.

Outcome Assessment

Women diagnosed with incident cancer after baseline were identified through linkage to the Cancer Registry of Norway, based on codes from the International Classification of Diseases for Oncology, Third Edition (ICD-O-3):[18] breast cancer [ICD-O-3 code C50], colorectal cancer [C18-20], lung cancer [C34], endometrial cancer [C54], ovarian cancer [C56], pancreatic cancer [C25], and kidney cancer [C64]. Breast, endometrial, and ovarian cancers were considered only among postmenopausal women due to 1) known etiological and risk-factor profile differences between pre- and postmenopausal statuses and 2) inadequate numbers of premenopausal cancer cases. Information on emigration and mortality was obtained through linkage to the Norwegian National Population Register and the national Cause of Death Registry, respectively. The end of the follow-up was 15 December 2018.

Statistical Analysis

Descriptive statistics were computed using means and standard deviations or medians and interquartile ranges, depending on distribution. Cox proportional hazard regression models were used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for associations between the HLI score and the incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, postmenopausal ovarian, pancreatic, and kidney cancer. Age was used as the underlying time-scale,[19] whereby study entry was defined as age at baseline or age at menopause for postmenopausal breast, endometrial, and ovarian cancer models if the participant was not postmenopausal at baseline. Women were categorized as postmenopausal if they

Table 1 Healthy Lifestyle Index Scoring System Combining Five Lifestyle Factors Additively

Score	Physical Activity Level (10-Point Scale)	BMI	Smoking Status	Alcohol Consumption (g/Day)	Diet (0–18 Score)
0	1–3	≥30	Current, ≥15 cig/day	≥20	0–6
1	4	27.0–29.9	Current, <15 cig/day	10.0–19.9	7
2	5	25.0–26.9	Former,<10yrs since cessation	5.0–9.9	8–9
3	6	23.0–24.9	Former,≥10yrs since cessation	>0.0–4.9	10
4	7–10	<23	Never	0	11–18

Abbreviations: BMI, body mass index; g/day, grams per day; cig/day, cigarettes smoked per day; yrs, years.

reported that their menstruation had stopped, reported use of hormone replacement therapy, or if they were ≥53 years of age, to maintain consistency with previous publications from the NOWAC study[20,21] and the Million Women Study.[22] Exit time was defined as age at cancer diagnosis, death, emigration, or end of follow-up, whichever occurred first.

The HLI score was first modelled as a continuous variable to estimate HRs corresponding to a one-point increase in the score. Categorical analyses were also carried out by dividing the HLI score into 4 groups (0–5 HLI group, 6–10 HLI group, 11–15 HLI group, and 16–20 HLI group), using the 11–15 HLI group as a reference. The proportional hazards assumption was assessed by Schoenfeld residuals.[23] The reverse Kaplan–Meier method was used to calculate the median follow-up duration.[24]

The selection of covariates for adjusted models was done on a per-outcome basis, including known risk factors for the exposure, the outcome, or both.[25] As such, all models were adjusted for education (years) and height (centimeters). Models for postmenopausal breast and postmenopausal endometrial cancer were additionally adjusted for age at menarche (years), use of oral contraceptives (ever, never), parity (0, 1–2, >2), breastfeeding (cumulative months 0, <12 months, ≥12 months), and use of hormone replacement therapy (current, former, never). Models for postmenopausal breast cancer were further adjusted for the history of breast cancer in first-degree relatives (yes, no). The above analyses were conducted on multiple imputed data (described later).

Nonlinear dose–response relationships between the HLI score and the incidence of the included cancer types were modelled with restricted cubic splines. The Akaike information criterion (AIC) was used to evaluate the number of knots and positioning for the best fit. Nonlinearity was assessed through visual inspection of plots and comparison of linear and nonlinear model AIC values.

Several sensitivity analyses were conducted using the HLI score as a continuous variable. To evaluate the driving contributions of each lifestyle factor to the overall observed associations, reduced models were created, each of which excluded one lifestyle factor from the HLI score, thus producing five separate, reduced models for each cancer type. The factor that was excluded from the HLI score was included as a confounder in their respective model, and HRs from these reduced models were compared to those from the full models for each cancer type. This was conducted in both linear and nonlinear models in which the HLI was modelled as a continuous variable. Possible reverse causation was assessed by excluding cancer cases diagnosed within 2 years of baseline. To assess if underweight individuals captured by the healthiest BMI category biased associations for BMI to the null, those with BMI <18.5 were excluded. To assess whether the association with colorectal cancer was driven by differing associations of specific sites, associations were also analyzed separately for the incidence of colon and rectal cancer.

Multiple Imputation

Missing information among covariates was handled using multiple imputation (MI) using chained equations under the assumption that data were missing at random. All covariates required for analysis (lifestyle factors and potential confounders), cancer incidence, and the Nelson Aalen cumulative hazard estimator were included in the MI model. A fully conditional specification was applied, allowing the univariate imputation method and predictors set for each incomplete variable to be specified.[26] Missing information was replaced with values from 20 MI datasets with five iterations.

MI was performed on physical activity (1–10 scale); weight; height; smoking status; current smoking intensity (number of cigarettes smoked per day on average); time since smoking cessation (year); alcohol consumption (ethanol); daily grams of whole grains, fruit, vegetables, dairy, red meat and processed meat; years of education; and age at menarche (missing >19%). The remaining covariates were complete and therefore only used for prediction purposes. The HLI scores were then generated for each participant in all 20 MI datasets (96,869 participants × 20 datasets). HRs and 95% CIs were estimated by pooling estimates and standard errors from the 20 MI datasets using Rubin's Rule to account for between imputation variance.[27] All data treatment and statistical analyses were conducted in RStudio Version 1.3.959 with R Version 4.0.3.[28]

Results

After a median follow-up of 20.0 years and 15.2 postmenopausal years, there were 4286 postmenopausal breast, 1591 colorectal, 1416 lung, 1043 postmenopausal endometrial, 531 postmenopausal ovarian, 382 pancreatic, and 345 kidney cancer cases diagnosed. The majority (58%) of participants were in the 11–15 HLI group. The 0–5 HLI group was the least common (1%), while

the 6–10 (22%) and 16–20 HLI groups (19%) were evenly populated. The mean age was 51.6. Overall, 49% of participants were relatively physically active (physical activity level ≥6). Mean BMI was 24.7, mean alcohol consumption was 1.98 g/day and the mean diet score was 9 (Table 2).

Table 2 Baseline Characteristics of Participants by Healthy Lifestyle Index (HLI) Group, NOWAC Cohort, N = 81,554, Complete-Case Analyses

	All	HLI Group			
		0–5	6–10	11–15	16–20
Number of participants	81,554	878	17,847	47,435	15,394
Number of incident cancer cases					
Postmenopausal breast	3397	39	825	2014	519
Colorectal	1213	18	281	715	199
Lung	1006	33	349	557	67
Postmenopausal endometrial	807	13	228	444	122
Postmenopausal ovarian	425	3	101	241	73
Pancreatic	284	4	102	136	42
Kidney	268	3	74	156	35
Physical activity level (% >6)	48.9	1.8	16.3	50.9	83.3
BMI, mean (SD)	24.7 (3.9)	30.4 (4.2)	27.4 (4.5)	24.4 (3.4)	22.4 (2.2)
Smoking status, %					
Never	36.3	0.5	13.8	35.1	68.3
Former	34.9	14.4	34.7	38.1	26.6
Current	28.8	85.1	51.5	26.8	5.0
Alcohol consumption (g/day), median (IQ1, IQ3)	2.0 (0.4–5.3)	7.0 (2.0–12.2)	3.0 (1.0–7.9)	2.0 (0.6–5.3)	1.0 (0.0–2.9)
Diet score, median(IQ1, IQ3)	9 (7–11)	6 (6–8)	8 (6–9)	9 (8–10)	11 (9–12)
Age at baseline, mean(SD)	51.6 (6.4)	51.6 (5.6)	51.6 (6.2)	51.6 (6.4)	51.5 (6.7)
Height (cm), mean(SD)	166.3 (5.7)	166.2 (5.7)	166.2 (5.7)	166.2 (5.6)	166.4 (5.7)
Weight (kg), mean(SD)	68.4 (11.5)	84.1 (12.6)	75.6 (13.1)	67.4 (10.2)	61.9 (7.1)
Energy intake (KJ/day), mean(SD)	7076.9 (1900.3)	6602.5 (1853.6)	6747.9 (1806.3)	7096.3 (1866.1)	7425.8 (1982.3)
Education (years), mean(SD)	12.3 (3.4)	11.6 (3.1)	11.9 (3.3)	12.3 (3.4)	12.9 (3.6)
Age at menarche, mean(SD)	13.3 (1.4)	12.8 (1.4)	13.1 (1.4)	13.3 (1.4)	13.5 (1.4)
Oral contraception use (% ever)	54.3	59.5	58.4	54.3	48.9
Parity (%)					
Nulliparous	8.5	11.2	8.9	8.3	8.5
1–2	53.4	55.6	56.8	53.5	48.9
3+	38.2	33.2	34.4	38.3	42.5
Breastfeeding (%)					
0 months	54.1	58.0	54.8	53.7	54.6
0–12 months	24.5	29.9	27.6	24.8	19.6
>12 months	21.4	12.2	17.6	21.5	25.8
Hormone replacement therapy use (%)					
Never	66.7	58.1	62.5	66.5	72.3
Former	11.3	17.1	12.9	11.2	9.4
Current	22.0	24.8	24.6	22.2	18.3

Abbreviations: NOWAC, Norwegian women and cancer study; BMI, body mass index; SD, standard deviation; g/day, grams per day; IQ, interquartile; cig/day, cigarettes smoked per day; cm, centimetres; kg, kilograms; KJ, kilojoules.

Table 3 Linear Associations Between Healthy Lifestyle Index (HLI) Score and Incidence of Common Cancer Types, NOWAC Cohort (1996–2018), N = 96,869

	HLI score			
	0–5	6–10	11–15	16–20
	Cases (N)	Cases (N)	Cases (N)	Cases (N)
	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Postmenopausal	41	907	2162	567
Breast[a,b,c]	1.13 (0.85–1.50)	1.12 (1.03–1.20)	1.00 (ref)	0.83 (0.76–0.91)
Colorectal[a]	19	293	751	218
	1.46 (0.95–2.26)	1.04 (0.92–1.19)	1.00 (ref)	0.87 (0.75–1.00)
Lung[a]	36	376	613	72
	3.15 (2.29–4.34)	1.63 (1.45–1.84)	1.00 (ref)	0.39 (0.31–0.49)
Postmenopausal	13	248	473	128
Endometrial[a,b]	1.60 (0.95–2.69)	1.39 (1.20–1.62)	1.00 (ref)	0.78 (0.65–0.94)
Postmenopausal	3	101	241	73
Ovarian[a,b]	1.11 (0.46–2.67)	1.06 (0.86–1.33)	1.00 (ref)	0.94 (0.73–1.20)
Pancreatic[a]	4	108	143	45
	1.49 (0.57–3.91)	1.87 (1.48–2.37)	1.00 (ref)	0.94 (0.69–1.28)
Kidney[a]	3	79	163	37
	1.26 (0.47–3.36)	1.28 (0.99–1.67)	1.00 (ref)	0.75 (0.53–1.04)

Notes: [a]Results from analyses conducted on multiple imputation data, adjusted for education and height. [b]Additionally adjusted for age at menarche, use of oral contraceptives, parity, breastfeeding, and use of hormone replacement therapy. [c]Additionally adjusted for family history of breast cancer in a first-degree relative.
Abbreviations: NOWAC, Norwegian women and cancer study; N, number; HR, hazard ratio; CI, confidence interval.

Results from the MI models (Table 3) showed that the magnitude and direction of the effects were similar to those observed in complete-case analyses [see Table S1 in Supplementary File]. After adjustment for covariates, estimates from the linear analysis using MI data (Figure 1) showed HRs of 0.97 (95% CI: 0.96–0.98) for postmenopausal breast cancer, 0.98 (95% CI: 0.96–1.00) for colorectal cancer, 0.86 (95% CI: 0.84–0.87) for lung cancer, 0.93 (95% CI: 0.91–0.95) for postmenopausal endometrial cancer, 0.99 (95% CI: 0.96–1.02) for postmenopausal ovarian cancer, 0.92 (95% CI: 0.89–0.95) for pancreatic cancer, and 0.94 (95% CI: 0.91–0.97) for kidney cancer for every 1-point increase in HLI score. No considerable difference in HRs was observed for colon and rectal cancers (results not shown). When the HLI score was modelled as a categorical variable, HRs for the 16–20 HLI group compared to the 11–15 HLI group were 0.83 (95% CI: 0.76–0.91) for postmenopausal breast cancer, 0.88 (95% CI: 0.77–1.00) for colorectal cancer, 0.39 (95% CI: 0.31–0.49) for lung cancer, 0.78 (95% CI: 0.64–0.94) for postmenopausal endometrial cancer, 0.94 (95% CI: 0.73–1.20) for postmenopausal ovarian cancer, 0.94 (95% CI: 0.69–1.28) for pancreatic cancer, and 0.72 (95% CI: 0.52–1.00) for kidney cancer (Table 3).

Analyses of nonlinearity in associations demonstrated that AIC estimates were lowest for all outcomes, except for lung cancer incidence, when three knots at the percentiles were applied, positioned at the defined HLI score boundaries (1, 20) and at the HLI score median (13), compared to four- and five-knot models positioned at percentiles or three-, four-, and five-knot models positioned at equal intervals [see Table S2 and Figure S2–8 for plot comparisons in Supplementary File]. The AIC estimate was lowest when five knots positioned at the percentiles were applied to the lung cancer incidence model; however, there were visual indications of overfitting [see Figure S4A in Supplementary File]. Therefore, three knots were applied for all outcomes. The resulting

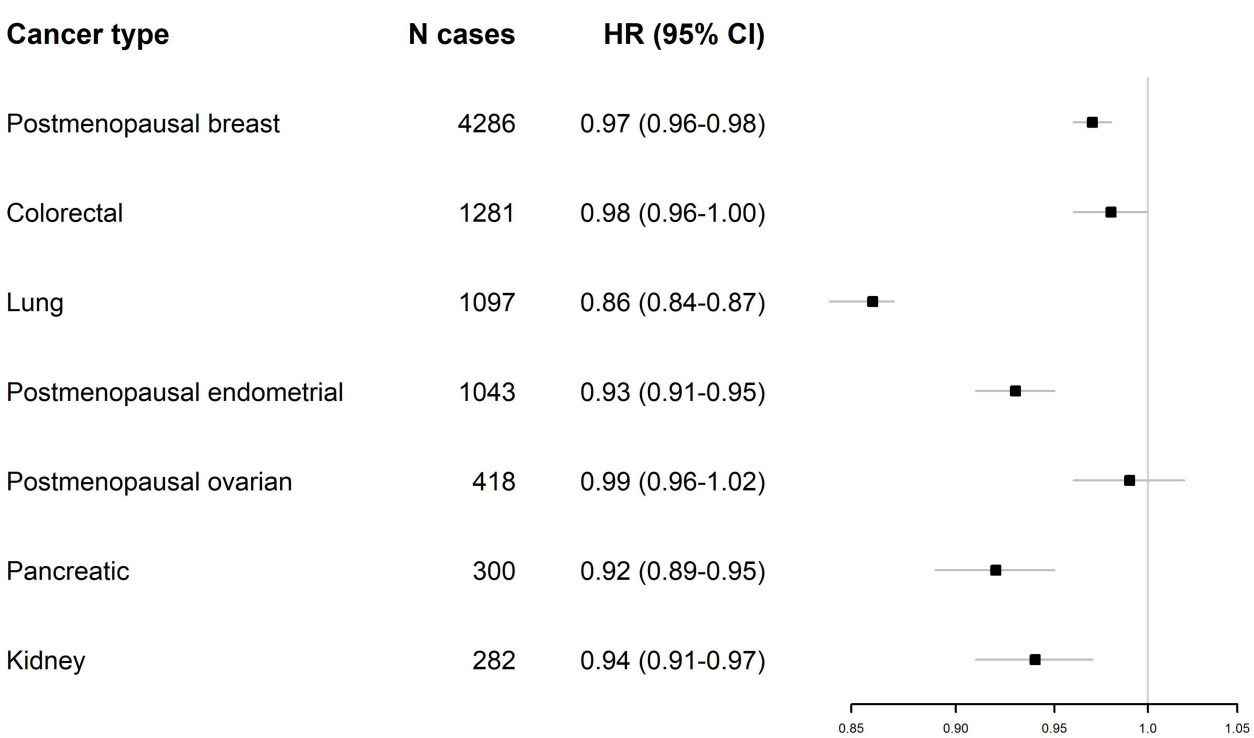


Figure 1 Forest plot of linear associations between healthy lifestyle index (HLI) score and incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, pancreatic and kidney cancers, NOWAC (1996–2018), N = 96,869. HRs and 95% CIs correspond to a 1-point increase on the HLI score. Estimates were obtained from multiple imputation data, employed Cox proportional hazard regression, adjusted for education and height.
Abbreviations: NOWAC, Norwegian women and cancer study; HR, hazard ratio; CI, confidence interval.

plots indicated nonlinearity for lung cancer incidence, with relatively linear decreases in incidence until reaching a plateau at an HLI score of approximately 16 (Figure 2). Nonlinearity was also indicated for postmenopausal breast cancer incidence, with a strengthening of the negative HR gradient at HLI scores above 13 compared to below. The plots for the incidence of colorectal, postmenopausal endometrial, pancreatic, and kidney cancer did not display indications of nonlinearity.

Exclusion of single lifestyle factors from HLI scores did not affect estimates of the incidence of postmenopausal breast, colorectal, and postmenopausal ovarian cancer [see Table S3 in Supplementary File]. For lung cancer incidence, the HR increased to 1.02 (95% CI: 0.99–1.03) when smoking status was excluded, and the HR decreased to 0.78 (95% CI: 0.76–0.79) when BMI was excluded. For postmenopausal endometrial cancer incidence, the HR decreased to 0.89 (95% CI: 0.87–0.91) when smoking status was excluded, and the HR increased to 0.97 (95% CI: 0.95–0.99) when BMI was excluded. For pancreatic cancer incidence, the HR increased to 0.97 (95% CI: 0.93–1.01) when smoking status was excluded. For kidney cancer incidence, the HR increased to 0.96 (95% CI: 0.92–1.00) when BMI was excluded. Among nonlinear models [see Figure S9–15 in Supplementary File], exclusion of BMI and diet resulted in linear associations for postmenopausal breast cancer incidence. Exclusion of smoking status in the nonlinear lung cancer incidence model resulted in no association. Removal of cancers diagnosed within 2 years of baseline and excluding those with BMI <18.5 did not alter the results considerably (results not shown).

Discussion

In this Norwegian national prospective cohort study, we identified inverse associations between our a priori-defined HLI score and the incidence of all included cancer types, except for postmenopausal ovarian cancer. Our examinations indicated that higher HLI scores were associated with lower lung cancer incidence, whereby there were smaller differences in lung cancer incidences among the healthiest participants. There were indications that differences in postmenopausal breast cancer incidence could be greater for women with HLI scores above the median (13). We consider our study population to have a high adherence to

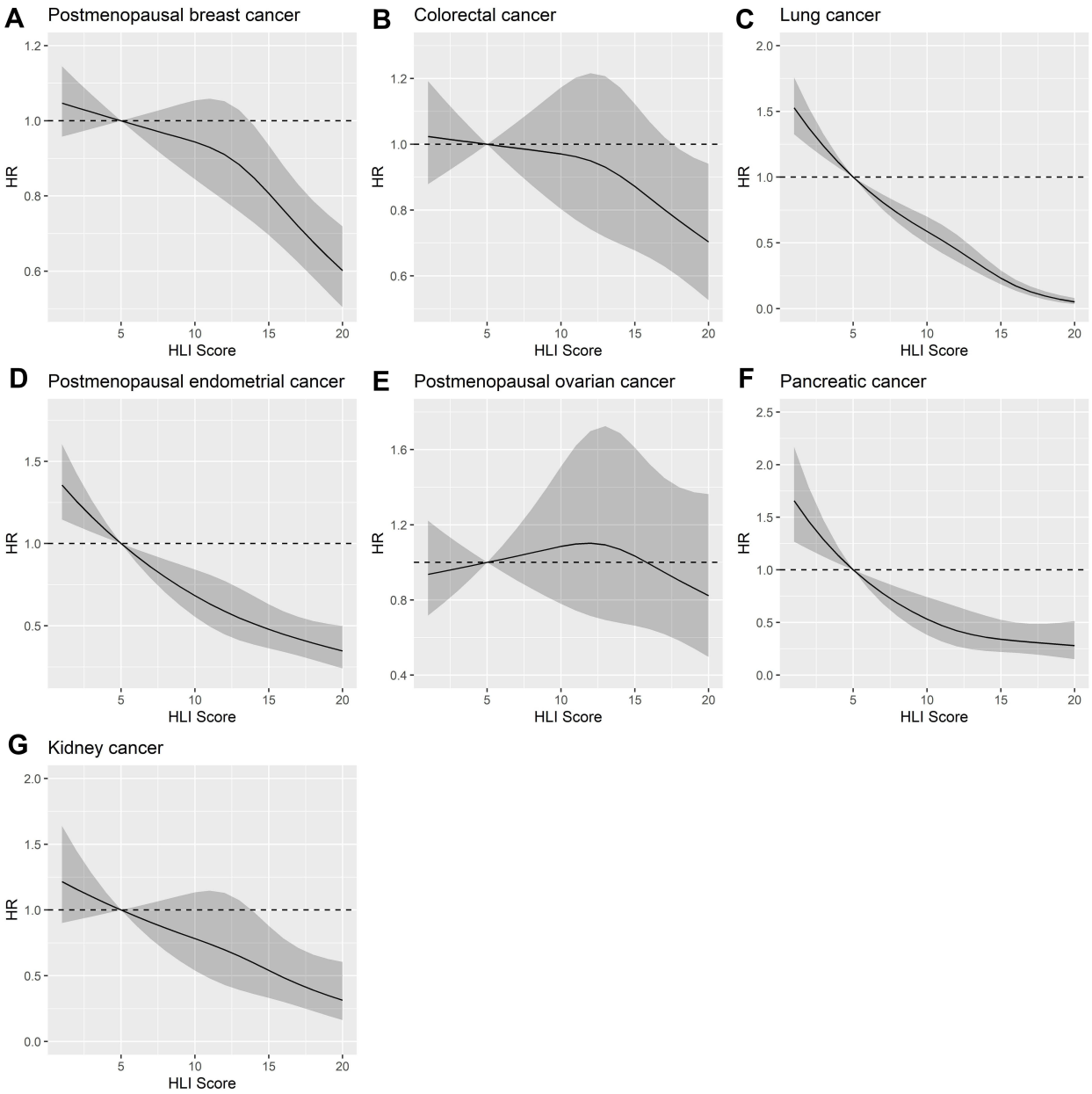


Figure 2 Nonlinear associations between the healthy lifestyle index (HLI) score and incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, pancreatic and kidney cancers, NOWAC (1996–2018), N = 96,869. Obtained by applying restricted cubic splines with three knots to the healthy lifestyle index (HLI) score from complete-case analysis data. All models (**A-G)** were adjusted for education and height. Models (**A, D**, and **E**) were additionally adjusted for age at menopause, use of oral contraceptives, parity, breastfeeding, and use of hormone replacement therapy. Model (**A**) was additionally adjusted for family history of breast cancer in a first-degree relative.

Abbreviations: NOWAC, Norwegian women and cancer study; HLI, healthy lifestyle index; HR, hazard ratio.

healthy lifestyles, with 77% of participants having HLI scores above 10.

Postmenopausal Breast Cancer

The incidence of postmenopausal breast cancer decreased by 3% for every 1-point increase in HLI score, and our results suggest a nonlinear relationship, based on a comparison of AIC values. Visually, we observed that the inverse association was more pronounced at HLI scores above 13. Our linear estimate is consistent with the results of two other studies, which reported a 3% and 4% decrease, respectively, in breast cancer incidence for

each 1-point increase in HLI score.[29,30] Other studies on the association between a combined lifestyle measure and breast cancer incidence observed a lower incidence among those with healthier lifestyles.[31–36] However, these studies are less comparable to ours due to the scoring system employed.

We also conducted multiple sensitivity analyses; each one excluded a different, single lifestyle factor from the HLI score. These analyses resulted in inverse associations that were similar to those observed in the main analyses, suggesting that no specific lifestyle factor drove the observed associations to the HLI score. BMI and diet were suggested as the main contributors to the nonlinear trend in the association, as a linear trend was observed when these factors were excluded. Previous publications from the NOWAC study have observed positive associations between BMI[21] and smoking[37] with cancer incidence; however, there is no evidence from the NOWAC study that physical activity[38] or food groups and dietary patterns[39] are associated with postmenopausal breast cancer incidence. Lifestyle factors known to be associated with postmenopausal breast cancer in the NOWAC cohort, namely, BMI and smoking status, did not fully explain our observed association, suggesting that the additive effect of multiple healthy lifestyle factors is important for postmenopausal breast cancer prevention, even though some single lifestyle factors are only weakly associated with cancer incidence.

Colorectal Cancer

We observed an inverse association between HLI score and colorectal cancer incidence, suggesting a 2% decrease in incidence for each 1-point increase in HLI score, and a 13% decrease in incidence when the 16–20 HLI group was compared to the 11–15 HLI group. Allowing for nonlinearity in the HLI indicated that the score was more strongly associated with cancer incidence among women with the highest HLI scores at baseline, similar to our observation for postmenopausal breast cancer risk. However, the large amount of uncertainty in the present study makes it difficult to establish a clear interpretation of this trend, and there are currently no other studies exploring nonlinearity with which we could compare our results. Previous prospective studies that assessed the overall lifestyle with a similar additive exposure score observed comparable effect estimates in linear models, but only with greater precision.[40–42]

Sensitivity analyses indicated that single lifestyle factors did not explain the inverse association we observed. In previous publications of the NOWAC study, clear associations were not observed between single lifestyle factors and colorectal cancer incidence.[21,43–45] However, our results suggest that healthy lifestyle factors, in sum, are inversely associated, and thus, together, could reduce colorectal cancer incidence.

Lung Cancer

We observed a strong, inverse association with lung cancer incidence, with a 14% decrease in incidence for each 1-unit increase in HLI score. This association plateaued at the upper end of HLI scores. To our knowledge, no other study has explored nonlinearity in the relationship between overall lifestyle and lung cancer incidence. Compared to our result, an inverse association of lesser magnitude was observed in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort, with a 9% decrease in the incidence of tobacco-related cancers for each 1-unit increase in a similarly constructed 20-unit index.[46] This difference in the observed associations could be explained by the fact that several cancer types were pooled in the EPIC analyses. In accordance with this study, a strong inverse association was also observed in a Chinese population;[47] however, a 5-point HLI and categorical analysis were employed, and thus the strength of associations could not be compared.

Results from sensitivity analyses suggested that smoking status fully explained the observed inverse association. We also observed that smoking status explained the plateau in the nonlinear model. This is consistent with a consensus that smoking is the greatest contributor to lung cancer incidence[48,49] in the NOWAC cohort and, by design, it is impossible for those with HLI scores above 17 to be current smokers. A study employing American Cancer Society (ACS) recommendations, not including smoking, to construct a combined score also observed a null association with lung cancer incidence.[36] However, when employing a combined exposure score based on WCRF/AICR recommendations in the EPIC cohort, researchers observed an inverse association between combined lifestyle factors and lung cancer incidence without the influence of smoking. Since dietary factors are weighted heavily in the WCRF/AICR adherence score, carcinogenic effects from a high consumption of preserved and red meat may have been detected.[50] We observed that a higher BMI was inversely associated with lung cancer

incidence, which is consistent with the "obesity paradox".[51] While reverse causation due to weight loss associated with the early stages of cancer was ruled out in our study, it is unclear whether the residual confounding effect of smoking cessation could explain the paradoxical observation.

Postmenopausal Endometrial Cancer

The incidence of postmenopausal endometrial cancer decreased by 7% for each 1-point increase in HLI score, with no indication of a departure from linearity. Our linear observations are highly consistent with previous studies. A 6% and 5% risk decrease per 1-unit increase in HLI, based on 20-unit indices similar to ours, was observed in cohorts from the United States[52] and Canada.[30] General adherence to WCRF/AICR and ACS lifestyle recommendations was also inversely associated with endometrial cancer risk in prospective cohort studies.[36,53]

From sensitivity analyses, we observed that the HLI was still inversely associated with postmenopausal endometrial cancer incidence when BMI was excluded from the HLI score, although to a lesser degree than in the main analyses. This suggests that BMI contributed considerably to the association, which is consistent with the known dose–response association between overweight/obesity and cancer risk in the NOWAC study[21] and with the current consensus that higher BMI is a risk factor for postmenopausal endometrial cancer.[54] Excluding smoking status from the HLI score strengthened the association, indicating that smoking was protective. While smoking confers risk for most non-communicable diseases, including cancers, cardiovascular disease, and diabetes, its relationship with endometrial cancer risk appears to divert from this pattern.[55] However, our results suggest that higher physical activity, lower alcohol consumption, healthier diet, and especially lower BMI are, in sum, protective against postmenopausal endometrial cancer incidence. In line with this, higher physical activity levels[20] and lower BMI[21] have been associated with decreased postmenopausal endometrial cancer incidence in the NOWAC cohort. Smoking habits, alcohol consumption, and aspects of diet have not been previously investigated with respect to endometrial cancer in the NOWAC study.

Postmenopausal Ovarian Cancer

We observed a null association between the HLI score and postmenopausal ovarian cancer incidence. There was no indication of a nonlinear trend in the association. Null associations between comparable HLI scores were also observed in cohorts from Canada,[30] the United States[52] and France.[34] Cohort studies employing scores based on WCRF/AICR and ACS recommendations also observed null associations.[36,53] As such, there is little evidence in the published literature that overall lifestyle is associated with ovarian cancer incidence, compared to the stronger associations observed for the other cancers we explored.

Sensitivity analyses indicated that no single lifestyle factor suppressed the association. In the NOWAC cohort, ever smoking was not differentially associated with ovarian cancer incidence across histological subtypes and invasiveness,[56] nor was BMI[21] or physical activity[57] observed to be associated with overall ovarian cancer. This lends further evidence to a minimal or absence of association between lifestyle factors and ovarian cancer incidence in the NOWAC cohort. However, heterogeneity in the etiology of ovarian cancer subtypes may have attenuated the magnitude and reduced the precision of estimates. This should be investigated further in the NOWAC cohort and other populations.

Pancreatic Cancer

Pancreatic cancer incidence decreased by 8% for each 1-point increase in HLI score, with no indication of a departure from linearity in the association. To our knowledge, two studies on overall lifestyle and pancreatic cancer have been conducted. Naudin et al[58] observed a similar association in the EPIC cohort when employing a 20-unit HLI comparable to ours: risk decreased by 21% for each 3-unit increase in the HLI score, corresponding to a decrease of approximately 7% per 1-unit increase. In an American cohort, Jiao et al[59] also observed an inverse association: compared to the lowest HLI score – 0, the highest score – 5 – was associated with a 58% decrease in pancreatic cancer incidence.

Sensitivity analyses indicated that smoking status fully explained the inverse association we observed between the HLI score and pancreatic cancer incidence. However, Naudin et al[58] observed that healthier lifestyles, in addition to smoking habits, were associated with decreased pancreatic cancer incidence in the EPIC cohort. Indeed, obesity and alcohol consumption are also known risk factors for pancreatic cancer, although

smoking is recognized as the primary modifiable risk factor.[60] It is possible that we did not observe these associations due to a lack of power or our choice of body fatness measure, since weight gain, not baseline BMI, has been strongly associated with pancreatic cancer risk in the NOWAC cohort.[21]

Kidney Cancer

Kidney cancer incidence decreased by 6% for each 1-point increase in HLI score, with no indication of departure from linearity in the association. To our knowledge, there are no published studies that have examined the association between a similar combined score to ours and kidney cancer incidence to which we can compare our observations. While two other studies based on WCRF and ACS recommendation adherence scores also observed inverse associations, their combined score did not include smoking.[36,53]

Our sensitivity analyses demonstrated that BMI was a strong contributor, in comparison to the other lifestyle factors included the HLI, to the inverse association observed in the main analyses. Indeed, higher BMI has been associated with kidney cancer incidence in the NOWAC study.[21] Although smoking is a well-established risk factor for kidney cancer,[61] there was no evidence that it was an especially strong contributor to the association in our study, nor has it been studied as a single risk factor for kidney cancer in the NOWAC study previously. The assessment of other single lifestyle factors in relation to kidney cancer has not been undertaken in the NOWAC study, and these factors, such as physical activity and alcohol consumption, are emerging as important protective or risky behaviors.[61] Nevertheless, there is evidence from the present study to suggest that the sum of considered lifestyle factors, not just BMI, are important for kidney cancer prevention.

Public Health Implications and Interpreting a Combined Exposure

The associations between single lifestyle factors and the incidence of different cancer types have been thoroughly examined in previous publications from the NOWAC study. In the present report, we provide an alternate approach for assessing risk, by combining several relevant lifestyle factors into an additive exposure score, as well as exploring linear and nonlinear associations. Given the representativeness of the NOWAC cohort to the Norwegian population,[62] we can reasonably suggest that the majority of Norwegian women have high HLI scores, with 77% of the NOWAC cohort scoring above 10. Although these women can be considered healthy, our observations suggest that, regardless of baseline HLI score, healthier overall lifestyle, and thus greater adherence to public health recommendations, is protective against postmenopausal breast colorectal, lung, postmenopausal endometrial, pancreatic and kidney cancer. While our findings indicated no added benefit of healthy behaviors, besides smoking reduction for lung and pancreatic cancer, these combined behaviors had a meaningful protective impact on the incidence of postmenopausal breast, postmenopausal endometrial, postmenopausal ovarian, and kidney cancers. Further, limitations in our data may have rendered the remaining associations, without the contribution of smoking, undetectable.

The use of an additive exposure score provides a more holistic assessment of modifiable risk factors compared to single risk factors, as lifestyle behaviors co-exist. We therefore constructed an a priori index based on tangible increments of lifestyle behaviors, as far as our data would allow. With the aim of simplicity, we chose to apply a single, additive index across several cancer types, as this approach offered ease of interpretation, reproducibility across cancer types and populations,[63] and effective public health messaging. Our results demonstrate that the HLI score is a valuable representation of combined lifestyle factors when evaluating the effects of several modifiable lifestyle factors in a population-based study. For example, the HLI score made it possible to determine whether an overall healthy lifestyle was protective of colorectal cancer, when individual lifestyle factors did not show such associations in the NOWAC cohort.

However, the use of our HLI has its challenges, given the inevitable loss of information that occurs when combining factors into an additive score. We were unable to discern which lifestyle factors specifically contributed to the risk difference at a specific HLI score, nor whether a reduction in one risky behavior could offset the increase in another.[7] The additive score also assumes linearity in its unit increments, whereas equal distance between units on the HLI may not represent proportional increments of a behavior.

Strengths

The main strengths of this study are its prospective design, large sample size, long follow-up time, and linkage to

national registries. An assessment of the external validity of the NOWAC study concluded that the NOWAC cohort is adequately representative of Norwegian women.[62] Recruitment through random sampling within the Central Population Register of Norway minimized sampling bias. Although education levels in the NOWAC cohort are somewhat higher than the national average, there were no considerable differences in cancer incidence or lifestyle factors compared to national reports.[62] As such, we can assume that the distribution of participants across HLI scores in the NOWAC study represents the distribution of Norwegian women. Linkage to registries allowed us to be highly confident in the ascertainment of all incident and prevalent cancer cases. Further, MI was used to avoid potential bias created by listwise deletion and to conserve sample size.[26]

Limitations

This study also has several limitations. Given that lifestyle information is self-reported, it is possible that exposure misclassification introduces bias in our estimates of association. It has been widely acknowledged that research participants tend to underreport their food intake, alcohol consumption, and weight, and overreport variables like height.[64] Although these tendencies were confirmed in a NOWAC study that compared energy and alcohol consumption in the FFQ to repeated 24-hour dietary recalls, the FFQ still performed well on ranking high and low consumers.[17] As such, the dietary component of the HLI, which scored participants based on relative intake of food groups, is expected to have achieved an adequate ranking.

Underreporting of weight was also confirmed in a NOWAC validation study, in which the largest tendency to underreport occurred among overweight women, and the largest degree of underreporting occurred among obese women.[15] Our risk estimates may thus be attenuated due to misclassification of BMI. It is also possible that waist circumference or waist–hip ratio are more accurate indicators of body fatness, and thus metabolic risk, than BMI,[65] suggesting that stronger associations may have been observed if waist circumference or waist–hip ratio were used. The physical activity report measure was not informed directly by questions pertaining to dose (intensity, frequency, and duration of physical activity). Therefore, differences in physical activity may not have been fully captured, resulting in attenuated estimates. Nevertheless, the physical activity measure demonstrated an adequate Spearman's rank correlation coefficient in the range of 0.36–0.46 with objective criteria in the validation study and is considered relevant to rank the physical activity levels at a population level.[14]

We did not have information on family history of colorectal, endometrial, pancreatic, and kidney cancers. Due to this and other unmeasured variables, residual confounding may be introduced. Lastly, the models used in this study assumed that lifestyle exposures at baseline were held constant until the participants were censored. Potential changes in lifestyle may be relevant for assessing the relationship between combined lifestyle factors and cancer incidence.

Conclusion

This is the first prospective study to examine the linear and nonlinear relationship between combined lifestyle factors and the incidence of common cancers in a Norwegian population. Based on our results, healthier lifestyle, as assessed by the HLI score, was associated with lower incidence of postmenopausal breast, colorectal, lung, postmenopausal endometrial, pancreatic, and kidney cancer among women, although the magnitude and linearity varied. The adoption of an overall healthy lifestyle should therefore have a considerable impact on decreasing the cancer burden among women. Limiting smoking is the single most important component of the overall lifestyle for minimizing lung and pancreatic cancer incidence.

Ethics and Consent

The NOWAC study has been given appropriate approval for collection and handling of questionnaire data by the Regional Committee for Medical and Health Research Ethics (REK) (reference: REK NORD 141/2008) and the Norwegian Data Inspectorate. All participants provided written voluntary and informed consent. The guidelines according to the Declaration of Helsinki were followed.

Acknowledgments

We would like to acknowledge all the NOWAC participants for their valuable contribution to this study. In addition, thanks are extended to Guri Skeie, Torill E Jensen, and Sabine Naudin for their discussions and input.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

Disclosure

The authors report no conflicts of interest in this work.

References

1. GBD. Causes of death collaborators 2017. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392 (10159):1736–1788.
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249.
3. Cancer Registry of Norway. *Cancer in Norway 2018: Cancer Incidence, Mortality, Survival and Prevalence in Norway*. Oslo: Cancer Registry of Norway; 2018.
4. World Health Organization (WHO). Cancer. Available from: https://www.who.int/news-room/fact-sheets/detail/cancer. Accessed October 9, 2020.
5. Weiderpass E. Lifestyle and Cancer Risk. *J Prev Med Public Health*. 2010;43(6):459–471. doi:10.3961/jpmph.2010.43.6.459
6. World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR). *Diet, Nutrition, Physical Activity and Cancer: A Global Perspective*: WCRF/AICR; 2018.
7. Friedenreich CM, McTiernan A. Combining variables for cancer risk estimation: is the sum better than the parts? *Cancer Prevent Res*. 2018;11(6):313–316. doi:10.1158/1940-6207.CAPR-18-0102
8. Pronk NP, Anderson LH, Crain AL, et al. Meeting recommendations for multiple healthy lifestyle factors. Prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. *Am J Prev Med*. 2004;27(2 Suppl):25–33. doi:10.1016/j.amepre.2004.04.022
9. Anand P, Kunnumakkara AB, Sundaram C, et al. Cancer is a preventable disease that requires major lifestyle changes. *Pharm Res*. 2008;25(9):2097–2116. doi:10.1007/s11095-008-9661-9
10. Zhang Y-B, Pan X-F, Chen J, et al. Combined lifestyle factors, incident cancer, and cancer mortality: a systematic review and meta-analysis of prospective cohort studies. *Br J Cancer*. 2020;122 (7):1085–1093. doi:10.1038/s41416-020-0741-x
11. Naudin S, Solans Margalef M, Saberi Hosnijeh F, et al. Healthy lifestyle and the risk of lymphoma in the European prospective investigation into cancer and nutrition study. *Int j Cancer*. 2020;147 (6):1649–1656. doi:10.1002/ijc.32977
12. Lund E, Dumeaux V, Braaten T, et al. Cohort Profile: the Norwegian Women and Cancer Study—NOWAC—Kvinner og kreft. *Int J Epidemiol*. 2007;37(1):36–41. doi:10.1093/ije/dym137
13. Landsforeningen for kosthold og helse. *Mål Og Vekt for Matvarer (Norwegian Weight and Measurement Table)*. Oslo; 1995.
14. Borch KB, Ekelund U, Brage S, Lund E. Criterion validity of a 10-category scale for ranking physical activity in Norwegian women. *Int J Behav Nutr Phys Act*. 2012;9:2. doi:10.1186/1479-5868-9-2
15. Skeie G, Mode N, Henningsen M, Borch KB. Validity of self-reported body mass index among middle-aged participants in the Norwegian Women and Cancer study. *Clin Epidemiol*. 2015;7:313–323. doi:10.2147/CLEP.S83839
16. Parr CL, Veierod MB, Laake P, Lund E, Hjartaker A. Test-retest reproducibility of a food frequency questionnaire (FFQ) and estimated effects on disease risk in the Norwegian Women and Cancer Study (NOWAC). *Nutr J*. 2006;5:4. doi:10.1186/1475-2891-5-4
17. Hjartåker A, Andersen LF, Lund E. Comparison of diet measures from a food-frequency questionnaire with measures from repeated 24-hour dietary recalls. The Norwegian Women and Cancer Study. *Public Health Nutr*. 2007;10(10):1094–1103. doi:10.1017/S1368980007702872
18. World Health Organization (WHO). *International Classification of Diseases for Oncology:. Third* ed. Geneva: Switzerland; 2013.
19. Korn EL, Graubard BI, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am J Epidemiol*. 1997;145(1):72–80. doi:10.1093/oxfordjournals.aje.a009034
20. Borch KB, Weiderpass E, Braaten T, Jareid M, Gavrilyuk OA, Licaj I. Physical activity and risk of endometrial cancer in the Norwegian Women and Cancer (NOWAC) study. *Int j Cancer*. 2017;140(8):1809–1818. doi:10.1002/ijc.30610
21. da Silva M, Weiderpass E, Licaj I, Lissner L, Rylander C. Excess body weight, weight gain and obesity-related cancer risk in women in Norway: the Norwegian Women and Cancer study. *Br J Cancer*. 2018;119(5):646–656. doi:10.1038/s41416-018-0240-5
22. Beral V, Banks E, Reeves G, Bull D. Breast cancer and hormone-replacement therapy in the Million Women Study. *Lancet*. 2003;362(9382):419–427. doi:10.1016/S0140-6736(03)14596-5
23. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982;69(1):239–241. doi:10.1093/biomet/69.1.239
24. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials*. 1996;17(4):343–346. doi:10.1016/0197-2456(96)00075-X
25. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol*. 2019;34(3):211–219. doi:10.1007/s10654-019-00494-6
26. van Buuren S. *Flexible Imputation of Missing Data*. 2nd ed. Boca Raton, Florida: Taylor & Francis Group; 2018.
27. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Hoboken: WILEY; 2009.
28. R Foundation for Statistical Computing. *R: A Language and Environment for Statistical Computing [Computer Program]*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
29. Arthur R, Wassertheil-Smoller S, Manson JE, et al. The combined association of modifiable risk factors with breast cancer risk in the women's health initiative. *Cancer Prevent Res*. 2018;11(6):317–326. doi:10.1158/1940-6207.CAPR-17-0347
30. Arthur R, Kirsh VA, Kreiger N, Rohan T. A healthy lifestyle index and its association with risk of breast, endometrial, and ovarian cancer among Canadian women. *Cancer Causes Control*. 2018;29 (6):485–493. doi:10.1007/s10552-018-1032-1
31. McKenzie F, Ferrari P, Freisling H, et al. Healthy lifestyle and risk of breast cancer among postmenopausal women in the European prospective investigation into cancer and nutrition cohort study. *Int j Cancer*. 2015;136(11):2640–2648. doi:10.1002/ijc.29315
32. McKenzie F, Ellison-Loschmann L, Jeffreys M, Firestone R, Pearce N, Romieu I. Healthy lifestyle and risk of breast cancer for indigenous and non-indigenous women in New Zealand: a case control study. *BMC Cancer*. 2014;14:12. doi:10.1186/1471-2407-14-12
33. Sánchez-Zamorano LM, Flores-Luna L, Angeles-Llerenas A, et al. Healthy lifestyle on the risk of breast cancer. *Cancer Epidemiol Biomarkers Prevent*. 2011;20(5):912–922. doi:10.1158/1055-9965.EPI-10-1036
34. Dartois L, Fagherazzi G, Boutron-Ruault MC, Mesrine S, Clavel-Chapelon F. Association between five lifestyle habits and cancer risk: results from the E3N cohort. *Cancer Prevent Res*. 2014;7 (5):516–525. doi:10.1158/1940-6207.CAPR-13-0325
35. Romaguera D, Gracia-Lavedan E, Molinuevo A, et al. Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case-control study. *Int j Cancer*. 2017;141(1):83–93. doi:10.1002/ijc.30722
36. Kabat GC, Matthews CE, Kamensky V, Hollenbeck AR, Rohan TE. Adherence to cancer prevention guidelines and cancer incidence, cancer mortality, and total mortality: a prospective cohort study. *Am J Clin Nutr*. 2015;101(3):558–569. doi:10.3945/ajcn.114.094854

37. Gram IT, Little MA, Lund E, Braaten T. The fraction of breast cancer attributable to smoking: the Norwegian women and cancer study 1991–2012. *Br J Cancer*. 2016;115(5):616–623. doi:10.1038/bjc.2016.154
38. Borch KB, Lund E, Braaten T, Weiderpass E. Physical activity and the risk of postmenopausal breast cancer - the Norwegian Women and Cancer Study. *J Negat Results Biomed*. 2014;13:3. doi:10.1186/1477-5751-13-3
39. Engeset D, Dyachenko A, Ciampi A, Lund E. Dietary patterns and risk of cancer of various sites in the Norwegian European Prospective Investigation into Cancer and Nutrition cohort: the Norwegian Women and Cancer study. *Eur j Cancer Prevent*. 2009;18(1):69–75. doi:10.1097/CEJ.0b013e328305a091
40. Aleksandrova K, Pischon T, Jenab M, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. *BMC Med*. 2014;12:168. doi:10.1186/s12916-014-0168-4
41. Kirkegaard H, Johnsen NF, Christensen J, Frederiksen K, Overvad K, Tjønneland A. Association of adherence to lifestyle recommendations and risk of colorectal cancer: a prospective Danish cohort study. *BMJ*. 2010;341:c5504. doi:10.1136/bmj.c5504
42. Wang K, Ma W, Wu K, et al. Healthy lifestyle, endoscopic screening, and colorectal cancer incidence and mortality in the United States: a nationwide cohort study. *PLoS Med*. 2021;18(2):e1003522. doi:10.1371/journal.pmed.1003522
43. Gram IT, Braaten T, Lund E, Le Marchand L, Weiderpass E. Cigarette smoking and risk of colorectal cancer among Norwegian women. *Cancer Causes Control*. 2009;20(6):895–903. doi:10.1007/s10552-009-9327-x
44. Oyeyemi SO, Braaten T, Botteri E, Berstad P, Borch KB. Exploring geographical differences in the incidence of colorectal cancer in the Norwegian Women and Cancer Study: a population-based prospective study. *Clin Epidemiol*. 2019;11:669–682. doi:10.2147/CLEP.S207413
45. Oyeyemi SO, Braaten T, Licaj I, Lund E, Benjaminsen borch K. Physical activity patterns and the risk of colorectal cancer in the Norwegian women and cancer study: a population-based prospective study. *BMC Cancer*. 2018;18(1):1216. doi:10.1186/s12885-018-5092-0
46. McKenzie F, Biessy C, Ferrari P, et al. Healthy lifestyle and risk of cancer in the European prospective investigation into cancer and nutrition cohort study. *Medicine*. 2016;95(16):e2850. doi:10.1097/MD.0000000000002850
47. He Y, Bai Y, Wei S, et al. Healthy lifestyle and cancer risk among Chinese population in the Dongfeng-Tongji cohort. *Ann Med*. 2020;52(7):393–402. doi:10.1080/07853890.2020.1798017
48. Hansen MS, Licaj I, Braaten T, Lund E, Gram IT. The fraction of lung cancer attributable to smoking in the Norwegian women and cancer (NOWAC) Study. *Br J Cancer*. 2020;2:548.
49. Malhotra J, Malvezzi M, Negri E, La Vecchia C, Boffetta P. Risk factors for lung cancer worldwide. *Eur Respir J*. 2016;48(3):889. doi:10.1183/13993003.00359-2016
50. Xue XJ, Gao Q, Qiao JH, Zhang J, Xu CP, Liu J. Red and processed meat consumption and the risk of lung cancer: a dose-response meta-analysis of 33 published studies. *Int J Clin Exp Med*. 2014;7(6):1542–1553.
51. Mavridis K, Michaelidou K. The obesity paradox in lung cancer: is there a missing biological link? *J Thorac Dis*. 2019;11(Suppl 3):S363–s366. doi:10.21037/jtd.2018.12.69
52. Arthur R, Wassertheil-Smoller S, Rohan TE, et al. Associations of a healthy lifestyle index with the risks of endometrial and ovarian cancer among women in the Women's Health Initiative Study. *Am J Epidemiol*. 2019;188(2):261–273. doi:10.1093/aje/kwy249
53. Romaguera D, Vergnaud AC, Peeters PH, et al. Is concordance with World cancer research fund/American Institute for cancer research guidelines for cancer prevention related to subsequent risk of cancer? Results from the EPIC study. *Am J Clin Nutr*. 2012;96(1):150–163. doi:10.3945/ajcn.111.031674
54. Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*. 2007;335(7630):1134. doi:10.1136/bmj.39367.495995.AE
55. Zhou B, Yang L, Sun Q, et al. Cigarette smoking and the risk of endometrial cancer: a meta-analysis. *Am J Med*. 2008;121(6):501–508.e503. doi:10.1016/j.amjmed.2008.01.044
56. Licaj I, Jacobsen BK, Selmer RM. Smoking and risk of ovarian cancer by histological subtypes: an analysis among 300 000 Norwegian women. *Br J Cancer*. 2017;116(2):270–276. doi:10.1038/bjc.2016.418
57. Weiderpass E, Margolis KL, Sandin S, et al. Prospective study of physical activity in different periods of life and the risk of ovarian cancer. *Int j Cancer*. 2006;118(12):3153–3160. doi:10.1002/ijc.21800
58. Naudin S, Viallon V, Hashim D, et al. Healthy lifestyle and the risk of pancreatic cancer in the EPIC study. *Eur J Epidemiol*. 2020;35(10):975–986. doi:10.1007/s10654-019-00559-6
59. Jiao L, Mitrou PN, Reedy J, et al. A combined healthy lifestyle score and risk of pancreatic cancer in a large cohort study. *Arch Intern Med*. 2009;169(8):764–770. doi:10.1001/archinternmed.2009.46
60. Lowenfels AB, Maisonneuve P. Epidemiology and risk factors for pancreatic cancer. *Best Pract Res Clin Gastroenterol*. 2006;20(2):197–209. doi:10.1016/j.bpg.2005.10.001
61. Chow W-H, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. *Nat Rev Urol*. 2010;7(5):245–257. doi:10.1038/nrurol.2010.46
62. Lund E, Kumle M, Braaten T, et al. External validity in a population-based national prospective study–the Norwegian Women and Cancer Study (NOWAC). *Cancer Causes Control*. 2003;14(10):1001–1008. doi:10.1023/B:CACO.0000007982.18311.2e
63. Burggraf C, Teuber R, Brosig S, Meier T. Review of a priori dietary quality indices in relation to their construction criteria. *Nutr Rev*. 2018;76(10):747–764. doi:10.1093/nutrit/nuy027
64. Willett W. *Nutritional Epidemiology*. 3rd ed. Oxford: Oxford University Press; 2013.
65. Aleksandrova K, Nimptsch K, Pischon T. Influence of obesity and related metabolic alterations on colorectal cancer risk. *Curr Nutr Rep*. 2013;2(1):1–9. doi:10.1007/s13668-012-0036-9

Figure S1. Flow chart of study participants in the Norwegian Women and Cancer Study (NOWAC)

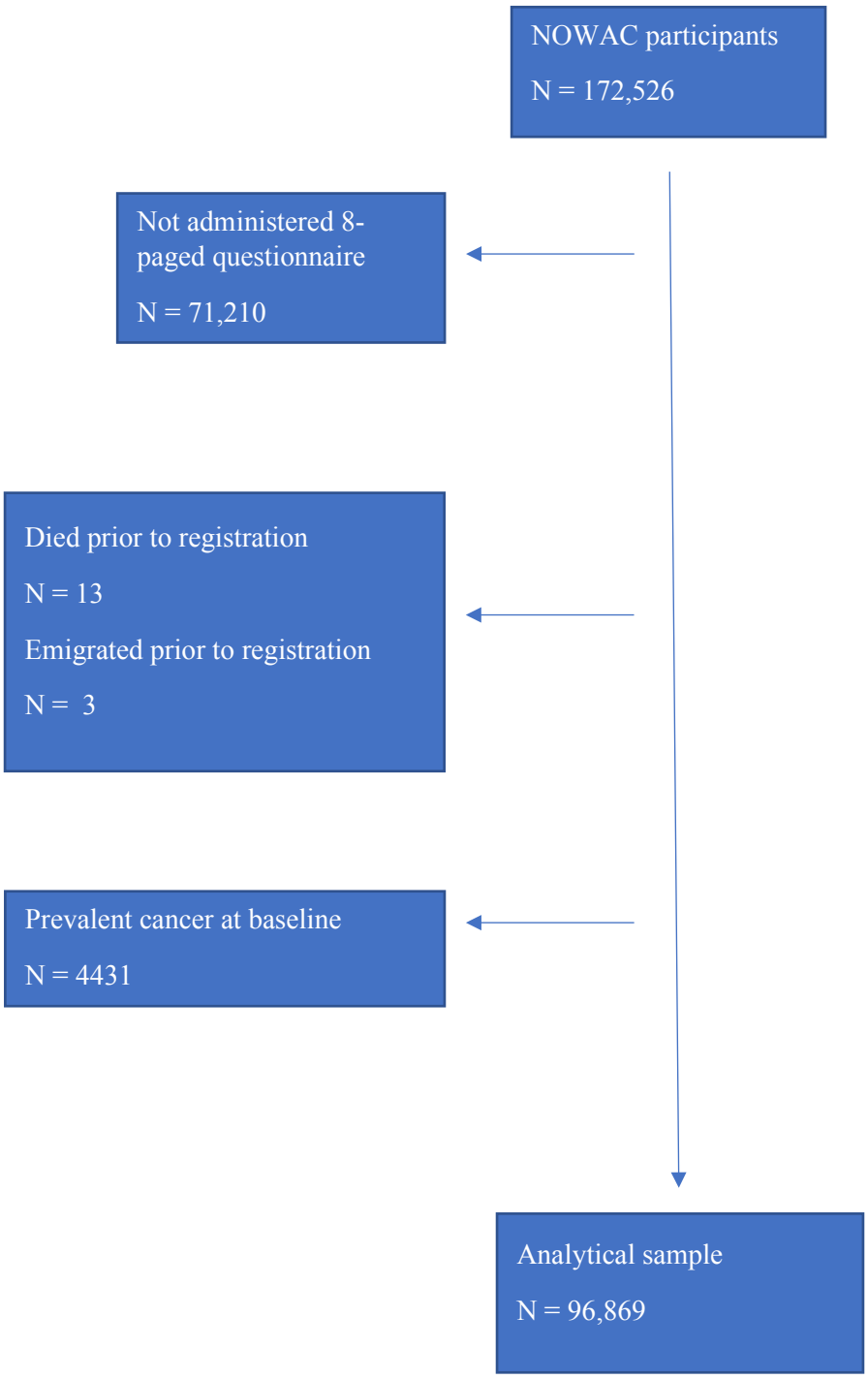


Table S1. Associations between the healthy lifestyle index (HLI) score and the incidence of common cancer types in the Norwegian Women and Cancer Study (NOWAC), 1996-2018, complete-case analysis

	HLI score				
	0-5	**6-10**	**11-15**	**16-20**	**Continuous**
	Cases (N) HR (95% CI)	Cases (N) HR (95% CI)	Cases (N) HR (95% CI)	Cases (N) HR (95% CI)	Cases (N) HR (95% CI)
Postmenopausal breast[a,b,c]					
	39	825	2014	519	3397
	1.05 (0.77-1.45)	1.08 (1.00-1.18)	1.00 (ref)	0.82 (0.75-0.91)	0.97 (0.96-0.98)
Colorectal[a]					
	18	281	715	199	1213
	1.46 (0.91-2.33)	1.07 (0.93-1.23)	1.00 (ref)	0.86 (0.73-1.01)	0.98 (0.96-1.00)
Lung[a]					
	33	349	557	67	1006
	3.30 (2.33-4.70)	1.67 (1.46-1.92)	1.00 (ref)	0.38 (0.29-0.49)	0.85 (0.83-0.87)
Postmenopausal endometrial[a,b]					
	13	228	444	122	807
	1.67 (0.96-2.90)	1.40 (1.19-1.64)	1.00 (ref)	0.84 (0.69-1.03)	0.93 (0.91-0.95)
Postmenopausal ovarian[a,b]					
	3	101	241	73	425
	0.67 (0.21-2.10)	1.11 (0.88-1.40)	1.00 (ref)	0.91 (0.70-1.19)	0.99 (0.96-1.02)
Pancreatic[a]					
	4	102	136	42	284
	1.72 (0.64-4.65)	2.03 (1.57-2.63)	1.00 (ref)	0.93 (0.65-1.32)	0.91 (0.88-0.95)
Kidney[a]					
	3	74	156	35	268
	1.06 (0.34-3.34)	1.26 (0.96-1.67)	1.00 (ref)	0.72 (0.50-1.03)	0.93 (0.90-0.97)

[a]Employed Cox proportional hazard regression models, adjusted for education and height

[b]Additionally adjusted for age at menarche, use of oral contraceptives, parity, breastfeeding, and use of hormone replacement therapy.

[c]Additionally adjusted for family history of breast cancer in a first-degree relative

Abbreviations: HR = hazard ratio, CI = confidence interval

This section displays Akaike's information criteria (AIC) values and plots generated for Cox proportional hazards regression models to estimate the associations between the HLI score, modelled as a nonlinear term, and incidences of common cancer types in the NOWAC cohort (1996-2018) on complete-case analysis data. Restricted cubic splines (RCS) were employed to model the HLI as a nonlinear term, where terms modelled with three, four, and five knots located at the percentiles have been tested. AIC values and plots of linear associations are also included for comparison.

Table S2. AIC values for linear and nonlinear models (3-5 RCS knots) among common cancer types

	Postmenopausal breast	Colorectal	Lung	Postmenopausal endometrial	Postmenopausal ovarian	Pancreatic	Kidney
5 knots	72488.83	25154.07	20430.72	17099.34	8888.43	5759.57	5603.20
4 knots	72486.92	25152.04	20433.95	17100.08	8886.32	5759.52	5602.23
3 knots	72484.97	25150.12	20431.63	17099.16	8885.36	5758.14	5600.22
Linear	72486.99	25149.02	20444.73	17097.21	8884.21	5757.35	5598.73

All of the following plots use the HLI score of 5 as the reference. Hazard ratios (HRs) are present on the vertical axes and HLI scores are present on the horizontal axes. Shaded regions represent 95% confidence intervals.

Figure S2. Postmenopausal breast cancer

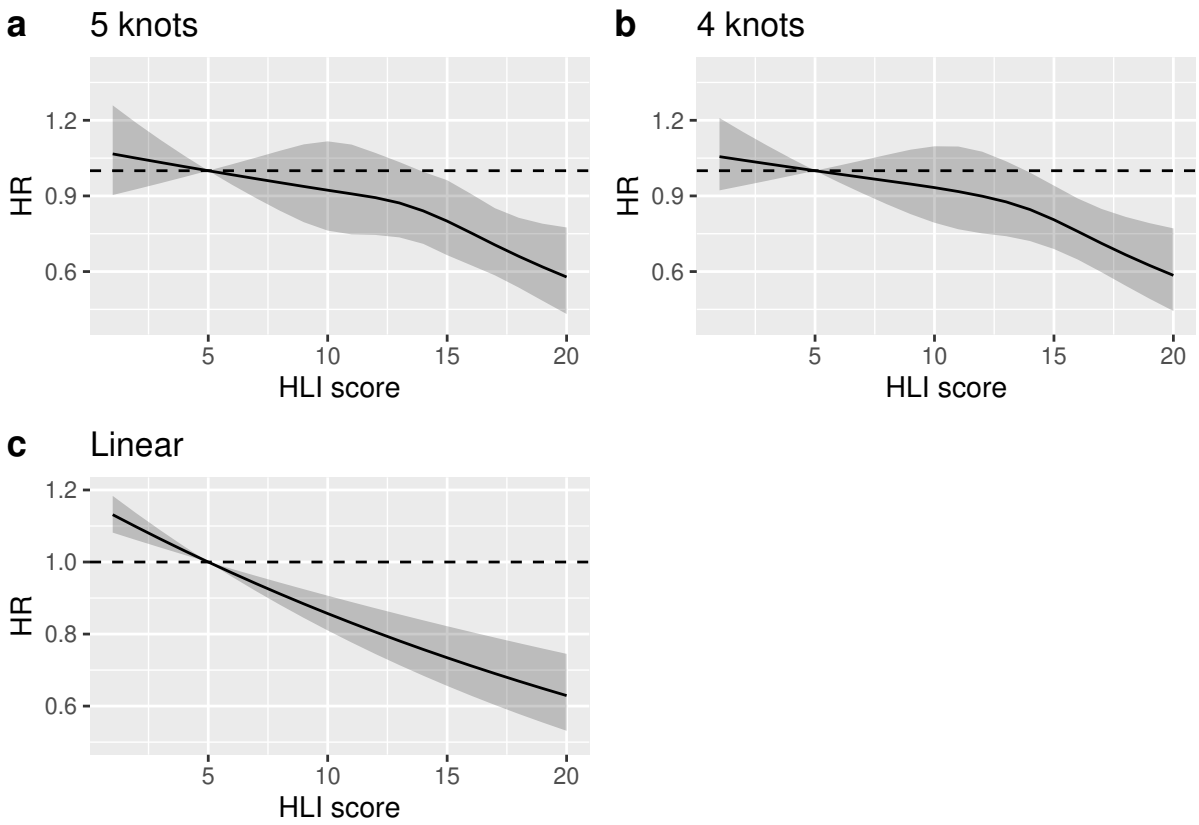


Figure S3. Colorectal cancer

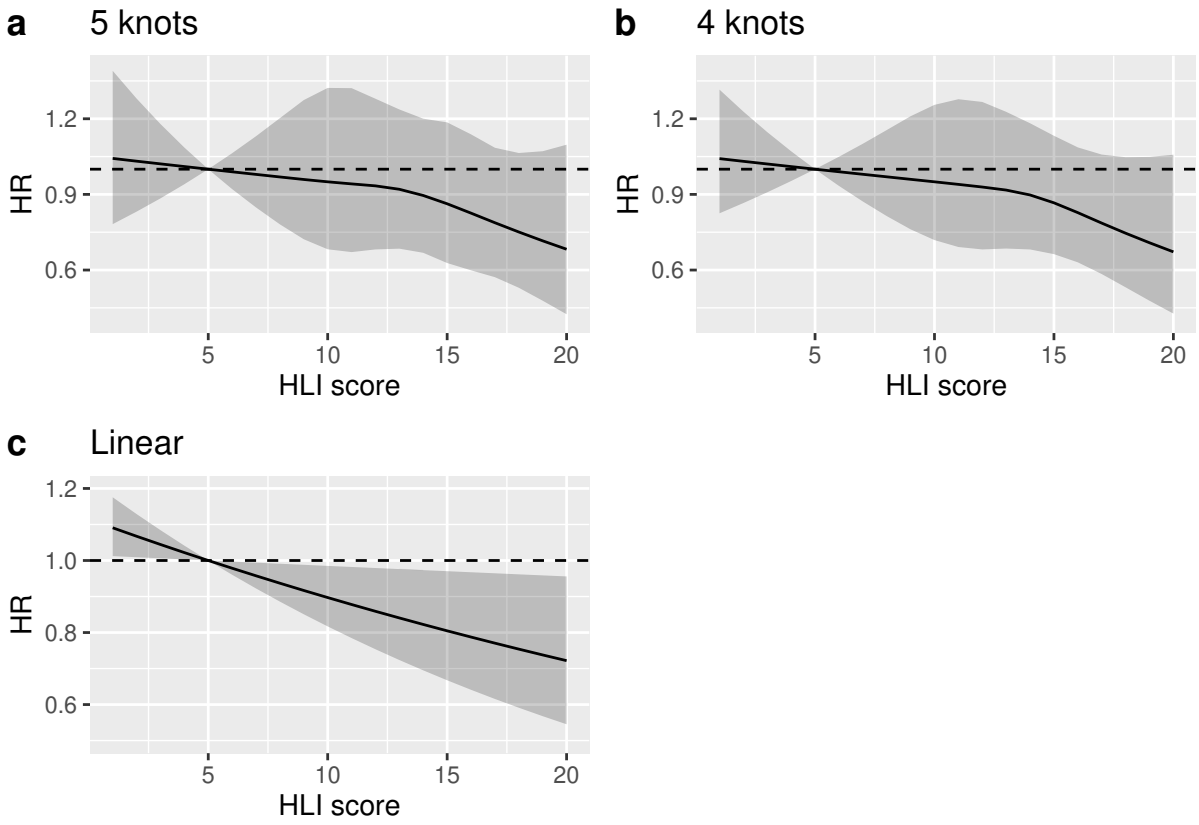


Figure S4. Lung cancer

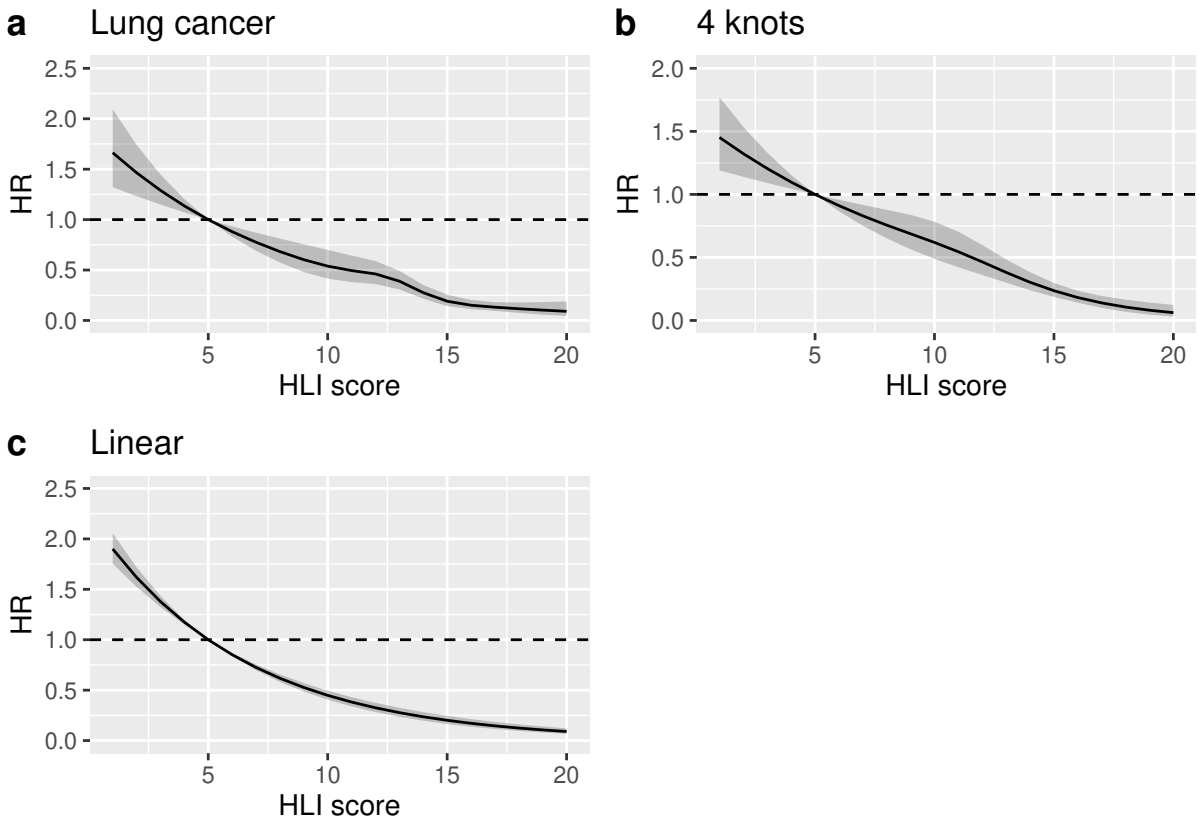


Figure S5. Postmenopausal endometrial cancer

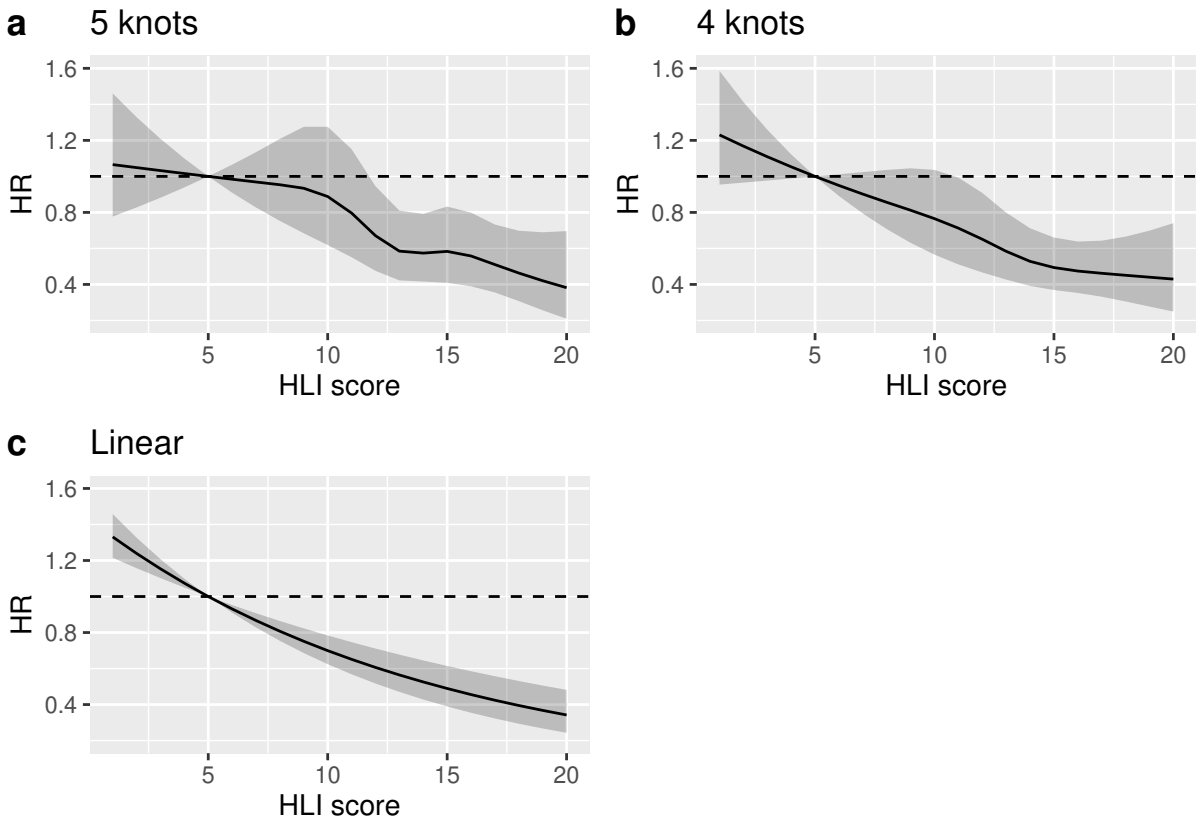


Figure S6. Postmenopausal ovarian cancer

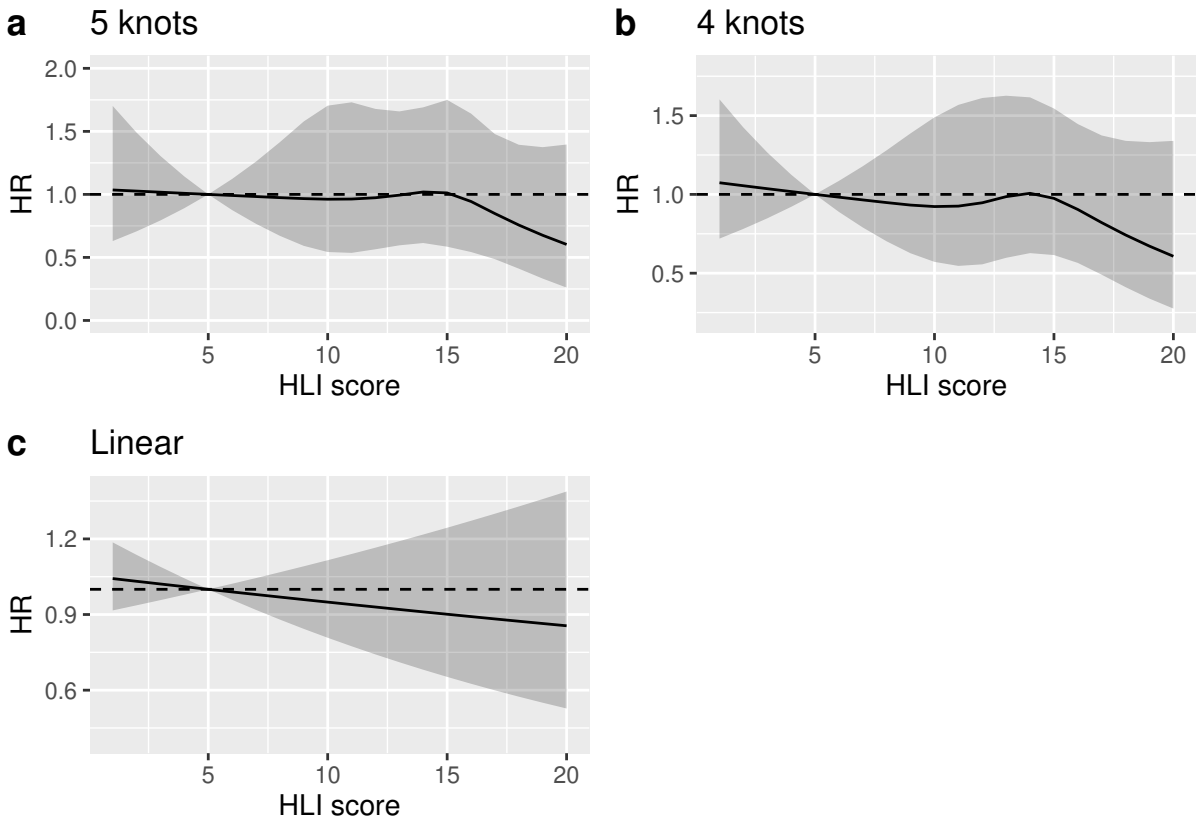


Figure S7. Pancreatic cancer

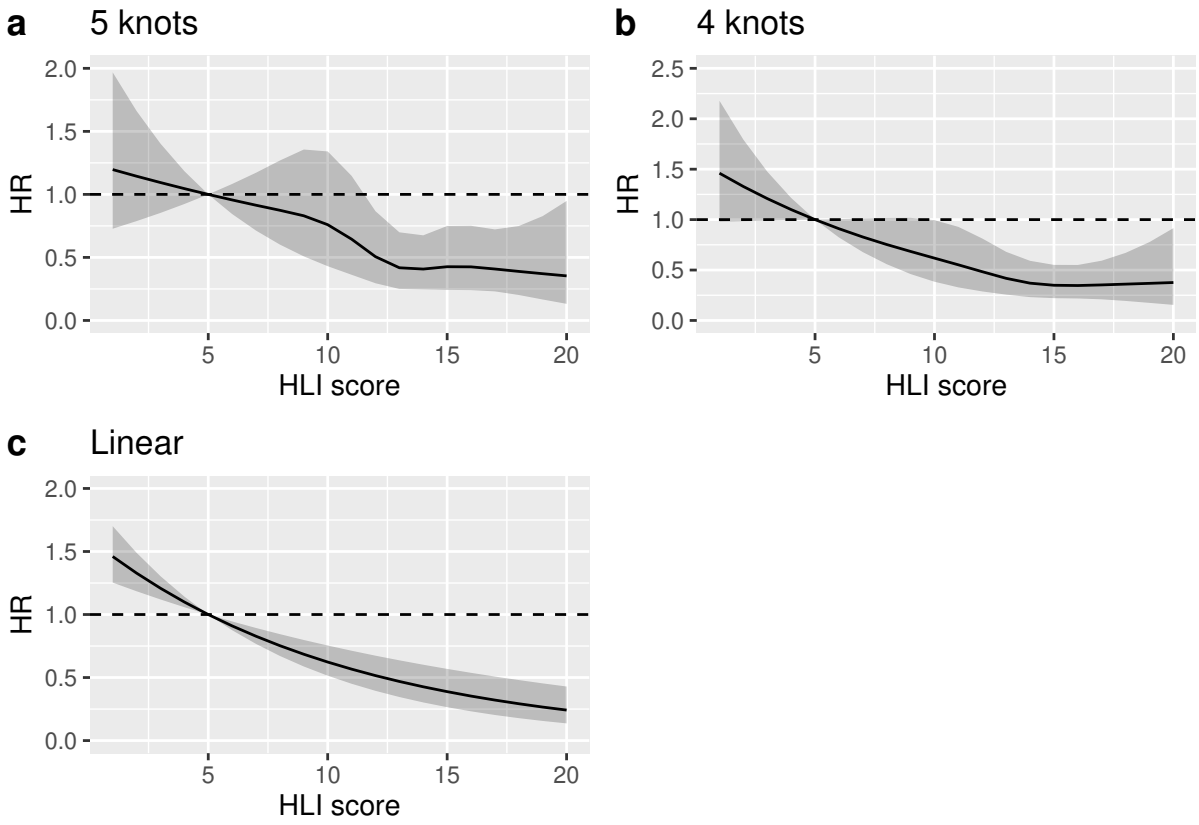


Figure S8. Kidney cancer

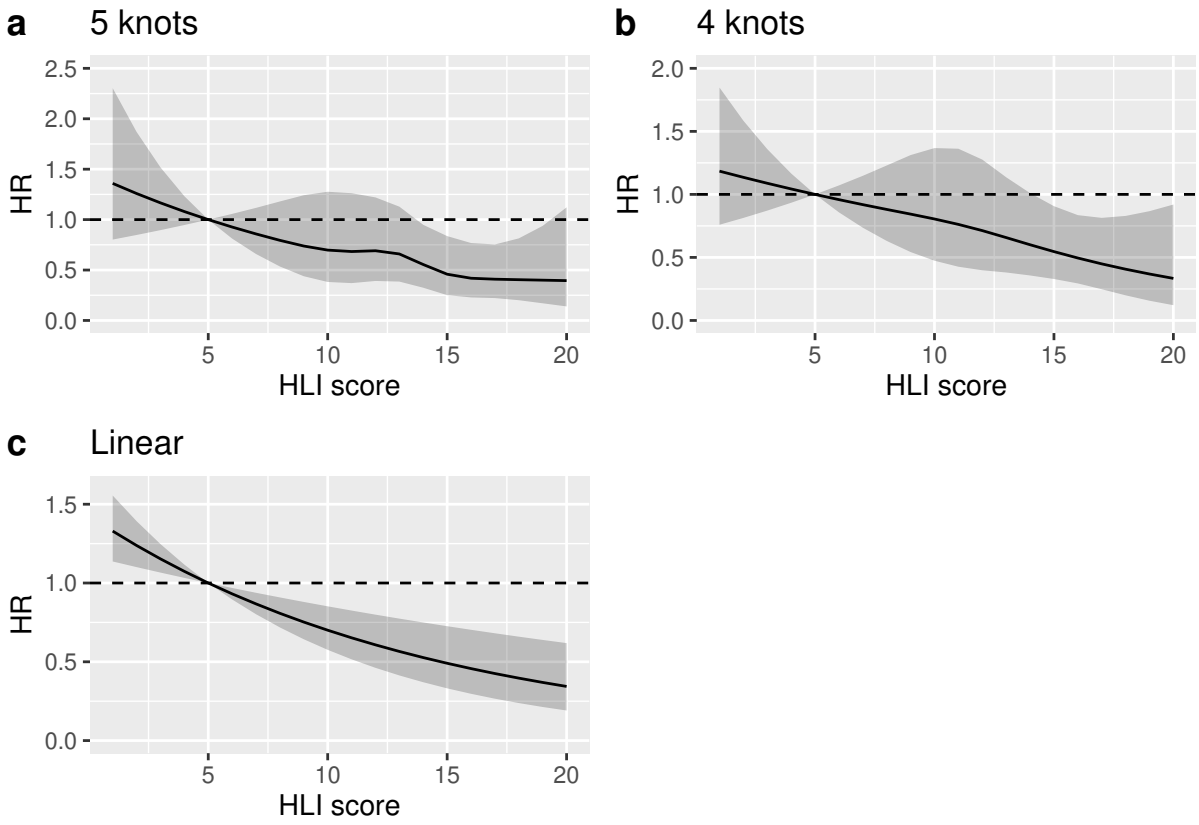


Table S3. Linear associations between the healthy lifestyle index (HLI) score excluding, in turn, each lifestyle factor, and the incidence of common cancer types in the Norwegian Women and Cancer Study (1996-2018)

	HR (95% CI)[a]
Postmenopausal breast[c,d,e]	
HLI[b]	0.97 (0.96-0.98)
HLI excluding physical activity level	0.97 (0.95-0.98)
HLI excluding BMI	0.96 (0.95-0.98)
HLI excluding smoking status	0.97 (0.96-0.98)
HLI excluding alcohol consumption	0.97 (0.96-0.98)
HLI excluding diet	0.96 (0.95-0.97)
Colorectal[c]	
HLI[b]	0.98 (0.96-1.00)
HLI excluding physical activity level	0.97 (0.95-0.99)
HLI excluding BMI	0.98 (0.96-1.00)
HLI excluding smoking status	0.99 (0.97-1.01)
HLI excluding alcohol consumption	0.98 (0.97-1.00)
HLI excluding diet	0.98 (0.96-1.00)
Lung[c]	
HLI[b]	0.86 (0.84-0.87)
HLI excluding physical activity level	0.83 (0.82-0.85)
HLI excluding BMI	0.78 (0.76-0.79)
HLI excluding smoking status	1.01 (0.99-1.03)
HLI excluding alcohol consumption	0.86 (0.85-0.83)
HLI excluding diet	0.81 (0.80-0.83)
Postmenopausal endometrial[c,d]	
HLI[b]	0.93 (0.91-0.95)
HLI excluding physical activity level	0.93 (0.91-0.96)
HLI excluding BMI	0.97 (0.95-0.99)
HLI excluding smoking status	0.89 (0.87-0.91)
HLI excluding alcohol consumption	0.92 (0.90-0.94)
HLI excluding diet	0.93 (0.91-0.95)
Postmenopausal ovarian[c,d]	
HLI[b]	0.99 (0.96-1.02)
HLI excluding physical activity level	0.98 (0.94-1.01)
HLI excluding BMI	0.98 (0.95-1.01)
HLI excluding smoking status	1.01 (0.97-1.04)
HLI excluding alcohol consumption	0.99 (0.96-1.02)
HLI excluding diet	1.00 (0.96-1.03)
Pancreatic[c]	
HLI[b]	0.92 (0.89-0.95)
HLI excluding physical activity level	0.90 (0.87-0.94)
HLI excluding BMI	0.91 (0.87-0.94)
HLI excluding smoking status	0.97 (0.93-1.01)
HLI excluding alcohol consumption	0.91 (0.88-0.95)
HLI excluding diet	0.90 (0.86-0.93)

Kidney[c]	
HLI[b]	0.94 (0.91-0.97)
HLI excluding physical activity level	0.94 (0.90-0.98)
HLI excluding BMI	0.96 (0.92-1.00)
HLI excluding smoking status	0.94 (0.91-0.98)
HLI excluding alcohol consumption	0.92 (0.89-0.96)
HLI excluding diet	0.93 (0.89-0.97)

[a] HRs correspond to 1-point increase on the HLI score, performed on multiple imputed data

[b] Estimates for HLI (including physical activity level, BMI, smoking status, alcohol consumption, and diet) in the full model also appear in Table 3. of main article

[c]Employed Cox proportional hazard regression models, adjusted for education and height

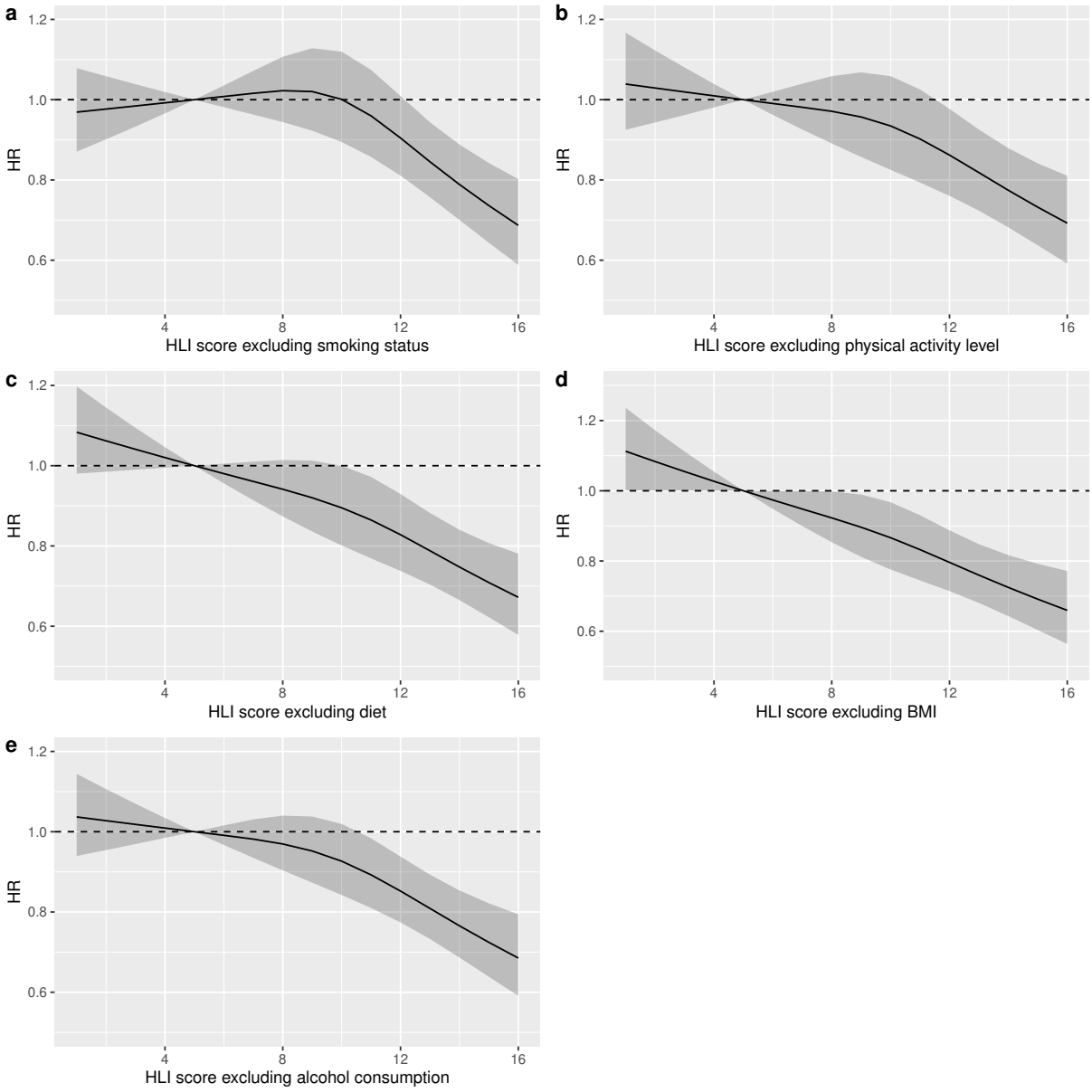
[d]Additionally adjusted for age at menarche, use of oral contraceptives, parity, breastfeeding, and use of hormone replacement therapy.

[e]Additionally adjusted for family history of breast cancer in a first-degree relative

Abbreviations: HR = hazard ratio, CI = confidence interval, BMI = body mass index

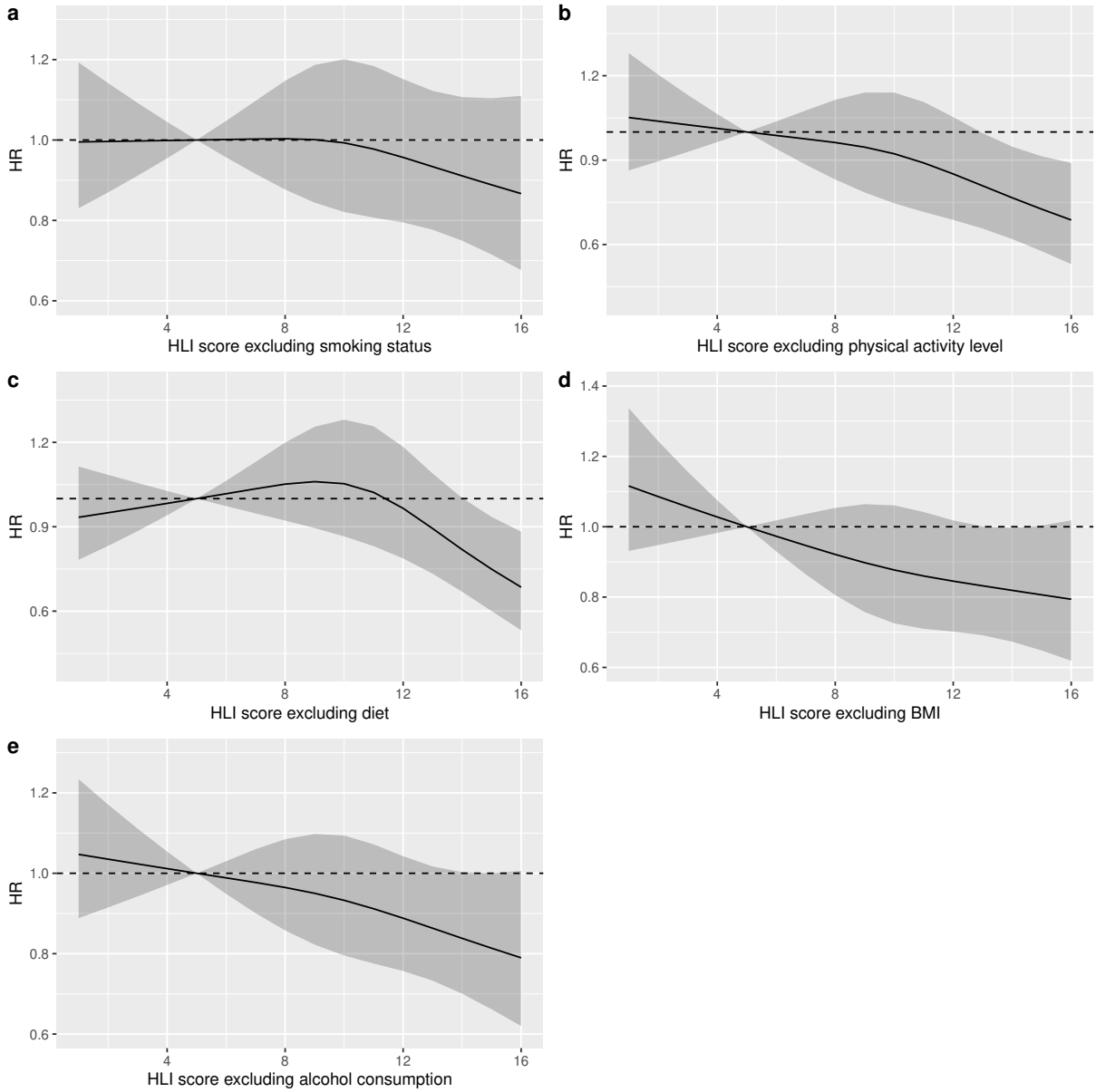
This section displays results from sensitivity analyses conducted on associations with the HLI modelled as a nonlinear term using restricted cubic splines (RCS) with three knots in the NOWAC cohort (1996-2018) on complete-cases analysis data. Single lifestyle factors were excluded from the HLI score, creating five reduced models. Thus, five estimated associations were obtained for each cancer outcome. Here, we visualise these results in plots.

Figure S9. Postmenopausal breast cancer incidence and its association with HLI score excluding single lifestyle factors. Models are adjusted for education, height, age at menarche, ever use of oral contraceptives, parity, breastfeeding, using of hormone replacement therapy, and history of breast cancer n a first-degree relative. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



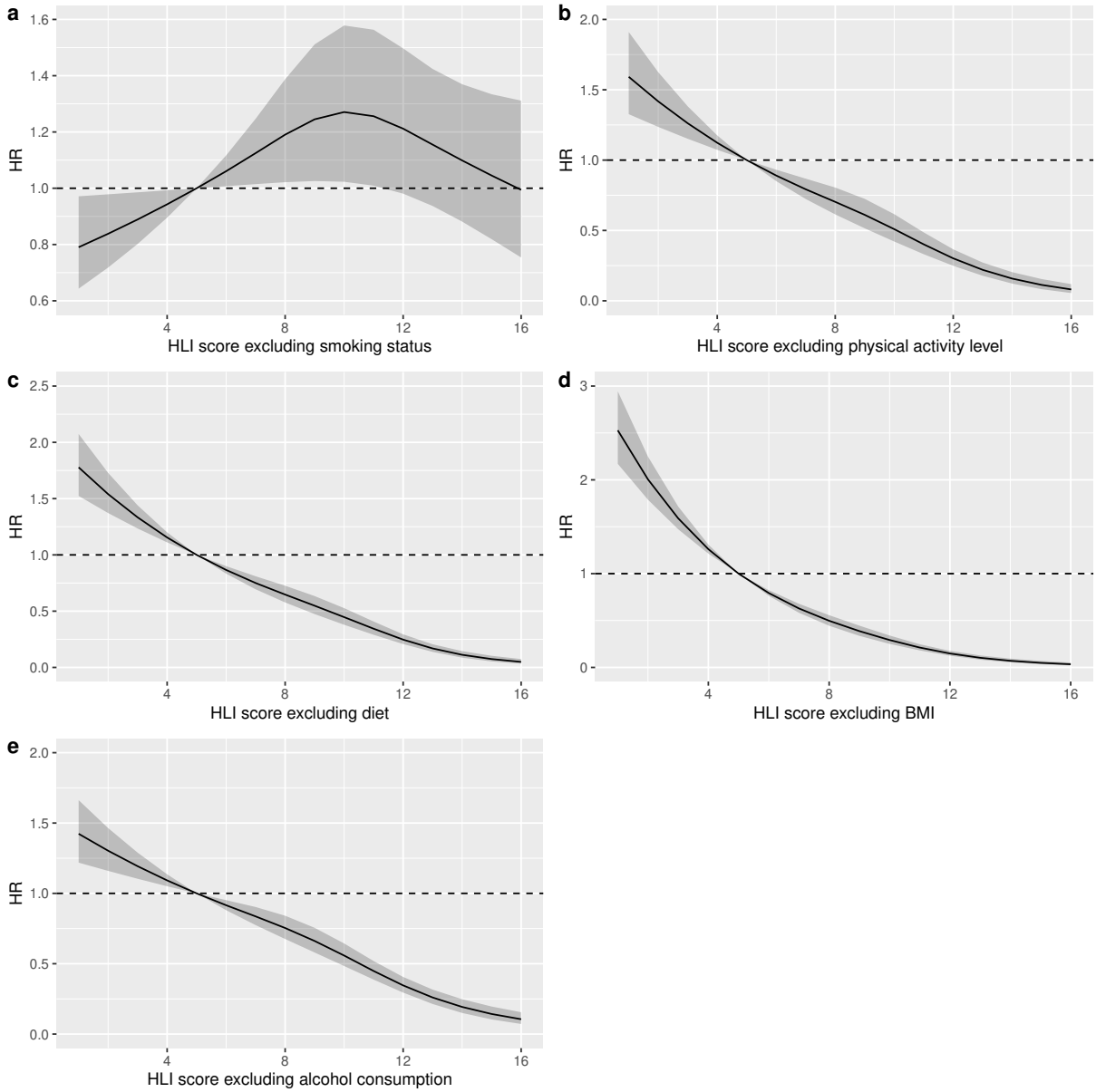
Abbreviations: HR = hazard ration, BMI = body mass index

Figure S10. Colorectal cancer incidence and its association with HLI score excluding single lifestyle factors. Models are adjusted for education and height. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



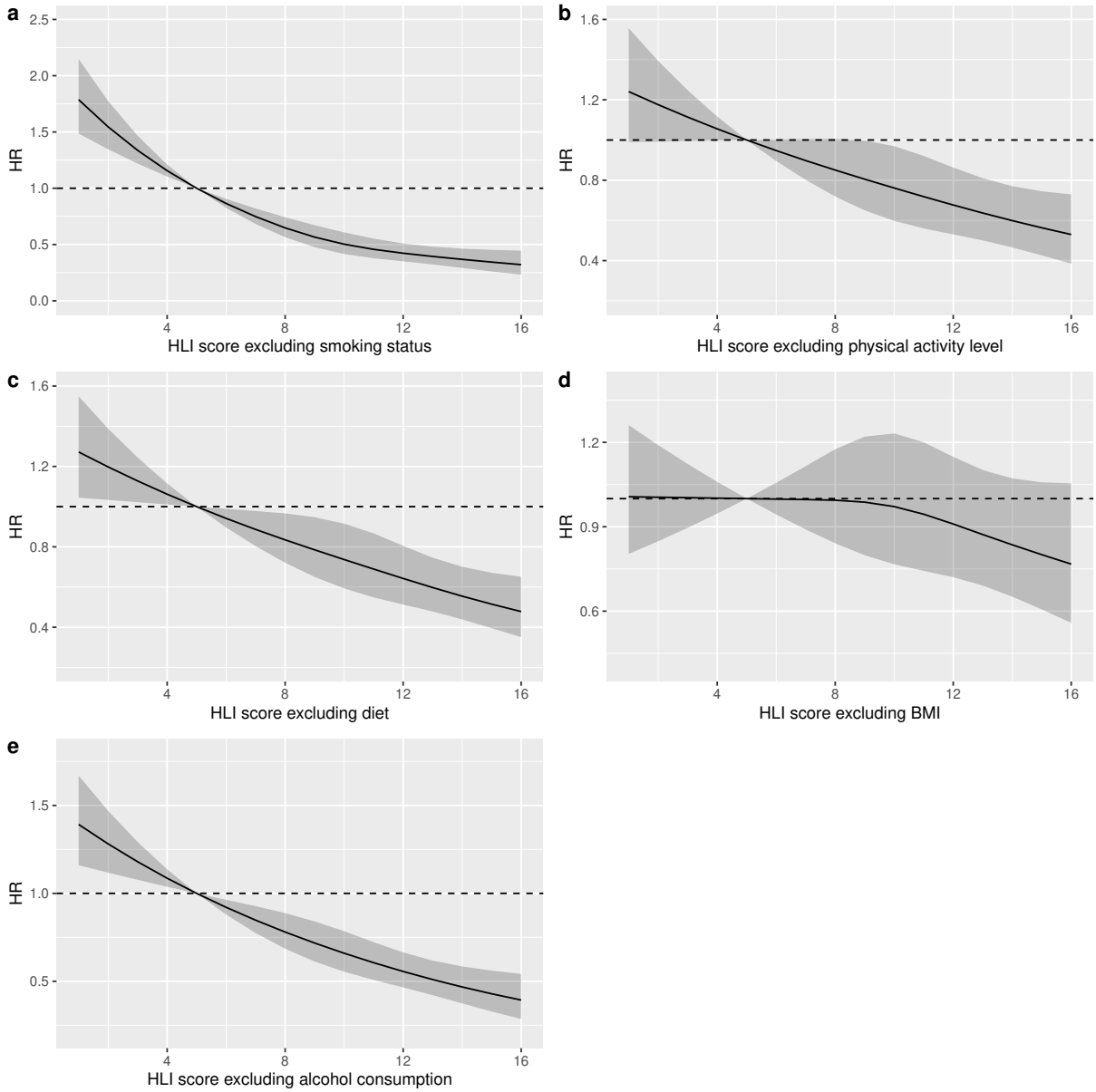
Abbreviations: HR = hazard ration, BMI = body mass index

Figure S11. Lung cancer incidence and its association with HLI score excluding single lifestyle factors. Models are adjusted for education and height. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



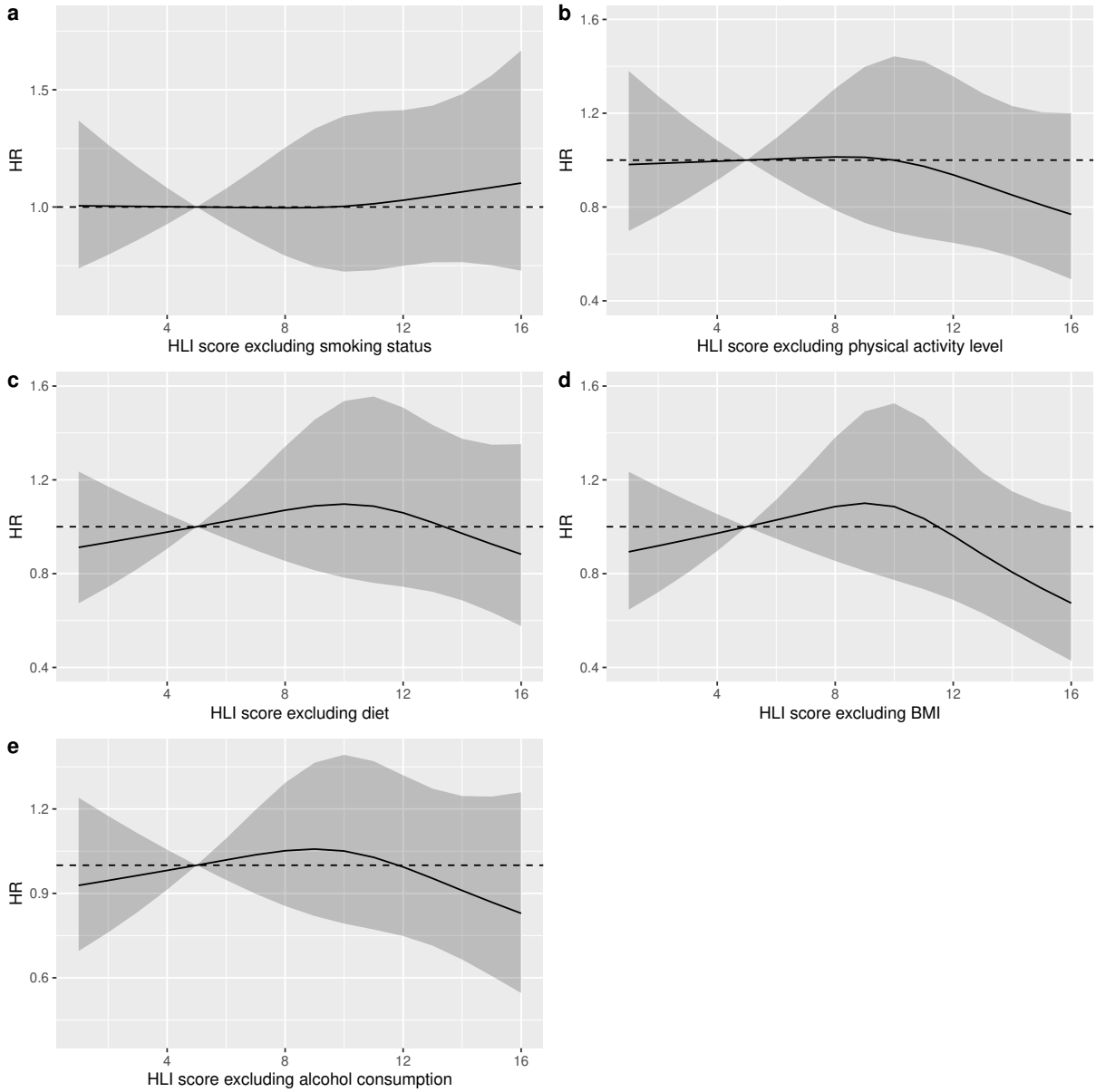
Abbreviations: HR = hazard ration, BMI = body mass index

Figure S12. Postmenopausal endometrial cancer incidence and its association with HLI score excluding single lifestyle factors. Models are adjusted for education, height, age at menarche, ever use of oral contraceptive, parity, breastfeeding, use of hormone replacement therapy. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



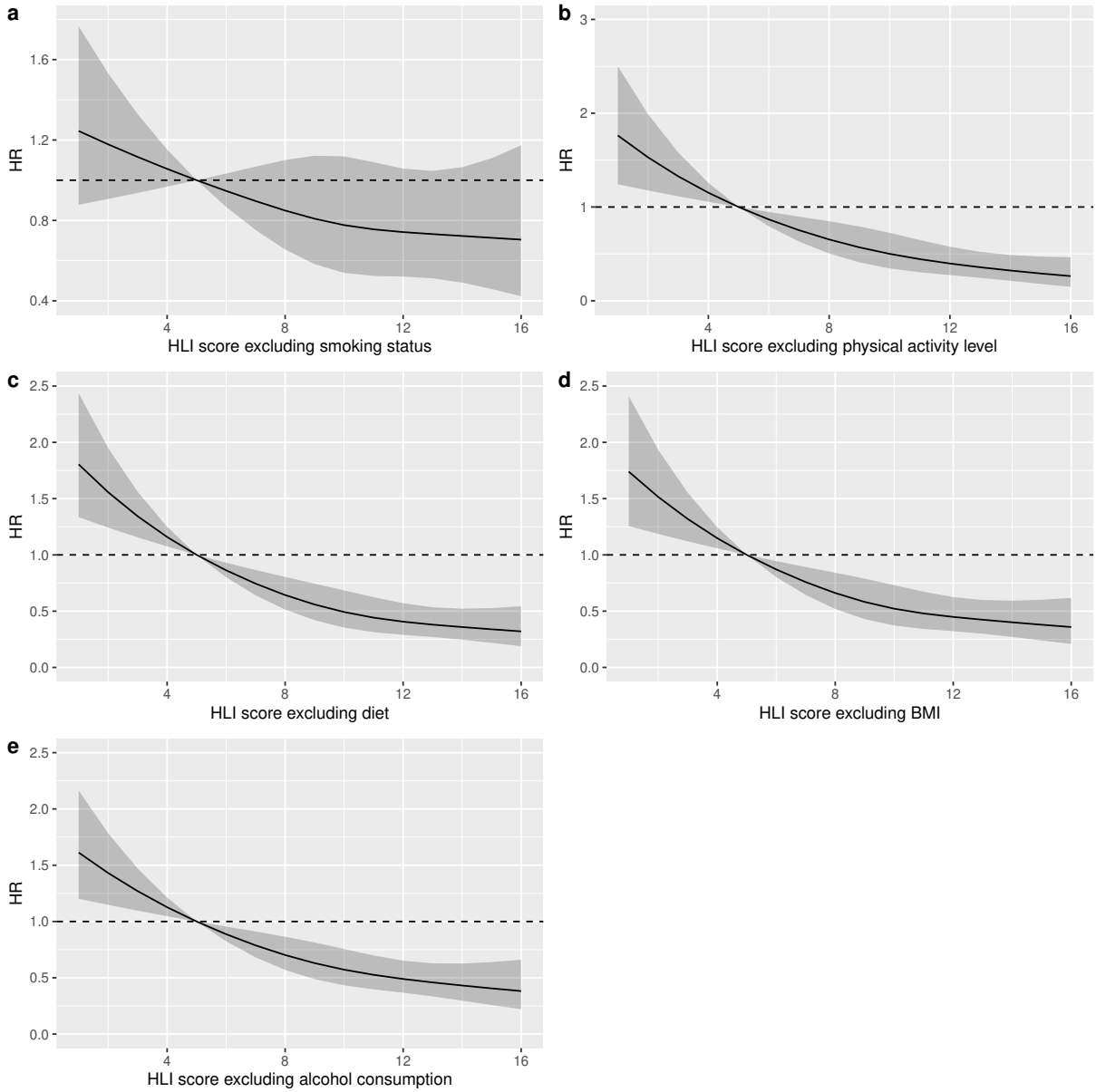
Abbreviations: HR = hazard ration, BMI = body mass index

Figure S13. Postmenopausal ovarian cancer incidence and its association with HLI score excluding single lifestyle factors. Models are adjusted for education, height, age at menarche, ever use of oral contraceptive, parity, breastfeeding, use of hormone replacement therapy. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



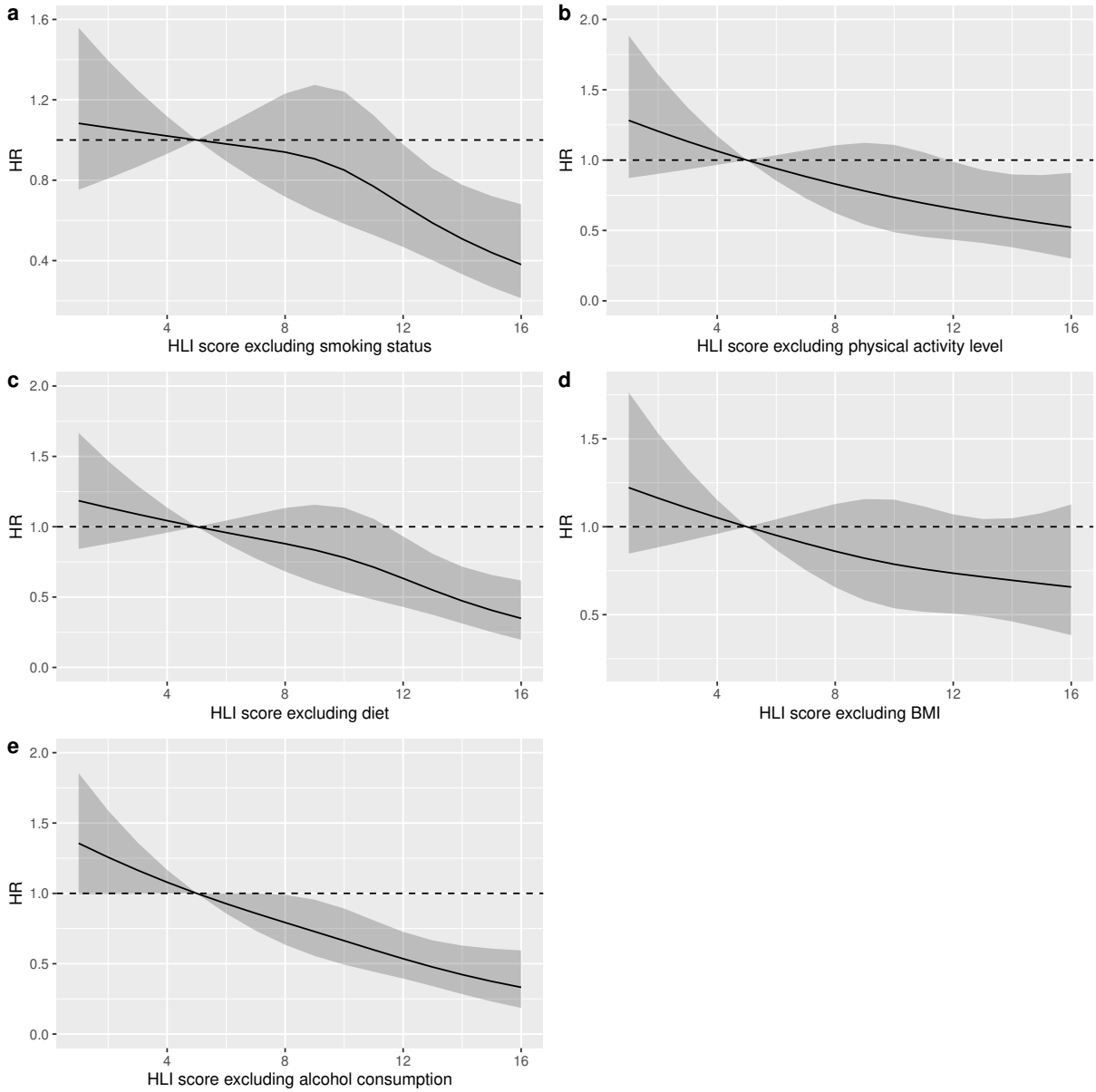
Abbreviations: HR = hazard ration, BMI = body mass index

Figure S14. Pancreatic cancer incidence and its association with HLI score excluding single lifestyle factors. Models are adjusted for education and height. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



Abbreviations: HR = hazard ration, BMI = body mass index

Figure S15. Kidney cancer incidence and its association with HLI score excluding single lifestyle components. Models are adjusted for education and height. The shaded regions represent the 95% confidence intervals and the reference HLI score is set to 5.



Abbreviations: HR = hazard ration, BMI = body mass index

Paper II

BMC Public Health

RESEARCH ARTICLE Open Access

Overall lifestyle changes in adulthood are associated with cancer incidence in the Norwegian Women and Cancer Study (NOWAC) – a prospective cohort study

Sairah L. F. Chen[1*], Therese H. Nøst[1], Edoardo Botteri[2,3], Pietro Ferrari[4], Tonje Braaten[1], Torkjel M. Sandanger[1] and Kristin B. Borch[1]

Abstract

Background Cancer is a leading cause of premature death worldwide and incidence is expected to rise in the coming decades. Many cohort studies, measuring lifestyle factors at one time-point, have observed that overall healthy lifestyles were inversely related to cancer incidence. However, there is little knowledge on the impact of lifestyle modification within adulthood.

Methods Using the Norwegian Women and Cancer study, two repeated self-reported assessments of lifestyle behaviours were used to calculate healthy lifestyle index scores at each time-point ($N = 66\,233$). The associations between change in healthy lifestyle index score and lifestyle-related cancer incidence, including alcohol-, tobacco-, obesity-, and reproductive-related, and site-specific breast and colorectal cancer incidence were estimated using Cox proportional hazard regression models. To assess nonlinearity in the dose–response relationships, restricted cubic spline models were used.

Results Independent of baseline lifestyle, positive lifestyle changes were inversely related to the incidence of overall lifestyle-related cancers, as well as alcohol-related, tobacco-related, obesity-related, and reproductive-related cancers, but not breast and colorectal site-specific cancers. An association between lifestyle worsening and cancer incidence compared to stable lifestyle was observed.

Conclusions This study provides evidence that overall lifestyle changes among cancer-free women between the ages of 41 and 76 impact the incidence of many cancer types. Regardless of baseline lifestyle, there was a negative dose–response relationship between magnitude of positive lifestyle change and the incidence of overall lifestyle-related cancers. We observed that underlying this trend was an especially clear association between lifestyle worsening and increased risk compared to stable lifestyle. For adult women, maintaining a stable healthy lifestyle and lifestyle improvement are important for preventing the occurrence of many cancer types.

Keywords Lifestyle, Cancer, Women, Change, Prevention, Composite index, Repeated measurements, Cohort

*Correspondence:
Sairah L. F. Chen
sairah.chen@uit.no
Full list of author information is available at the end of the article

Background

Cancer is a major public health concern. As a leading cause of premature death worldwide [1] and projected to surpass premature deaths caused by cardiovascular diseases, the cancer burden is and will be devastating. The International Agency for Research on Cancer (IARC) estimated 19.3 million new cancer cases in 2020. Due to population growth and aging, the predicted number of new cancer cases will increase by 47% from 2022 to 2040 [2]. An increase in the prevalence of unhealthy lifestyle behaviours will intensify this burden.

Based on evidence collected primarily in high-income countries, approximately 40% of cancer cases are preventable [3]. Studies focused on individual lifestyle factors from baseline assessments constitute much of the evidence linking key modifiable factors, including lack of physical activity levels, overweight and obesity, smoking, alcohol intake, and poor dietary habits, to increased cancer risk [4]. The assumption that lifestyle measured at one time point during adulthood will be maintained throughout time is pervasive and indeed pragmatic in epidemiology. Prevention strategies rightfully seek to shift populations towards healthy behaviours throughout the life course. However, the estimates of risk difference founding this public health engagement lacks an important dimension – the impact of lifestyle modification within adulthood of the individual.

For single risk factor changes during adulthood, smoking cessation is perhaps the most established lifestyle modification known to prevent especially lung and upper aerodigestive tract cancers [5, 6]. Weight gain is associated with higher risk of postmenopausal breast and endometrial cancer according to several studies [7–9], but the results are inconsistent with respect to other cancers and weight loss [7, 10, 11]. Improved and stable cardiorespiratory fitness is inversely associated with overall cancer incidence compared to reduced cardiorespiratory fitness [12]. In the prospective Norwegian Women and Cancer Study (NOWAC), increased physical activity over assessments collected 6 to 8 years apart was inversely associated with only colon cancer risk [13]. Alcohol cessation has been shown to be associated with lower risk of several cancers [14–16], yet studies investigating the impact of graded changes in alcohol intake are few and inconclusive [17, 18]. To our knowledge, there are no studies exploring the association between changes in dietary habits alone and cancer risk. However, a randomised study observed that smoking cessation combined with dietary intervention reduced the risk of lifestyle-related cancers among men at high risk for cancer [19]. Changing several lifestyle factors has only been investigated in one additional study, observing that Swedish women who maintained or improved their lifestyle were at lower risk for lifestyle-related cancer compared to those who had consistently poor lifestyle [17]. However, the study did not include diet, which is an important element of lifestyle as it relates to cancer risk [4].

More evidence is required to understand the impact of lifestyle changes, involving individual and combined factors, on cancer risk. In this study the association between changing several lifestyle factors combined during adulthood, as measured by the healthy lifestyle index (HLI) score, on lifestyle-related cancer incidence was investigated in a cohort of Norwegian women.

Methods

Study sample and data collection

The NOWAC study has been described in detail previously [20] and has been used to investigate a wide range of lifestyle factors and health outcomes. In brief, the NOWAC study is a nationwide, prospective cohort consisting of approximately 172 000 adult female participants. Women invited to participate in the NOWAC study were randomly sampled from the Norwegian Central Person Register between 1991 and 2007 in multiple sub-cohorts. Consenting participants completed a self-administered questionnaire at enrolment and were invited to complete a maximum of three follow-up self-administered questionnaires, where each questionnaire was distributed between 2 and 11 years apart. All questionnaires, including follow-up questionnaires, collected information on socio-demographic characteristics, reproductive and hormonal factors, self-reported health, physical activity level, height, weight, smoking habits, dietary habits, and family history of breast cancer. Questionnaires consisted of either 4 or 8 pages depending on the sub-cohort, with the 8-page questionnaire containing a detailed food frequency questionnaire (FFQ). The first completed 8-page questionnaire was used as the baseline measurement for the present study (Q1) (Additional File 1). The subsequent completed 8-page follow-up questionnaire was used as the follow-up measurement (Q2) (Additional File 2). Participants that did not complete at least two 8-page questionnaires were excluded. In this study, Q1 was administered from 1996 to 2004 and Q2 was administered from 2002 to 2014.

The Norwegian personal identity number assigned to every resident of Norway and its linkage to the Cancer Registry of Norway, Cause of Death Register, and National Population register allowed for complete follow-up for all participants. Women who had died ($n=3$), emigrated ($n=2$) or had been diagnosed with cancer ($n=5018$) before Q2 were excluded (see Additional File 3 for sample flow chart). A total of 66 233 participants were included in the analysis where 44 403 participants had complete information on lifestyle factors at two

timepoints. The timeline of the final sample is shown in Fig. 1.

Assessing lifestyle change

A healthy lifestyle index (HLI) was used to quantify overall lifestyle quality at Q1 and Q2. The construction of the HLI score in the NOWAC cohort was presented previously [21]. Briefly, the HLI used for this analysis consisted of five modifiable lifestyle factors – physical activity level, body fatness assessed by BMI (kg/m^2), smoking behaviour, alcohol consumption (grams/day), and a dietary score. Physical activity level was reported by participants on a 1 to 10 scale ranging from not active to very active, where participants were asked to consider the entirety of activity at work, outside work, at home, exercise, and other forms of physical activity. Smoking behaviour was measured by smoking status, time since cessation for former smokers, and current number of cigarettes smoked per day. Each lifestyle factor was assigned a score ranging from 0 to 4, which were summed to a total HLI score that ranged from 0 to 20, where higher scores indicated a healthier lifestyle. See Additional File 4 for details on HLI construction. The HLI score change was the difference between HLI score at Q2 and Q1, where positive score changes represented lifestyle improvement and negative score changes represented lifestyle worsening.

Outcome ascertainment

Follow-up time began at the end of Q2 and lasted until December 2018. Date of death and emigration were obtained through linkage to the Central Population Registry of Norway. Cancer diagnosis and date of diagnosis were obtained through linkage to the Cancer Registry of Norway based on codes from the International Classification of Diseases, Tenth Revision (ICD-10). The present study investigated all cancers considered to be lifestyle-related, constituting this study's total cancer cases, and several cancer subgroupings including alcohol-related, tobacco-related, obesity-related, and reproductive-related cancers based on the IARC monograph on known causes and prevention by organ site (Additional File 5) [22]. Breast and colorectal cancer incidence were also investigated separately.

Statistical analysis

Cox proportional hazards regression model, with age as the time scale, was used to estimate hazard ratios (HR) and 95% confidence intervals (CI). Age at entry was participants' age at Q2 and age at exit was age at cancer diagnosis, death, emigration, or age in 31 December 2018, whichever occurred first. Associations were estimated between continuous (per 1 SD increase) and categorical change in HLI score and incidence of alcohol-related, tobacco-related, obesity-related, breast- and reproductive-related, and lifestyle-related cancer incidence. Seven categories for HLI score change were used: ≤ 3, 2, and 1 point decrease, stable, 1, 2, and ≥ 3 point increase. The proportional hazards assumption was tested using Schoenfeld residuals. Potential non-linear associations were tested with restricted cubic splines, modelled with three knots located at the predictor minimum and maximum, and the remaining at the equidistant percentile (50th), as recommended by Harrell [23]. Likelihood ratio tests were used to compare goodness of fit between non-linear and linear models.

The confounders included in the models were based on previous literature and determined a priori. They included education (years), height (centimetres), HLI

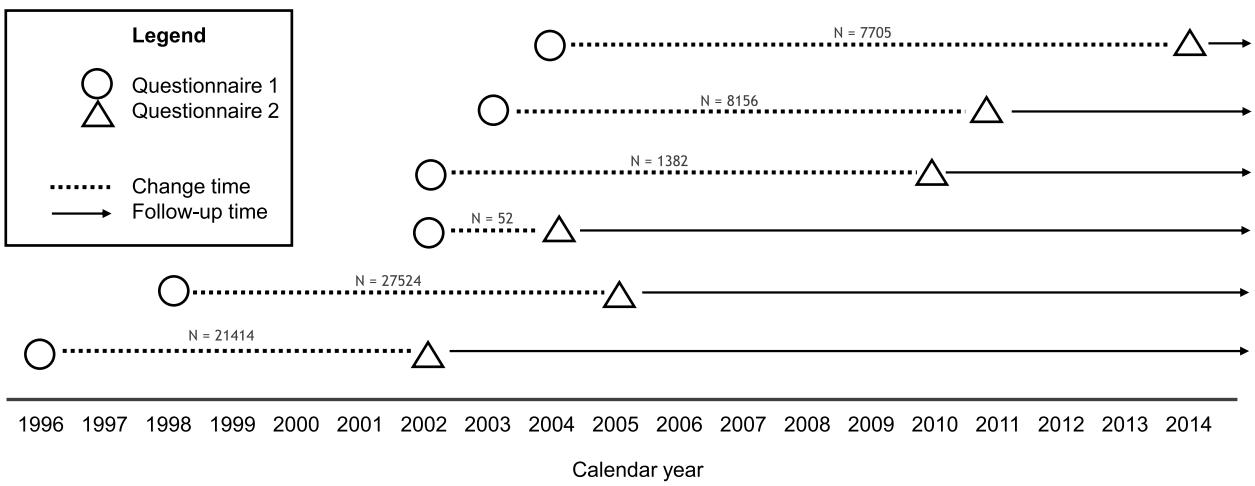


Fig. 1 Timing of data collections and start of follow-up for the analytical sample, *N* = 66 233, Norwegian Women and Cancer Study (NOWAC)

score at Q1 (continuous) and calendar year at Q2 (continuous). Alcohol-related, obesity-related, reproductive-related, breast, and lifestyle-related cancer models were additionally adjusted for age at menarche (years), menopausal status (premenopausal/postmenopausal), breastfeeding (cumulative months $0, <=12, >12$ months), hormone replacement therapy use (current, former, never), oral contraceptive use (ever, never), parity (0, 1–2, > 2), and history of breast cancer in first degree relatives (yes, no). Perimenopausal women were considered premenopausal. Women missing on menopausal status were reported as postmenopausal if age 53 or older at the time of Q2.

The associations between individual index components, modelled as single lifestyle factor scores (continuous), and all outcomes were estimated using the same outcome-based adjustment sets described above. All single lifestyle factor scores were included in the model as they were the exposures of interest and mutually adjusted for one another. Correlation between HLI score changes across years between Q1 and Q2 was assessed with the Pearson correlation coefficient.

Sensitivity analysis

To assess the contribution of each lifestyle factor to the associations between HLI score change (continuous) and cancer outcomes, the scores for physical activity, BMI, smoking, alcohol, and diet were excluded, one by one, from the HLI. Potential period effects were tested by performing analysis within two stratified enrolment year groups (1996–98, 2002–04) detailed in Fig. 1. The presence of effect modifications by age at Q1, age at Q2, and Q1 HLI score categories (HLI score 0–11, 12–13, 14–15, 16–20) were tested by modelling interaction terms and comparing models including the interaction term to the model without the interaction term using the likelihood ratio test. The first two years of follow-up time were excluded to test the impact of intentional or unintentional lifestyle changes due to morbid conditions, included pre-diagnosed cancer. The association between HLI change and all remaining cancers (not lifestyle-related) was estimated to test the viability of the lifestyle-related cancer grouping.

Multiple imputation

Missing data among variables constituting the HLI score at Q1 and/or Q2, and covariates for 21 830 participants were handled by multiple imputation chained equations (MICE) under the assumption that data were missing at random [24]. All covariates included in the cancer models and the Nelson Aalen cumulative hazard estimator were included in the MICE model. MICE analysis employed fully conditional specification, whereby each incomplete variable was modelled iteratively by a series of multivariable regression models [25]. A total of 100 datasets were generated with 10 iterations each. Parameter estimates in the Cox models from each imputed dataset were averaged through the Rubin's rule [26] to account for uncertainty in the MICE models to impute missing values. Model parameters were also estimated in complete-case analyses. Descriptive statistics for each imputed variable were compared between observed and imputed values. Convergence of MICE models were assessed by visual inspection of plots of the mean HLI score change against iteration number for each MI dataset (not shown).

All data treatment and statistical analysis were conducted in RStudio Version 1.2.959 with R Version 4.0.3 [27]. All statistical hypotheses were tested two-sided, allowing a Type I error rate of 5%.

Results

At the start of follow-up (Q2), the mean age was 58.2 years, 46% of participants reported a physical activity level ≥ 6 on the NOWAC 1–10 scale, mean BMI was 25.4 (kg/m^2), 20.7% were current smokers, median daily intake of alcohol was 2.09 g/day (mean: 4.0 g/day; IQR: 0.6, 5.8), the median expanded diet score was 9, and the median HLI score was 13 (Table 1). The mean time between Q1 and Q2 was 7 years (range: 2 – 11) with a mean HLI score change of -0.2 (range: -11 to 14). There was no correlation between the number of years between Q1 and Q2 and HLI score change ($r = -0.06$). The largest proportion of participants exhibited an HLI score difference of zero (17.2%), followed by decrease of three points (16.1%), decrease of one point (16.0%), increase of one point (15.2%), decrease of two points (12.7%), increase of three points (12.3%), and increase of two points (10.5%). The distributions of HLI change scores within Q1 HLI groups (Q1 HLI score group 0–11, 12–13, 14–15, 16–20) were different across Q1 HLI groups, reflecting the constraints of maximum and minimum change on the HLI and thus the probability distribution (Additional File 6).

The median follow-up time was 14.2 years during which 6 384 lifestyle-related cancer cases occurred, reflecting the total number of cancer cases. Within overlapping cancer groupings, there were 3 512 alcohol-related, 2 931 tobacco-related, 4 788 obesity-related, 3 385 reproductive-related, 2 384 breast, and 839 colorectal cancer cases that occurred.

The estimates obtained from MICE data models were within ± 5% of those obtained from complete-case data models for continuous exposure models and demonstrated a similar trend for categorical exposure models (see complete-case results in Additional File 7). Therefore, all presented estimates were obtained from MICE

Table 1 Characteristics of the study population at the start of follow-up (Questionnaire 2) according to healthy lifestyle index score change in the Norwegian Women and Cancer Study (*N* = 66,233)

		HLI score change							
	Total (*N* = 66,233)	**Decrease 3 or more (*N* = 7169)**	**Decrease 2 (*N* = 5636)**	**Decrease 1 (*N* = 7099)**	**Stable (*N* = 7638)**	**Increase 1 (*N* = 6743)**	**Increase 2 (*N* = 4673)**	**Increase 3 or more (*N* = 5445)**	**Missing, N (%)**
Age(years)	58.2 (6.3)	57.2 (5.9)	57.3 (6.0)	57.2 (5.9)	57.3 (6.0)	57.3 (6.1)	57.4 (6.2)	57.3 (6.0)	0(0)
Education (years)	12.3 (3.5)	12.5 (3.4)	12.7 (3.4)	12.6 (3.4)	12.7 (3.5)	12.6 (3.4)	12.6 (3.4)	12.5 (3.4)	3471 (5%)
HLI score at Q2, median (IQR)	136 (11, 15)	10 (9, 12)	12 (10, 14)	13 (10, 14)	13 (11, 15)	14 (12, 15)	14 (12, 16)	14 (13, 16)	16,343 (25%)
Physical activity score change	0.1 (1.4)	-1.2 (1.3)	-0.5 (1.1)	-0.2 (1.1)	0.1 (1.0)	0.5 (1.1)	0.8 (1.1)	1.5 (1.3)	11,969 (18%)
BMI score change	-0.3 (0.8)	-0.7 (0.8)	-0.5 (0.7)	-0.4 (0.7)	-0.2 (0.7)	-0.1 (0.7)	-0.0 (0.7)	0.2 (0.8)	3713 (6%)
Smoking score change	0.1 (0.5)	-0.0 (0.6)	0.0 (0.5)	0.1 (0.5)	0.1 (0.5)	0.1 (0.5)	0.2 (0.5)	0.3 (0.6)	5939 (9%)
Alcohol score change	-0.1 (0.7)	-0.4 (0.7)	-0.2 (0.7)	-0.2 (0.6)	-0.1 (0.6)	-0.0 (0.6)	0.0 (0.6)	0.2 (0.7)	2920 (4%)
Diet score change	0.0 (1.6)	-1.5 (1.4)	-0.8 (1.2)	-0.4 (1.2)	0.1 (1.1)	0.6 (1.2)	1.0 (1.2)	1.6 (1.3)	6914 (10%)
Height (cm)	166.0 (5.7)	166.2 (5.7)	166.4 (5.7)	166.1 (5.6)	166.3 (5.7)	166.2 (5.6)	166.3 (5.7)	166.4 (5.6)	1705 (3%)
Physical activity level (> = 6)[a], N (%)	30,391 (46%)	2182 (30%)	2457 (44%)	2457 (44%)	4262 (56%)	4064 (60%)	3041 (65%)	3956 (73%)	8572 (13%)
Body mass index (kg/m^2)	25.4 (4.2)	26.3 (4.1)	25.7 (4.1)	25.5 (4.1)	25.1 (4.1)	25.1 (4.2)	24.9 (4.1)	25.0 (4.1)	2633 (4%)
Smoking status, N (%)									2913 (4%)
Never	22,653 (34%)	2475 (35%)	2189 (39%)	2726 (38%)	3007 (39%)	2550 (38%)	1724 (37%)	1650 (30%)	
Former	26,942 (41%)	2986 (42%)	2263 (40%)	2907 (41%)	3090 (40%)	2751 (41%)	2012 (43%)	2687 (49%)	
Current	13,725 (21%)	1708 (24%)	1184 (21%)	1466 (21%)	1541 (20%)	1442 (21%)	937 (20%)	1108 (20%)	
Alcohol intake (g/day)	4.0 (5.0)	5.1 (5.8)	4.5 (5.2)	4.4 (5.3)	4.2 (4.9)	4.0 (4.9)	3.9 (4.5)	3.8 (4.6)	3387 (5%)
Diet score (0–18)	8.8 (2.5)	7.3 (2.3)	8.0 (2.4)	8.5 (2.5)	9.0 (2.5)	9.3 (2.3)	9.6 (2.2)	9.9 (2.1)	4786 (7%)
Postmenopausal, N (%)	52,110 (79%)	5415 (76%)	4270 (76%)	5346 (75%)	5798 (76%)	5048 (75%)	3528 (75%)	4111 (76%)	0(0)
Age at menarche (years)	13.3 (1.4)	13.2 (1.4)	13.3 (1.4)	13.3 (1.4)	13.3 (1.4)	13.3 (1.4)	13.3 (1.4)	13.3 (1.4)	921 (1%)
Hormone replacement therapy status, N (%)									0(0)
Never	45,276 (68%)	4970 (69%)	3865 (69%)	4924 (69%)	5260 (69%)	4521 (67%)	3158 (68%)	3640 (67%)	
Former	6852 (10%)	705 (10%)	546 (10%)	650 (9%)	737 (10%)	705 (10%)	479 (10%)	550 (10%)	
Current	14,105 (21%)	1494 (21%)	1225 (22%)	1525 (21%)	1641 (21%)	1517 (22%)	1036 (22%)	1255 (23%)	
Oral contraceptive ever use, N (%)	35,451 (54%)	4205 (59%)	3262 (58%)	3993 (56%)	4344 (57%)	3823 (57%)	2669 (57%)	3143 (58%)	0(0)
Parity, N (%)									0(0)
0	5411 (8%)	576 (8%)	475 (8%)	580 (8%)	611 (8%)	585 (9%)	398 (9%)	465 (9%)	
1–2	34,666 (52%)	3825 (53%)	2987 (53%)	3850 (54%)	4149 (54%)	3667 (54%)	2524 (54%)	3056 (56%)	

Table 1 (continued)

		HLI score change							
	Total (*N* = 66,233)	**Decrease 3 or more (*N* = 7169)**	**Decrease 2 (*N* = 5636)**	**Decrease 1 (*N* = 7099)**	**Stable (*N* = 7638)**	**Increase 1 (*N* = 6743)**	**Increase 2 (*N* = 4673)**	**Increase 3 or more (*N* = 5445)**	**Missing, N (%)**
> 2	26,156 (39%)	2768 (39%)	2174 (39%)	2669 (38%)	2878 (38%)	2491 (37%)	1751 (37%)	1924 (35%)	
Cumulative breastfeeding duration (months), N (%)									0(0)
0	34,402 (52%)	3950 (55%)	3229 (57%)	4063 (57%)	4351 (57%)	3827 (57%)	2696 (58%)	3023 (56%)	
< = 12	16,797 (25%)	1622 (23%)	1251 (22%)	1585 (22%)	1716 (22%)	1599 (24%)	1116 (24%)	1350 (25%)	
> 12	15,034 (23%)	1597 (22%)	1156 (21%)	1451 (20%)	1571 (21%)	1317 (20%)	861 (18%)	1072 (20%)	
Family history of breast cancer in the first degree, N (%)	5180 (8%)	559 (8%)	421 (7%)	570 (8%)	571 (7%)	506 (8%)	366 (8%)	385 (7%)	(0)

Values are mean (SD) unless otherwise specified

[a] Presents the physical activity level on the NOWAC 1–10 scale

data models, except for those in figures displaying HLI score change modelled with restricted cubic splines, where estimates from complete-case data models were described.

After adjusting for covariates, for every 1 SD increase in HLI score change, the HR was 0.93 (95% CI: 0.90–0.96) for lifestyle-related cancers, 0.96 (95% CI: 0.91–0.99) for alcohol-related cancers, 0.92 (95% CI: 0.88–0.96) for tobacco-related cancers, 0.94 (95% CI: 0.91–0.98) for obesity-related cancers, 0.90 (95% CI: 0.84–0.98) for reproductive-related cancers, 0.96 (95% CI: 0.91–1.01) for breast cancer, and 0.98 (95% CI: 0.90–1.07) for colorectal cancer (Table 2). When the HLI score change was modelled using restricted cubic splines, there were no indications of nonlinearity (all p-values > 0.05) in the adjusted associations for all lifestyle-related cancers (Fig. 2) nor for all other outcomes (Additional File 8).

Decreased HLI scores appeared to be statistically significant associated with an increased incidence of lifestyle-related cancer, while increased HLI scores were not (Fig. 2). These results were reflected in the categorical analysis of HLI score change, where decreases of three or more HLI units were associated with a HR of 1.16 (95%CI: 1.05–1.27) and increases of three or more HLI units were associated with a HR of 0.93 (95% CI: 0.84–1.03).

When individual lifestyle factors were excluded one by one, most associations changed less than 5% compared to associations with the HLI including all five lifestyle factors, with some exceptions (Table 2). For tobacco-related cancer incidence, the HR increased by 6.5% to 0.98 (95% CI: 0.94–1.03) when smoking was removed from the HLI and decreased by 5.4% to 0.87 (95% CI: 0.83–0.91) when BMI was removed from the HLI. For reproductive-related cancer incidence, the HR increased by 8.9% to 0.98 (95% CI: 0.90–1.06) when BMI was removed from the HLI and increased by 5.6% to 0.95 (95% CI: 0.88–1.02) when physical activity was removed from the HLI.

In the analysis of individual HLI factors, the HR for physical activity score change (per 1 unit increase) was 0.96 (95% CI: 0.94–0.98) for lifestyle-related cancer incidence and 0.94 (95% CI: 0.89–0.99) for reproductive-related cancer incidence. The HR for BMI score change (per 1 unit increase) was 0.96 (95% CI: 0.92–0.99) for obesity-related cancers and 0.86 (95% CI: 0.79–0.94) for reproductive-related cancers. The HR for smoking score change (per 1 unit increase) was 0.94 (95% CI: 0.88–1.00) for tobacco-related cancer incidence. The HR for alcohol score change (per 1 unit increase) was 0.94 (95% CI: 0.89–0.99) for alcohol-related cancers and 0.94 (95% CI: 0.88–1.00) for breast cancer incidence.

Tests for interaction for age at Q1, age at Q2, and Q1 HLI score category with HLI score change were not significant in adjusted models (all p-values > 0.05). When stratified on enrolment year, there was less than 5% change in estimates compared to the estimate obtained in the main analysis. The estimate for lifestyle-related cancer incidence was unchanged when excluding the first two years of follow-up. There was a null association observed between HLI score change and non-lifestyle-related cancer incidence (HR: 1.02; 95% CI: 0.96, 1.09).

Table 2 Associations between healthy lifestyle index score change and lifestyle-related, alcohol-related, tobacco-related, obesity-related, reproductive-related, breast, and colorectal cancer incidence in the Norwegian Women and Cancer Study ($n = 66{,}233$), imputed analysis

		Lifestyle-related cancer incidence[a]	Alcohol-related cancer incidence[a]	Tobacco-related cancer incidence	Obesity-related cancer incidence[a]	Reproductive-related cancer incidence[a]	Breast cancer incidence[a]	Colorectal cancer incidence
Cases		6354	3512	2931	4788	3385	2384	839
Continuous HLI score change	1-SD (2.6 HLI points) increase	0.93(0.90–0.96)	0.96(0.91–0.99)	0.92(0.88–0.96)	0.94(0.91–0.98)	0.90(0.84–0.98)	0.96(0.91–1.01)	0.98(0.90–1.07)
Categorical HLI score change	< = -3	1.16(1.05–1.27)	1.06(0.94–1.20)	1.27(1.10–1.45)	1.10(0.99–1.22)	1.21(0.96–1.54)	1.01(0.86–1.18)	1.23(0.94–1.61)
	-2	1.10(0.99–1.23)	1.10(0.96–1.26)	1.13(0.97–1.31)	1.09(0.97–1.22)	1.14(0.88–1.48)	1.09(0.92–1.29)	1.07(0.78–1.44)
	-1	1.03(0.93–1.13)	1.02(0.90–1.15)	1.08(0.94–1.25)	1.00(0.90–1.11)	1.04(0.81–1.33)	0.99(0.85–1.16)	1.10(0.84–1.45)
	0							
	1	0.99(0.89–1.09)	0.97(0.85–1.10)	1.06(0.92–1.23)	0.97(0.87–1.08)	0.93(0.72–1.20)	0.93(0.79–1.05)	1.09(0.82–1.44)
	2	0.96(0.86–1.07)	0.98(0.85–1.12)	1.02(0.87–1.18)	0.96(0.86–1.08)	0.88(0.66–1.17)	0.95(0.80–1.13)	1.10(0.81–1.49)
	> = 3	0.93(0.84–1.03)	0.92–0.81–1.05)	0.98(0.85–1.14)	0.92(0.82–1.03)	1.00(0.78–1.29)	0.89(0.75–1.05)	1.13(0.85–1.50)
HLI score change excluding one factor[b]	1-SD increase							
Excluding physical activity	2.0	0.95(0.93–0.98)	0.97(0.93–1.00)	0.94(0.91–0.98)	0.96(0.93–0.99)	0.95(0.88–1.02)	0.97(0.93–1.02)	0.98(0.90–1.06)
Excluding BMI	2.4	0.92(0.89–0.95)	0.95(0.91–0.99)	0.87(0.83–0.91)	0.96(0.93–0.99)	0.98(0.90–1.06)	0.96(0.91–1.01)	0.97(0.88–1.06)
Excluding smoking	2.5	0.95(0.92–0.98)	0.96(0.92–1.00)	0.98(0.94–1.03)	0.94(0.91–0.97)	0.87(0.81–0.94)	0.95(0.91–1.00)	1.00(0.92–1.09)
Excluding alcohol	2.5	0.93(0.90–0.96)	0.97(0.93–1.01)	0.92(0.88–0.96)	0.95(0.92–0.98)	0.89(0.82–0.96)	0.97(0.93–1.02)	0.99(0.91–1.08)
Excluding diet	2.0	0.93(0.90–0.96)	0.96(0.92–1.00)	0.93(0.89–0.97)	0.94(0.91–0.98)	0.90(0.84–0.97)	0.95(0.91–1.00)	0.98(0.90–1.07)
Single HLI factors[c]	1-unit increase (score 0–4)							
Physical activity score change		0.96(0.94–0.98)	0.98(0.95–1.01)	0.97(0.93–1.00)	0.98(0.95–1.00)	0.94(0.89–0.99)	0.97(0.94–1.01)	1.01(0.95–1.08)
BMI score change		0.98(0.95–1.02)	0.99(0.95–1.03)	1.04(0.99–1.09)	0.96(0.92–0.99)	0.86(0.79–0.94)	0.97(0.92–1.03)	1.03(0.93–1.13)
Smoking score change		0.98(0.93–1.03)	1.02(0.95–1.09)	0.94(0.88–1.00)	1.02(0.96–1.07)	1.08(0.95–1.24)	1.02(0.94–1.11)	0.97(0.84–1.12)
Alcohol score change		0.98(0.95–1.02)	0.94(0.89–0.99)	1.00(0.94–1.06)	0.97(0.93–1.02)	1.06(0.96–1.18)	0.94(0.88–1.00)	0.97(0.86–1.08)
Diet score change		0.99(0.97–1.01)	0.99(0.97–1.02)	0.99(0.97–1.02)	1.00(0.98–1.02)	1.00(0.95–1.05)	1.00(0.96–1.03)	1.00(0.95–1.06)

All models were adjusted for education (years), height (centimetres), HLI score at Q1 (continuous), and calendar year at Q2 (continuous)

[a] Models additionally adjusted for age at menarche (years), menopausal status (premenopausal/postmenopausal), breastfeeding (cumulative months 0, < = 12, > 12), hormone replacement therapy use (never/former/current), oral contraceptive use (never/ever), parity (0, 1–2, > 2), and history of breast cancer in a first degree relative (yes/no)

[b] Baseline HLI score was adjusted by separately adjusting for HLI score at Q1 excluding the factor in question and the individual factor score at Q1

[c] Mutually adjusted for all single factor HLI score changes and single factor HLI scores at Q1

Alcohol-related cancers including sites: upper aerodigestive [C01-C10], pharynx [C11-C14], esophagus [C15], colorectum [C18-C20], liver [C22-C24], larynx [C32], breast [C50],

Tobacco-related cancers including sites: upper aerodigestive [C01-C10], pharynx [C11-C14], esophagus [C15], stomach [C16], colorectum [C18-C20], liver [C22-C24], pancreas [C25], accessory sinus [C31], larynx [C32], trachea [C33], lung [C34], breast [C50], cervix [C53], ovarian [C56], kidney [C64-C66], bladder [C67], acute myeloid leukemia [C92]

Obesity-related cancers including sites: esophagus [C15], stomach [C16], colorectum [C18-C20], liver [C22-C24], pancreas [C25], breast [C50], uterine [C54-C55], ovarian [C56], kidney [C64-C66], thyroid [C73], multiple myeloma [C90],

Reproductive-related cancers including sites: vulva [C51] vagina [C52], cervix [C53], uterine [C54-C55], ovarian [C56], other female genital organs [C57-C58]

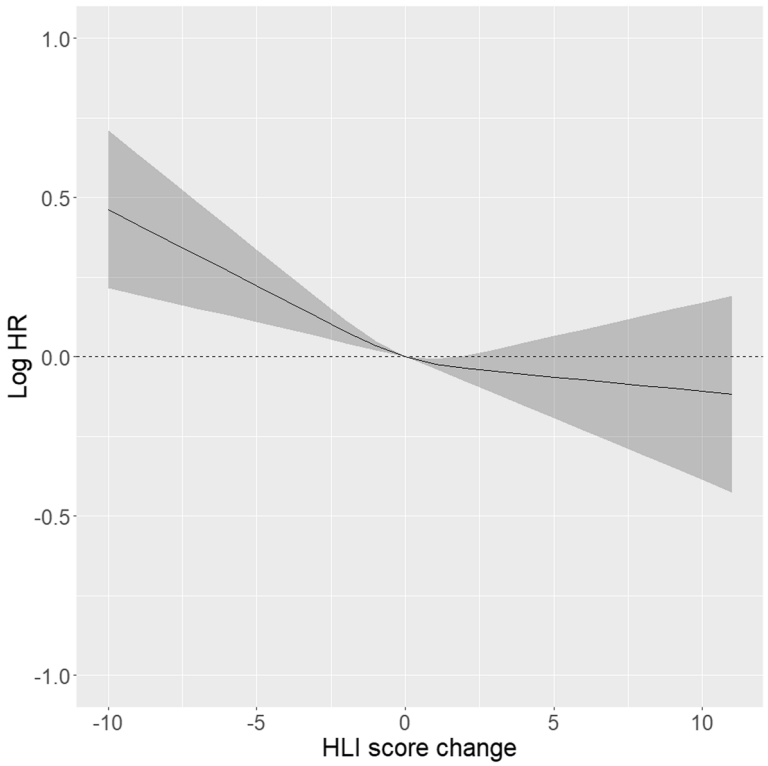


Fig. 2 Association between HLI score change modelled using restricted cubic splines and lifestyle-related cancer incidence

Discussion

In this study, lifestyle change was assessed by evaluating healthy lifestyle index scores at two timepoints that were on average 7 years apart for 66 233 Norwegian women in the period 1996–2014. Most participants did not report major lifestyle differences between baseline and follow-up, where approximately 38% of participants registered an HLI score improvement. We observed that lifestyle change equivalent to a 1 SD increase in HLI score change was associated with 7% lower incidence for lifestyle-related cancers, 4% lower incidence for alcohol-related cancers, 8% lower incidence for tobacco-related cancer, 6% lower incidence for obesity-related cancers, and 10% lower incidence for reproductive-related cancers. There was a 4% reduced incidence of breast cancer, although the 95% CI for the HR was 0.93 to 1.01. We did not observe an association between HLI score changes and colorectal cancer incidence. When evaluated as group comparisons, major lifestyle worsening corresponding to a decline of three or more HLI score points from baseline to follow-up compared to no HLI score change was associated with a 16% higher risk of overall lifestyle-related cancer. Lifestyle improvement of three or more HLI score points was associated with a 7% lower risk of lifestyle-related cancer, although a null association was also compatible with our data.

In general, lifestyle worsening was both more strongly and more likely associated with the incidence of total lifestyle-related cancers, tobacco-related cancers, obesity-related cancers, and reproductive-related cancers compared to lifestyle improvement. We observed this from results modelling HLI score change as a continuous measure using restricted cubic splines and from modelling HLI score change as group comparisons. However, since there were no clear indications of nonlinearity from the restricted cubic spline models according to visual inspection, this suggests that the linear estimates are robust. Additionally, although we observed the strongest associations for lifestyle worsening, we cannot assert with any confidence that lifestyle improvement is not related to reduced cancer incidence considering the lack of published studies assessing the effects of changes in lifestyle factors in combination.

There are a small number of published studies investigating the effect of changes in lifestyle behaviours combined, as single factors or overall, on cancer incidence. In a study conducted on a large cohort of Swedish women, Botteri et al. [17] observed that those who either improved their lifestyle or maintained their lifestyle had a reduced risk of lifestyle-related cancers compared to those who had consistently poor lifestyle [17]. However, as diet was not included in their HLI and their

assessment of lifestyle factor scores was different, more detailed comparison is challenging. A controlled intervention of lifestyle among men at risk for coronary heart disease in Norway observed a 32% risk reduction after 25 years of follow-up [19]. The selected sample and controlled design likely accounted for their stronger estimates compared to ours.

No single lifestyle factor was indicated as solely responsible for the HLI score change associations we observed. Therefore, in combination, changes in physical activity level, BMI, smoking habits, alcohol intake, and diet were related to cancer incidence. In the present study, physical activity score change was the only factor to demonstrate a clear association with lifestyle-related cancer incidence in the single factor analysis. Increasing physical activity level has previously been related to lower cancer incidence among mid-life adults, although the sample was limited to Norwegian men [12]. Oyeyemi et al. [13] observed that, in NOWAC, physical activity level increase was associated with lower colon cancer risk, but not for colorectal cancer, which is consistent with our results. Further, only stable high physical activity levels were associated with lower colon cancer incidence in a large US cohort [28]. Consistent with observations from the Norwegian-Swedish Women's Lifestyle and Health cohort, we did not observe an association between physical activity level change and breast cancer [29].

Several studies on BMI change – often equated with changes in weight – have identified that weight loss is associated with lower cancer risk [30–33]. However, weight loss has not been shown to influence cancer risk to the same degree or level of certainty as weight gain [7, 32, 34, 35]. Considering that we have identified BMI as an important contributor to the association between continuous HLI score and cancer incidence, the weak associations we observed between lifestyle improvement and lower cancer incidence are consistent with the literature on weight change. Unintentional weight loss as a prediagnostic symptom of cancer has been suggested as an explanation for the little to no risk reduction observed among those who lost weight. While the present study did not observe a difference in estimates after conducting sensitivity analysis that excluded the first two years of follow-up, it is possible that unintentional weight loss due to morbid conditions, including cancer, can emerge earlier than two years before diagnosis.

The benefits of smoking cessation for lung cancer [6], head and neck cancer [14], and oesophageal squamous cell carcinoma [36] risk reduction have been widely documented, and are consistent with our results.

We observed that alcohol was an important contributor to the HLI score change associations. Assessed as a single factor, increase in alcohol score change was associated with 6% lower incidence of alcohol-related cancers and breast cancer. A strong positive association between 5-year alcohol consumption increase and breast cancer risk, but not for alcohol reduction was observed among postmenopausal Danish women [18]. This supports our continuous estimate and could add weight to the potential lack of association between overall lifestyle improvement and lower cancer incidence we observed. Unlike our observations, there were no observed associations for alcohol change and incidence of alcohol-related cancers or breast cancer in EPIC [17]. Alcohol cessation has been associated with the lower risks of laryngeal, pharyngeal, and oesophageal cancers [14, 16] supporting our results for alcohol-related cancer. Our observations support the recommendation to reduce alcohol intake for the prevention of several types of cancer.

Diet had the least influence on lifestyle-related cancer incidence compared to other lifestyle factors, given almost unchanged estimates when it was removed from the index and markedly null estimates from the single factor analysis. To our knowledge, studies on dietary change and cancer risk at the individual level do not exist to provide comparison. However, this result is plausible given the lack of convincing evidence between some food groups included in the HLI and cancer incidence as summarised by the WCRF/AICR Continuous Update Project in 2018 [4].

We investigated colorectal cancer incidence as a specific outcome due to its exceptionally high incidence among Norwegian women compared to that of neighbouring and high-income countries [37]. Our study did not observe an association between lifestyle changes, in combination or among individual lifestyle factors, and colorectal cancer incidence. However, in general, the presence of strong and convincing associations between measured risk factors, whether at baseline or at multiple timepoints, and colorectal cancer continue to elude large population-based cohort studies in the Norwegian population [7, 13, 38, 39]. Nevertheless, HLI at baseline and colorectal cancer risk were inversely associated among women in NOWAC [21] and EPIC [40]. This may indicate that, in terms of lifestyle, healthy habits lived from the beginning of adulthood are most important for reducing colorectal cancer risk and/or that the true strength of association is so small that models are underpowered.

Lifestyle changes occurring among Norwegian women in their middle adult years during the period 1996 to 2014 was likely driven by several phenomena, including changes that occurred due to societal shifts in attitudes and availabilities as well as intentional or unintentional individual change. On average, NOWAC women reported increasing physical activity levels, increasing weight, reducing smoking, increasing alcohol intake,

and negligible dietary changes on the HLI from baseline until follow-up. Considering this population in its context, we would expect smoking habits to be reduced given increasing tobacco restrictions through the 1990s and 2000s [41]. Weight increase with age, specifically in adult years, is a universal occurrence. Further, national trends have shown that alcohol intake habits among young Norwegian women have been increasing over the past half-century, thus impacting their habits later in adulthood [42]. Due to this, we would expect birth cohorts to undergo systematically different lifestyle changes and for risk to possibly manifest differentially. However, we did not observe different risk estimates for continuous or categorical models between subgroups recruited early or late in the sampling time, despite the wide variation in age, time between baseline and follow-up, and calendar years. This increases our confidence that our estimates reflect risk differences largely attributable to HLI score change.

The findings from our study have major public health relevance. In this study, we provide evidence that overall lifestyle changes among cancer-free women between the ages of 41 and 76 impact the incidence of many cancer types. Importantly, the umbrella grouping of lifestyle-related cancer covers nearly all the most frequent cancers currently diagnosed among adult Norwegian women, including cancers of the breast, lung, colon, and endometrium. To-date, risk differences for lifestyle change have seldom been assessed but are key to making informed policy decisions for how cancer can be prevented in the already adult segment of the population. The importance of having a healthy baseline lifestyle is undeniable. However, our observations indicate that lifestyle changes over a period of five years during adulthood do impact cancer risk, regardless of baseline lifestyle. Further, our results emphasize the importance of avoiding lifestyle worsening. Considering that most Norwegian women in our cohort experienced negative HLI score changes, and thus lifestyle worsening, maintenance of lifestyle should be on the public health agenda.

Strengths

The minimalism of the HLI enables a broader assessment of lifestyle and an easy method for investigating lifestyle patterns and interaction between single factors. The use of this simple, composite exposure seems to effectively capture an association between lifestyle change and cancer incidence. This supports the use of the HLI as a composite exposure in epidemiological studies given the public health aim to prevent the occurrence of cancer cases.

Additional strengths of this study include its large, nationally representative sample of women in Norway with comprehensive measurements of lifestyle factors and other important characteristics at two timepoints. This data has enabled us to undertake, for the first time, an assessment of the effect of overall lifestyle changes – including physical activity level, BMI, smoking, alcohol, and diet – on cancer incidence. Linkage of participants to the national registries were instrumental in ensuring the follow-up of participants, including cancer case ascertainment, death, and emigration.

Limitations

There were limitations to the measurement of lifestyle change as a numeric difference between the HLI score measured at two timepoints. Firstly, the data does not inform when the lifestyle change(s) took place beyond recognition of net change between baseline and follow-up. Due to the long latency period of cancers, it is logical that changes occurring closer to baseline, and hence at a younger age, had a greater effect on the outcome compared to changes occurring closer to follow-up, or older age. Not being able to account for these differences likely biased our results to the null. Secondly, changes representing an increase in HLI score in one lifestyle factor concurrent with a decrease in HLI score in another would manifest as a major lifestyle change for the individual, but as a net zero HLI score change. A real example is the known weight gain that follows smoking reduction. Indeed, we observed that weight loss was associated with a higher incidence of tobacco-related cancer. It is therefore possible that our estimates were attenuated in such situations given that both changes are unlikely to represent the same risk compared to no change.

Recall bias is a concern when data is self-reported as it can lead to misclassification error. In NOWAC, height tends to be overestimated and weight tends to be underestimated among participants with overweight and, to a greater extent, obesity [43]. We expect misclassification to have occurred non-differentially across cases and non-cases, thus likely only attenuating rather than biasing our estimates. Under-reporting of unhealthy foods and alcohol has been confirmed in the FFQ used by NOWAC [44]. However, the ranking of individuals' intake was deemed adequate and the relative validity of the FFQ was observed to be in the same range as observed in other EPIC cohorts [44]. In addition, the FFQs were not identical at baseline and follow-up due to the addition of some food items to the follow-up FFQ that had become relevant for the Norwegian diet after baseline [45]. Although adjustment for energy intake by means of nutrient densities to calculate the diet score accounted for some of these differences, dietary change was likely underestimated. We cannot exclude the presence of residual confounding

bias in our risk estimates despite the adjustment of several risk factors. In addition, follow-up time may not have been long enough for the effects of lifestyle changes on cancer development/prevention to accrue.

Conclusions

This study supports lifestyle intervention as cancer preventive action in the already adult segment of the population. We provide evidence that overall lifestyle changes among cancer-free women between the ages of 41 and 76 impact the incidence of many cancer types. There was a negative dose–response relationship between magnitude of positive lifestyle change and the incidence of overall lifestyle-related cancers, as well as alcohol-related, tobacco-related, obesity-related, and reproductive-related cancers. We observed that underlying this trend was an especially clear association between lifestyle worsening and increased risk compared to stable lifestyle. The prevention of lifestyle worsening, maintenance of healthy lifestyle, and lifestyle improvement, belong on the public health agenda if the predicted trajectory of cancer incidence is to be dismantled.

Abbreviations

BMI	Body mass index
EPIC	European prospective investigation into cancer and nutrition
FFQ	Food frequency questionnaire
HLI	Healthy lifestyle index
ICD-10	International Classification of Disease, Version 10
MI	Multiple imputation
NOWAC	Norwegian Women and Cancer Study
Q1	Questionnaire 1 (baseline measurement)
Q2	Questionnaire 2 (follow-up measurement)

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-023-15476-3.

Additional file 1. Questionnaire 1 (baseline), Norwegian Women and Cancer Study (NOWAC).

Additional file 2. Questionnaire 2 (follow-up), Norwegian Women and Cancer Study (NOWAC).

Additional file 3. Sample flowchart, Norwegian Women and Cancer Study (NOWAC).

Additional file 4. Description of healthy lifestyle index (HLI) construction in the Norwegian Women and Cancer Study (NOWAC).

Additional file 5. Cancer types and associated International Classification of Disease, tenth revision (ICD-10) codes included in the study.

Additional file 6. Distribution of HLI score change across baseline HLI score categories.

Additional file 7. Associations between healthy lifestyle index score change and lifestyle-related, alcohol-related, tobacco-related, obesity-related, reproductive-related, breast, and colorectal cancer incidence in the Norwegian Women and Cancer Study ($n = 44404$), complete-case analysis.

Additional file 8. Associations between HLI score change modelled with restricted cubic splines and several cancer subgroupings.

Acknowledgements

We would like to thank all the NOWAC participants for their irreplaceable contribution to this study. Additional thanks are extended to Komodo Matta, for their discussions, and Arne Bastian Wiik, for their assistance with obtaining data.

Disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer /World Health Organization.

Authors' contributions

All authors have read and approved this manuscript, and agreed to be accountable for the contents of this manuscript. SLFC, KBB, THN, TBB, and TMS were responsible for the conception and design of the work. SLFC was responsible for the analysis of data and writing the manuscript. SLFC, KBB, THN, TMS, TBB, EB, and PF were involved in the interpretation of data and substantially revised the manuscript.

Funding

Open access funding provided by UiT The Arctic University of Norway (incl University Hospital of North Norway). Funding for open access publication and the resources, including personnel, required to design, conduct analyses, and write the manuscript was provided by UiT The Arctic University of Norway. However, UiT did not take part in the conceptualization, design, or reporting of the study.

Availability of data and materials

The datasets analysed during the current study are not publicly available due to local and national ethical and security policies but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The NOWAC study has been given appropriate approval for collection and handling of questionnaire data by the Regional Committee for Medical and Health Research Ethics (REK) for Northern Norway and the Norwegian Data Inspectorate. All participants provided informed consent.

Consent for publication

Not applicable.

Competing interests

The authors have no competing interests to declare.

Author details

[1]Department of Community Medicine, UiT The Arctic University of Norway, Hansine Hansens Veg 18, 9019 Tromsø, Norway. [2]Department of Research, Cancer Registry of Norway, Ullernchauseen 64, 0379 Oslo, Norway. [3]Section for Colorectal Cancer Screening, Cancer Registry of Norway, Ullernchauseen 64, 0379 Oslo, Norway. [4]Nutrition and Metabolism Branch, International Agency for Research On Cancer, World Health Organization, World Health Organization, 150 Cours Albert Thomas, 69372 CEDEX 08 Lyon, France.

Received: 7 July 2022 Accepted: 16 March 2023
Published online: 03 April 2023

References

1. Bray F, Laversanne M, Weiderpass E, Soerjomataram I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. Cancer. 2021;127(16):3029–30.
2. 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 Cancer J Clin. 2021;71(3):209–49.

3. Vineis P, Wild CP. Global cancer patterns: causes and prevention. The Lancet. 2014;383(9916):549–57.
4. World Cancer Research Fund/American Institute for Cancer Research: Diet, Nutrition, Physical Activity and Cancer: a Global Perspective. In: Continuous Update Project Expert Report. 2018.
5. Bosetti C, Gallus S, Peto R, Negri E, Talamini R, Tavani A, Franceschi S, La Vecchia C. Tobacco smoking, smoking cessation, and cumulative risk of upper aerodigestive tract cancers. Am J Epidemiol. 2008;167(4):468–73.
6. Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. BMJ. 2000;321(7257):323.
7. da Silva M, Weiderpass E, Licaj I, Lissner L, Rylander C. Excess body weight, weight gain and obesity-related cancer risk in women in Norway: the Norwegian Women and Cancer Study. Br J Cancer. 2018;119(5):646–56.
8. Keum N, Greenwood DC, Lee DH, Kim R, Aune D, Ju W, Hu FB, Giovannucci EL. Adult weight gain and adiposity-related cancers: a dose-response meta-analysis of prospective observational studies. J Natl Cancer Inst. 2015;107(2):djv088.
9. Christakoudi S, Pagoni P, Ferrari P, Cross AJ, Tzoulaki I, Muller DC, Weiderpass E, Freisling H, Murphy N, Dossus L, et al. Weight change in middle adulthood and risk of cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Int J Cancer. 2021;148(7):1637–51.
10. Bracci PM. Obesity and pancreatic cancer: overview of epidemiologic evidence and biologic mechanisms. Mol Carcinog. 2012;51(1):53–63.
11. Liu Y, Warren Andersen S, Wen W, Gao YT, Lan Q, Rothman N, Ji BT, Yang G, Xiang YB, Shu XO, et al. Prospective cohort study of general and central obesity, weight change trajectory and risk of major cancers among Chinese women. Int J Cancer. 2016;139(7):1461–70.
12. Robsahm TE, Heir T, Sandvik L, Prestgaard E, Tretli S, Erikssen JE, Falk RS. Changes in midlife fitness, body mass index, and smoking influence cancer incidence and mortality: a prospective cohort study in men. Cancer Med. 2019;8:4875.
13. Oyeyemi SO, Braaten T, Licaj I, Lund E, BenjaminsenBorch K. Physical activity patterns and the risk of colorectal cancer in the Norwegian Women and Cancer Study: a population-based prospective study. BMC Cancer. 2018;18(1):1216.
14. Marron M, Boffetta P, Zhang ZF, Zaridze D, Wünsch-Filho V, Winn DM, Wei Q, Talamini R, Szeszenia-Dabrowska N, Sturgis EM, et al. Cessation of alcohol drinking, tobacco smoking and the reversal of head and neck cancer risk. Int J Epidemiol. 2010;39(1):182–96.
15. Heckley GA, Jarl J, Asamoah BO. G-Gerdtham U: How the risk of liver cancer changes after alcohol cessation: a review and meta-analysis of the current literature. BMC Cancer. 2011;11(1):446.
16. Ahmad Kiadaliri A, Jarl J, Gavriilidis G, Gerdtham UG. Alcohol drinking cessation and the risk of laryngeal and pharyngeal cancers: a systematic review and meta-analysis. PLoS ONE. 2013;8(3): e58158.
17. Botteri E, Berstad P, Sandin S, Weiderpass E. Lifestyle changes and risk of cancer: experience from the Swedish women's lifestyle and health cohort study. Acta Oncol. 2021;60:1–8.
18. Dam MK, Hvidtfeldt UA, Tjønneland A, Overvad K, Grønbæk M, Tolstrup JS. Five year change in alcohol intake and risk of breast cancer and coronary heart disease among postmenopausal women: prospective cohort study. BMJ. 2016;353: i2314.
19 Botteri E, de Lange T, Tonstad S, Berstad P. Exploring the effect of a lifestyle intervention on cancer risk: 43-year follow-up of the randomized Oslo diet and antismoking study. J Int Med. 2018;284:282.
20. Lund E, Dumeaux V, Braaten T, Hjartåker A, Engeset D, Skeie G, Kumle M. Cohort profile: The Norwegian Women and Cancer Study—NOWAC—Kvinner og kreft. Int J Epidemiol. 2007;37(1):36–41.
21. Chen SLF, Braaten T, Borch KB, Ferrari P, Sandanger TM, Nøst TH. Combined lifestyle behaviors and the incidence of common cancer types in the Norwegian Women and Cancer Study (NOWAC). Clin Epidemiol. 2021;13:721–34.
22. International Agency for Research on Cancer (IARC). Human cancer: known causes and prevention by Organ Site. 2022 [updated 01/09/2021]. Available from: https://monographs.iarc.who.int/human_cancer_known_causes_and_prevention_organ_site/. Accessed 06 June 2022.
23. Harrell FE. Regression modeling strategies. 2nd ed. Cham, Switzerland: Springer International Publishing; 2015.
24. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med. 2011;30(4):377–99.
25. van Buuren S. Flexible imputation of missing data. Boca Raton: Taylor & Francis Group; 2018. Available from: https://stefvanbuuren.name/fimd/.
26. Rubin DB. 1.5. Multiple imputation. In: Multiple imputation for nonresponse in surveys. New York: Wiley; 1987. p. 15–8.
27. Posit team: RStudio: Integrated Development Environment for R. In. Boston, MA: Posit Software, PBC; 2022.
28. Wolin KY, Patel AV, Campbell PT, Jacobs EJ, McCullough ML, Colditz GA, Gapstur SM. Change in physical activity and colon cancer incidence and mortality. Cancer Epidemiol Biomarkers Prev. 2010;19(12):3000–4.
29. Margolis KL, Mucci L, Braaten T, Kumle M, TrolleLagerros Y, Adami HO, Lund E, Weiderpass E. Physical activity in different periods of life and the risk of breast cancer: the Norwegian-Swedish Women's lifestyle and health cohort study. Cancer Epidemiol Biomarkers Prev. 2005;14(1):27–32.
30. Parker ED, Folsom AR. Intentional weight loss and incidence of obesity-related cancers: the Iowa women's health study. Int J Obes Relat Metab Disord. 2003;27(12):1447–52.
31. Eliassen AH, Colditz GA, Rosner B, Willett WC, Hankinson SE. Adult weight change and risk of postmenopausal breast cancer. JAMA. 2006;296(2):193–201.
32. Radimer KL, Ballard-Barbash R, Miller JS, Fay MP, Schatzkin A, Troiano R, Kreger BE, Splansky GL. Weight change and the risk of late-onset breast cancer in the original Framingham cohort. Nutr Cancer. 2004;49(1):7–13.
33. Kawai M, Minami Y, Kuriyama S, Kakizaki M, Kakugawa Y, Nishino Y, Ishida T, Fukao A, Tsuji I, Ohuchi N. Adiposity, adult weight change and breast cancer risk in postmenopausal Japanese women: the Miyagi cohort study. B J Cancer. 2010;103(9):1443–7.
34. Teras LR, Goodman M, Patel AV, Diver WR, Flanders WD, Feigelson HS. Weight loss and postmenopausal breast cancer in a prospective cohort of overweight and obese US women. Cancer Causes Control : CCC. 2011;22(4):573–9.
35. Emaus MJ, van Gils CH, Bakker MF, Bisschop CN, Monninkhof EM, Bueno-de-Mesquita HB, Travier N, Berentzen TL, Overvad K, Tjønneland A, et al. Weight change in middle adulthood and breast cancer risk in the EPIC-PANACEA study. Int J Cancer. 2014;135(12):2887–99.
36. Wang QL, Xie SH, Li WT, Lagergren J. Smoking cessation and risk of esophageal cancer by histological type: systematic review and meta-analysis. J Natl Cancer Inst. 2017;109(12):1–10.
37. ECIS - European Cancer Information System. Estimates of cancer incidence and mortality in 2020, for all countries: European Union; 2022. Available from: https://ecis.jrc.ec.europa.eu. Accessed 23 Feb 2022.
38. Oyeyemi SO, Braaten T, Botteri E, Berstad P, Borch KB. Exploring geographical differences in the incidence of colorectal cancer in the Norwegian Women and Cancer Study: a population-based prospective study. Clin Epidemiol. 2019;11:669–82.
39. Bakken T, Braaten T, Olsen A, Kyro C, Lund E, Skeie G. Consumption of whole-grain bread and risk of colorectal cancer among Norwegian Women (the NOWAC Study). Nutrients. 2016;8(1):40.
40. Aleksandrova K, Pischon T, Jenab M, Bueno-de-Mesquita HB, Fedirko V, Norat T, Romaguera D, Knuppel S, Boutron-Ruault MC, Dossus L, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC Med. 2014;12:168.
41. Norwegian Directorate of Health. Statistikk og historikk om røyking, snus og e-sigaretter [Statistics and history for smoking, snus, and e-cigarettes]. Oslo: Norwegian Directorate of Health; 2019 [updated 18 January 2022]. Available from: https://www.helsedirektoratet.no/tema/tobakk-royk-og-snus/statistikk-om-royking-bruk-av-snus-og-e-sigaretter. Accessed 11 Feb 2022.
42. Vedøy TF, Skretting A: Bruk av alkohol blant kvinner [Alcohol use among women]. In., vol. 4. Oslo: State Institute for Substance Abuse Research (SIRUS); 2009.
43. Skeie G, Mode N, Henningsen M, Borch KB. Validity of self-reported body mass index among middle-aged participants in the Norwegian Women and Cancer Study. Clin Epidemiol. 2015;7:313–23.
44. Hjartåker A, Andersen LF, Lund E. Comparison of diet measures from a food-frequency questionnaire with measures from repeated 24-hour dietary recalls. The Norwegian Women and Cancer Study. Public Health Nutr. 2007;10(10):1094–103.
45. Skeie G, Hjartåker A, Braaten T, Lund E. Dietary change among breast and colorectal cancer survivors and cancer-free women in the Norwegian Women and Cancer cohort study. Cancer Causes Control : CCC. 2009;20(10):1955–66.

Publisher's Note

KVINNER OG KREFT

Høst 2003

KONFIDENSIELT

Hvis du samtykker i å være med, sett kryss for JA i ruten ved siden av. Dersom du ikke ønsker å delta kan du unngå purring ved å sette kryss for NEI og returnere skjemaet i vedlagte svarkonvolutt.
Vi ber deg fylle ut spørreskjemaet så nøye som mulig.

Skjemaet skal leses optisk. Vennligst bruk blå eller sort penn.
Du kan ikke bruke komma, bruk blokkbokstaver.

Med vennlig hilsen
Eiliv Lund
Professor dr. med

Jeg samtykker i å delta i spørreskjemaundersøkelsen JA ☐ NEI ☐

Forhold i oppveksten

I hvilken kommune har du bodd lengre enn ett år?

	Kommune:		Alder	
1. Fødested:		Fra	år til	år
2		Fra	år til	år
3		Fra	år til	år
4		Fra	år til	år
5		Fra	år til	år
6		Fra	år til	år
7		Fra	år til	år

Kroppstype i 1. klasse. (Sett ett kryss)

☐veldig tynn ☐tynn ☐normal ☐tykk ☐veldig tykk

Menstruasjonsforhold

Hvor gammel var du da du fikk menstruasjon første gang?

Hvor mange år tok det før menstruasjonen ble regelmessig?

☐ Ett år eller mindre ☐ Mer enn ett år
☐ Aldri ☐ Husker ikke

Har du regelmessig menstruasjon fremdeles?

☐ Ja ☐ Har uregelmessig menstruasjon
☐ Vet ikke (menstruasjon uteblitt pga. sykdom o.l.)
☐ Bruk av hormonpreparat med østrogen
☐ Nei

Hvis Nei;

har den stoppet av seg selv? ☐
operert vekk eggstokkene? ☐
operert vekk livmoren? ☐
annet? ☐

Alder da menstruasjonen opphørte?

Graviditeter, fødsler og amming

Har du noen gang vært gravid? Ja ☐ Nei ☐

Hvis Ja; fyll ut for hvert barn du har født opplysninger om fødselsår og antall måneder du ammet (fylles også ut for dødfødte eller for barn som er døde senere i livet). Dersom du ikke har født barn, fortsetter du ved neste spørsmål.

Barn	Fødselsår	Antall måneder med amming	Barn	Fødselsår	Antall måneder med amming
1			5		
2			6		
3			7		
4			8		

Bruk av hormonpreparater med østrogen i overgangsalderen

Har du noen gang brukt østrogentabletter/plaster? Ja ☐ Nei ☐

Hvis Ja; hvor mange år har du brukt østrogentabletter/plaster i alt?

Hvor gammel var du første gang du brukte østrogentabletter/plaster?

Bruker du tabletter/plaster nå? Ja ☐ Nei ☐

Hvor pålitelig anser du kildene nedenfor å være når det gjelder informasjon om østrogenbehandling?

	Lite pålitelig	Pålitelig	Meget pålitelig	Vet ikke/ usikker
Allmenpraktiserende lege	☐	☐	☐	☐
Gynekolog	☐	☐	☐	☐
Apotek	☐	☐	☐	☐
Radio/TV	☐	☐	☐	☐
Ukeblader/aviser	☐	☐	☐	☐
Slekt/venninner	☐	☐	☐	☐

Bruker du soyapreparater mot plager i overgangsalderen? Ja ☐ Nei ☐

UTFYLLENDE SPØRSMÅL TIL ALLE SOM HAR BRUKT ELLER BRUKER PREPARATER MED ØSTROGEN I FORM AV TABLETTER ELLER PLASTER.

Hvis du har svart «nei» på spørsmålene om hormonbruk i overgangsalderen, kan du gå videre til spørsmålene under «**P-piller**». Har du svart «ja», ber vi deg om å utdype dette nærmere ved å svare på spørsmålene nedenfor. For hver periode med sammenhengende bruk av samme hormonpreparat håper vi du kan si oss hvor gammel du var da du startet, hvor lenge du brukte det samme hormonpreparatet og navnet på dette. Dersom du har tatt opphold eller skiftet merke, skal du besvare spørsmålene for en ny periode. Dersom du ikke husker navnet på hormonpreparatet sett «usikker». For å hjelpe deg til å huske navnet på hormonpreparatene ber vi deg bruke den vedlagte brosjyre som viser bilder av hormonpreparater som har vært solgt i Norge. Vennligst oppgi også nummer på hormontabletten/plasteret som står i brosjyren.

Periode	Alder ved start	Brukt samme hormontablett/plaster/ Sammenhengende år	måned	Nr.	Hormontablett/ plaster/ (se brosjyre) Navn
1.					
2.					
3.					
4.					
5.					

P-pillebruk

Har du brukt p-piller eller minipiller? ..... Ja ☐ Nei ☐

Bruker du p-piller nå? ..... Ja ☐ Nei ☐

For p-pillebruk ønsker vi å få vite navnet på p-pillen, årstallet du startet å bruke den og hvor lenge du brukte dette merket sammenhengende. Dersom du har hatt opphold eller skiftet merke start på ny linje. For å hjelpe deg å huske navnet ber vi deg bruke den vedlagte brosjyren. Vennligst oppgi nummeret på p-pillen.

Periode	Alder ved start	Brukt samme hormontablett/plaster/ Sammenhengende år	måned	Nr.	Hormontablett/ plaster/ (se brosjyre) Navn
1.					
2.					
3.					
4.					
5.					
6.					

Hormonspiral

Har du noen gang brukt hormonspiral (Levonova)? ..... Ja ☐ Nei ☐

Hvis Ja; hvor mange hele år har du brukt hormonspiral i alt? ..... ☐☐

Hvor gammel var du første gang du fikk innsatt hormonspiral? ☐☐

Bruker du hormonspiral nå? ..... Ja ☐ Nei ☐

Østrogenpreparat til lokal bruk i skjeden

Har du noen gang brukt østrogen-krem/stikkpille? ..... Ja ☐ Nei ☐

Hvis Ja; bruker du krem/stikkpille nå? ..... Ja ☐ Nei ☐

Andre legemidler

Bruker du noen av disse legemidlene daglig nå?

Fontex, Fluoxetin ..... Ja ☐ Nei ☐

Cipramil, Citalopram ..... Ja ☐ Nei ☐

Seroxat, Paroxetin ..... Ja ☐ Nei ☐

Zoloft ..... Ja ☐ Nei ☐

Fevarin ..... Ja ☐ Nei ☐

Cipralex ..... Ja ☐ Nei ☐

Hvis Ja; hvor lenge har du brukt dette legemidlet sammenhengede? Måneder ☐☐ År ☐☐

Har du benyttet noen av disse legemidlene tidligere? Ja ☐ Nei ☐

Hvis Ja; hvor lenge har du benyttet disse legemidlene i alt? År ☐☐

Sykdom

Har du eller har du hatt noen av følgende sykdommer?

	Ja	Nei	Hvis ja: Alder ved start
Kreft	☐	☐	☐☐
Høyt blodtrykk	☐	☐	☐☐
Hjertesvikt/hjertekrampe	☐	☐	☐☐
Hjerteinfarkt	☐	☐	☐☐
Slag	☐	☐	☐☐
Sukkersyke (diabetes)	☐	☐	☐☐
Depresjon (oppsøkt lege)	☐	☐	☐☐

Selvopplevd helse

Oppfatter du din egen helse som; (Sett ett kryss)

Meget god ☐ God ☐ Dårlig ☐ Meget dårlig ☐

Røykevaner

Har du i løpet av livet røykt mer enn 100 sigaretter til sammen? ..... Ja ☐ Nei ☐

Hvor gammel var du da du tok din første sigarett? ☐☐

Hvis Ja, ber vi deg om å fylle ut for hver aldersgruppe i livet hvor mange sigaretter du i gjennomsnitt røykte pr. dag i den perioden.

Antall sigaretter hver dag

Alder	0	1-4	5-9	10-14	15-19	20-24	25+
10-14	☐	☐	☐	☐	☐	☐	☐
15-19	☐	☐	☐	☐	☐	☐	☐
20-29	☐	☐	☐	☐	☐	☐	☐
30-39	☐	☐	☐	☐	☐	☐	☐
40-49	☐	☐	☐	☐	☐	☐	☐
50+	☐	☐	☐	☐	☐	☐	☐

	Ja	Nei
Røyker du daglig nå?	☐	☐
Røykte noen av dine foreldre når du var barn?	☐	☐

Hvis Ja, hvor mange sigaretter røykte de til sammen pr. dag? ☐☐

Brystkreft i nærmeste familie

Har noen nære slektninger hatt brystkreft?

	Ja	Nei	Vet ikke	Alder ved start
Datter	☐	☐	☐	☐☐
Mor	☐	☐	☐	☐☐
Søster	☐	☐	☐	☐☐

Mammografiundersøkelse

Har du vært til undersøkelse av brystene med mammografi ..... Ja ☐ Nei ☐

Hvis Ja;

hvor gammel var du første gangen? (hele år) ..... ☐☐

Hvor mange ganger har du vært undersøkt?

-etter invitasjon fra Mammografiprogrammet ..... ☐☐

-etter henvisning fra lege ..... ☐☐

-uten henvisning fra lege ..... ☐☐

Har du silikoninnlegg i brystene? Ja ☐ Nei ☐

Hvis Ja;

hvor mange år har du hatt det? ..... ☐☐

Har du hatt silikoninnlegg tidligere? Ja ☐ Nei ☐

Hvis Ja;

hvorfor fjernet du innlegget? ☐

☐

Fysisk aktivitet

Vi ber deg angi din fysiske aktivitet etter en skala fra svært lite til svært mye. Skalaen nedenfor går fra 1-10. Med fysisk aktivitet mener vi både arbeid i hjemmet og i yrkeslivet, samt trening og annen fysisk aktivitet som turgåing o.l. Sett kryss over det tallet som best angir ditt nivå av fysisk aktivitet.

Alder	Svært lite									Svært mye
14 år	1	2	3	4	5	6	7	8	9	10
30 år	1	2	3	4	5	6	7	8	9	10
I dag	1	2	3	4	5	6	7	8	9	10

Hvor mange timer pr. dag i gjennomsnitt går eller spaserer du utendørs?

	sjelden aldri	mindre enn 1/2 time	1/2-1 time	1-2 timer	mer enn 2 timer
Vinter	☐	☐	☐	☐	☐
Vår	☐	☐	☐	☐	☐
Sommer	☐	☐	☐	☐	☐
Høst	☐	☐	☐	☐	☐

For hver av følgende aktiviteter du deltar i, ber vi deg oppgi hvor mange minutter pr. dag du bruker i gjennomsnitt til hver av aktivitetene.

Fritidsaktivitet	Vinter	Vår	Sommer	Høst
Se på TV	☐☐☐	☐☐☐	☐☐☐	☐☐☐
Lesing	☐☐☐	☐☐☐	☐☐☐	☐☐☐
Håndarbeid/hobby	☐☐☐	☐☐☐	☐☐☐	☐☐☐
Hagearbeid	☐☐☐	☐☐☐	☐☐☐	☐☐☐
Dusj/bad/egenpleie	☐☐☐	☐☐☐	☐☐☐	☐☐☐

Høyde og vekt

Hvor høy er du? (i hele cm.) ..... ☐☐☐

Hvor mye veide du da du var 18 år? (i hele kg.) ☐☐☐

Hvor mye veier du i dag? (i hele kg.) ..... ☐☐☐

Kosthold

Påvirker noen av følgende forhold kostholdet ditt?
(sett gjerne flere kryss)

- ☐ Er vegetarianer/veganer
- ☐ Har anoreksi
- ☐ Spiser ikke norsk kost til daglig
- ☐ Har allergi/intoleranse
- ☐ Har bulimi
- ☐ Kronisk sykdom
- ☐ Prøver å gå ned i vekt

Vi er interessert i å få kjennskap til hvordan kostholdet ditt er vanligvis. Kryss av for hvert spørsmål om hvor ofte du i gjennomsnitt siste året har brukt den aktuelle matvaren, og hvor mye du pleier å spise/drikke hver gang.

Hvor mange glass melk drikker du vanligvis av hver type? (Sett ett kryss pr. linje)

	aldri/ sjelden	1-4 pr. uke	5-6 pr. uke	1 pr. dag	2-3 pr. dag	4+ pr. dag
Helmelk (søt, sur)	☐	☐	☐	☐	☐	☐
Lettmelk (søt, sur)	☐	☐	☐	☐	☐	☐
Ekstra lettmelk	☐	☐	☐	☐	☐	☐
Skummet (søt, sur)	☐	☐	☐	☐	☐	☐

Hvor mange kopper kaffe/te drikker du vanligvis av hver sort? (Sett ett kryss for hver linje)

	aldri/ sjelden	1-6 pr. uke	1 pr. dag	2-3 pr. dag	4-5 pr. dag	6-7 pr. dag	8+ pr. dag
Kokekaffe	☐	☐	☐	☐	☐	☐	☐
Traktekaffe	☐	☐	☐	☐	☐	☐	☐
Pulverkaffe	☐	☐	☐	☐	☐	☐	☐
Espresso o.l.	☐	☐	☐	☐	☐	☐	☐
Svart te	☐	☐	☐	☐	☐	☐	☐
Grønn te	☐	☐	☐	☐	☐	☐	☐

Hvor mange glass vann drikker du vanligvis?
(Sett ett kryss for hver linje)

	aldri/ sjelden	1-3 pr. uke	4-6 pr. uke	1 pr. dag	2-3 pr. dag	4+ pr. dag
Springvann	☐	☐	☐	☐	☐	☐
Flaskevann u/kullsyre	☐	☐	☐	☐	☐	☐
Flaskevann m/kullsyre	☐	☐	☐	☐	☐	☐

Hvor mange glass appelsinjuice, saft og brus drikker du vanligvis? (Sett ett kryss for hver linje)

	aldri/ sjelden	1-3 pr. uke	4-6 pr. uke	1 pr. dag	2-3 pr. dag	4+ pr. dag
Appelsinjuice	☐	☐	☐	☐	☐	☐
Saft/brus med sukker	☐	☐	☐	☐	☐	☐
Saft/brus sukkerfri	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du yoghurt (1 beger)? (Sett ett kryss)

☐ Aldri/sjelden ☐ 1 pr. uke ☐ 2-3 pr. uke ☐ 4+ pr. uke

Hvor ofte spiser du kornblanding, havregryn eller müsli? (Sett ett kryss)

☐ Aldri/sjelden ☐ 1-3 pr. uke ☐ 4-6 pr. uke ☐ 1 pr. dag

Hvor mange skiver brød/rundstykker og knekkebrød/skonrokker spiser du vanligvis?
(1/2 rundstykke = 1 brødskive) (Sett ett kryss for hver linje)

	aldri/ sjelden	1-4 pr. uke	5-7 pr. uke	2-3 pr. dag	4-5 pr. dag	6+ pr. dag
Grovt brød	☐	☐	☐	☐	☐	☐
Kneipp/halvfint	☐	☐	☐	☐	☐	☐
Fint brød	☐	☐	☐	☐	☐	☐
Knekkebrød o.l.	☐	☐	☐	☐	☐	☐

Nedenfor er det spørsmål om bruk av ulike påleggstyper. Vi spør om hvor mange brødskiver med det aktuelle pålegget du pleier å spise. Dersom du også bruker matvarene i andre sammenhenger enn til brød (f. eks. til vafler, frokostblandinger, grøt), ber vi om at du tar med dette når du besvarer spørsmålene.

På hvor mange brødskiver bruker du? (Sett ett kryss pr. linje)

	0 pr. uke	1-3 pr. uke	4-6 pr. uke	1 pr. dag	2-3 pr. dag	4+ pr. dag
Syltetøy	☐	☐	☐	☐	☐	☐
Brun ost, helfet	☐	☐	☐	☐	☐	☐
Brunost, halvfet/mager	☐	☐	☐	☐	☐	☐
Hvitost, helfet	☐	☐	☐	☐	☐	☐
Hvitost, halvfet/mager	☐	☐	☐	☐	☐	☐
Kjøttpålegg, Leverpostei	☐	☐	☐	☐	☐	☐
Rekesalat, italiensk o.l.	☐	☐	☐	☐	☐	☐

På hvor mange brødskiver pr. uke har du i gjennomsnitt siste året spist? (Sett ett kryss pr. linje)

	0 pr. uke	1 pr.uke	2-3 pr.uke	4-6 pr.uke	7-9 pr.uke	10+ pr.uke
Makrell i tomat, røkt makrell	☐	☐	☐	☐	☐	☐
Kaviar	☐	☐	☐	☐	☐	☐
Sild/Ansjos	☐	☐	☐	☐	☐	☐
Laks (gravet/røkt)	☐	☐	☐	☐	☐	☐
Annet fiskepålegg	☐	☐	☐	☐	☐	☐

Hva slags fett bruker du vanligvis på brødet?
(Sett gjerne flere kryss)

- ☐ Bruker ikke fett på brødet
- ☐ Smør
- ☐ Hard margarin (f. eks. Per, Melange)
- ☐ Myk margarin (f. eks. Soft, Vita, Solsikke)
- ☐ Smørblandet margarin (f.eks. Bremyk)
- ☐ Brelett
- ☐ Lettmargarin (f. eks. Soft light, Letta)
- ☐ Middels lett margarin (f. eks. Olivero, Omega)

Dersom du bruker fett på brødet, hvor tykt lag pleier du smøre på? (En kuvertpakke med margarin veier 12 gram).
(Sett ett kryss)

☐ Skrapet (3 g) ☐ Tynt lag (5 g) ☐ Godt dekket (8 g) ☐ Tykt lag (12 g)

Hvor ofte spiser du frukt? (Sett ett kryss pr. linje)

	aldri/ sjelden	1-3 pr.mnd.	1 pr.uke	2-4 pr.uke	5-6 pr.uke	1 pr.dag	2+ pr. dag
Epler/pærer	☐	☐	☐	☐	☐	☐	☐
Appelsiner o.l.	☐	☐	☐	☐	☐	☐	☐
Bananer	☐	☐	☐	☐	☐	☐	☐
Annen frukt	☐	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du ulike typer grønnsaker?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1-3 pr.mnd.	1 pr.uke	2 pr.uke	3 pr.uke	4-5 pr.uke	6-7 pr. uke
Gulrøtter	☐	☐	☐	☐	☐	☐	☐
Kål	☐	☐	☐	☐	☐	☐	☐
Kålrot	☐	☐	☐	☐	☐	☐	☐
Brokkoli/blomkål	☐	☐	☐	☐	☐	☐	☐
Blandet salat	☐	☐	☐	☐	☐	☐	☐
Tomat	☐	☐	☐	☐	☐	☐	☐
Grønnsakblanding (frossen)	☐	☐	☐	☐	☐	☐	☐
Andre grønnsaker	☐	☐	☐	☐	☐	☐	☐

For de grønnsakene du spiser, kryss av for hvor mye du spiser hver gang. (Sett ett kryss for hver sort)

- gulrøtter	☐ 1/2 stk.	☐ 1 stk.	☐ 1 1/2 stk.	☐ 2+ stk.
- kål	☐ 1/2 dl	☐ 1 dl	☐ 1 1/2 dl	☐ 2+ dl
- kålrot	☐ 1/2 dl	☐ 1 dl	☐ 1 1/2 dl	☐ 2+ dl
- brokkoli/blomkål	☐ 1-2 buketter	☐ 3-4 buketter	☐ 5+ buketter	
- blandet salat	☐ 1 dl	☐ 2 dl	☐ 3 dl	☐ 4+ dl
- tomat	☐ 1/4	☐ 1/2	☐ 1	☐ 2+
- grønnsakblanding	☐ 1/2 dl	☐ 1 dl	☐ 2 dl	☐ 3+ dl

Hvor mange poteter spiser du vanligvis (kokte, stekte, mos)? (Sett ett kryss)

☐ Spiser ikke/spiser sjelden poteter
☐ 1-4 pr. uke ☐ 5-6 pr. uke ☐ 1 pr. dag ☐ 2 pr. dag
☐ 3 pr. dag ☐ 4+ pr. dag

Hvor ofte bruker du ris og spagetti/makaroni ?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1-3 pr. mnd.	1 pr. uke	2 pr. uke	3+ pr. uke
Ris	☐	☐	☐	☐	☐
Spagetti, makaroni	☐	☐	☐	☐	☐

Hvor ofte spiser du grøt ? (Sett ett kryss)

	aldri/ sjelden	1 pr. mnd.	2-3 pr. mnd.	1 pr. uke	2-6 pr. uke	1+ pr. dag
Risengrynsgrøt	☐	☐	☐	☐	☐	☐
Annen grøt (havre o.l.)	☐	☐	☐	☐	☐	☐

Fisk

Vi vil gjerne vite hvor ofte du pleier å spise fisk, og ber deg fylle ut spørsmålene om fiskeforbruk så godt du kan. Tilgangen på fisk kan variere gjennom året. Vær vennlig å markere i hvilke årstider du spiser de ulike fiskeslagene.

	aldri/ sjelden	like mye hele året	vintrer	vår	sommer	høst
Torsk, sei, hyse, lyr	☐	☐	☐	☐	☐	☐
Steinbit, flyndre, uer	☐	☐	☐	☐	☐	☐
Laks, ørret	☐	☐	☐	☐	☐	☐
Makrell	☐	☐	☐	☐	☐	☐
Sild	☐	☐	☐	☐	☐	☐
Annen fisk	☐	☐	☐	☐	☐	☐

Med tanke på de periodene av året der du spiser fisk, hvor ofte pleier du å spise følgende?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd.	2-3 pr. mnd.	1 pr. uke	2+ pr. uke
Kokt torsk, sei, hyse, lyr	☐	☐	☐	☐	☐
Stekt torsk, sei, hyse, lyr	☐	☐	☐	☐	☐
Steinbit, flyndre, uer	☐	☐	☐	☐	☐
Laks, ørret	☐	☐	☐	☐	☐
Makrell	☐	☐	☐	☐	☐
Sild	☐	☐	☐	☐	☐
Annen fisk	☐	☐	☐	☐	☐

Dersom du spiser fisk, hvor mye spiser du vanligvis pr. gang? (1 skive/stykke = 150 gram)

Kokt fisk (skive) ☐ 1 ☐ 1,5 ☐ 2 ☐ 3+

Stekt fisk (stykke) ☐ 1 ☐ 1,5 ☐ 2 ☐ 3+

Hvor mange ganger pr. år spiser du fiskeinnmat?
(Sett ett kryss pr. linje)

	0	1-3	4-6	7-9	10+
Rogn	☐	☐	☐	☐	☐
Fiskelever	☐	☐	☐	☐	☐

Dersom du spiser fiskelever, hvor mange spiseskjeer pleier du å spise hver gang? (Sett ett kryss)

1	2	3-4	5-6	7+
☐	☐	☐	☐	☐

Hvor ofte bruker du følgende typer fiskemat?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd.	2-3 pr. mnd.	1 pr. uke	2+ pr. uke
Fiskekaker/pudding/boller	☐	☐	☐	☐	☐
Plukkfisk/fiskegrateng	☐	☐	☐	☐	☐
Frityrfisk/fiskepinner	☐	☐	☐	☐	☐
Andre fiskeretter	☐	☐	☐	☐	☐

Hvor stor mengde pleier du vanligvis å spise av de ulike rettene? (Sett ett kryss for hver linje)

- fiskekaker/pudding/boller (stk.) (2 fiskeboller=1 fiskekake) ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- plukkfisk, fiskegrateng (dl) ☐ 1-2 ☐ 3-4 ☐ 5+
- frityrfisk, fiskepinner (stk.) ☐ 1-2 ☐ 3-4 ☐ 5-6 ☐ 7+

I tillegg til informasjon om fiskeforbruk er det viktig å få kartlagt hvilket tilbehør som blir servert til fisk. Hvor ofte bruker du følgende til fisk? (Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd.	2-3 pr. mnd.	1 pr. uke	2+ pr. uke
Smeltet smør	☐	☐	☐	☐	☐
Smeltet eller fast margarin/fett	☐	☐	☐	☐	☐
Seterrømme (35%)	☐	☐	☐	☐	☐
Lettrømme (20%)	☐	☐	☐	☐	☐
Saus med fett (hvit/brun)	☐	☐	☐	☐	☐
Saus uten fett (hvit/brun)	☐	☐	☐	☐	☐

For de ulike typene tilbehør du bruker til fisk, vær vennlig å kryss av for hvor mye du vanligvis pleier spise.

- smeltet smør (ss) ☐ 1/2 ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- smeltet margasin (ss) ☐ 1/2 ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- seterrømme (ss) ☐ 1/2 ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- lettrømme (ss) ☐ 1/2 ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- saus med fett (dl) ☐ 1/4 ☐ 1/2 ☐ 3/4 ☐ 1 ☐ 2+
- saus uten fett (dl) ☐ 1/4 ☐ 1/2 ☐ 3/4 ☐ 1 ☐ 2+

Hvor ofte spiser du skalldyr (f. eks. reker, krabbe og skjell)**?** (Sett ett kryss)

☐ Aldri/sjelden ☐ 1 pr. mnd ☐ 2-3 pr. mnd ☐ 1+ pr. uke

Andre matvarer

Hvor ofte spiser du reinkjøtt?

☐ Aldri/sjelden ☐ 1 pr. mnd. ☐ 2-3 pr. mnd. ☐ 1 pr. uke
☐ 2-3 pr. uke ☐ 4+ pr. uke

Hvor ofte spiser du følgende kjøtt- og fjærkreretter? (Sett ett kryss for hver rett)

	aldri/ sjelden	1 pr.mnd.	2-3 pr.mnd.	1 pr.uke	2+ pr.uke
Steik (okse, svin, får)	☐	☐	☐	☐	☐
Koteletter	☐	☐	☐	☐	☐
Biff	☐	☐	☐	☐	☐
Kjøttkaker, karbonader	☐	☐	☐	☐	☐
Pølser	☐	☐	☐	☐	☐
Gryterett, lapskaus	☐	☐	☐	☐	☐
Pizza med kjøtt	☐	☐	☐	☐	☐
Kylling	☐	☐	☐	☐	☐
Andre kjøttretter	☐	☐	☐	☐	☐

Dersom du spiser følgende retter, oppgi mengden du vanligvis spiser: (Sett ett kryss for hver linje)

- steik (skiver) ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- koteletter (stk.) ☐ 1/2 ☐ 1 ☐ 1,5 ☐ 2+
- kjøttkaker, karbonader (stk.) ☐ 1 ☐ 2 ☐ 3 ☐ 4+
- pølser (stk. à 150g) ☐ 1/2 ☐ 1 ☐ 1,5 ☐ 2+
- gryterett, lapskaus (dl) ☐ 1-2 ☐ 3 ☐ 4 ☐ 5+
- pizza m/kjøtt (stykke à 100 g) ☐ 1 ☐ 2 ☐ 3 ☐ 4+

Hvor mange egg spiser du vanligvis i løpet av en uke?(stekte, kokte, eggerøre, omelett) (Sett ett kryss)

☐ 0 ☐ 1 ☐ 2 ☐ 3-4 ☐ 5-6 ☐ 7+

Hvor ofte spiser du iskrem? (til dessert, krone-is osv.) Sett et kryss for hvor ofte du spiser iskrem om sommeren, og et kryss for resten av året)

	aldri/ sjelden	1-3. pr.	2-3 pr. mnd.	1 pr. uke	2+ pr. uke
-Om sommeren	☐	☐	☐	☐	☐
-Resten av året	☐	☐	☐	☐	☐

Hvor mye is spiser du vanligvis pr. gang? (Sett ett kryss)

☐ 1dl ☐ 2 dl ☐ 3 dl ☐ 4+ dl

Hvor ofte spiser du bakevarer som boller kaker, wienerbrød eller småkaker (Sett ett kryss pr. linje)

	aldri/ sjelden	1-3 mnd.	1 pr. uke	2-3 pr uke	4-6 pr. uke	1+ pr. dag
Gjærbakst (boller)	☐	☐	☐	☐	☐	☐
Wienerbrød, kringle	☐	☐	☐	☐	☐	☐
Kaker (bløtkaker)	☐	☐	☐	☐	☐	☐
Pannekaker	☐	☐	☐	☐	☐	☐
Vafler	☐	☐	☐	☐	☐	☐
Småkaker, kjeks	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du dessert? (Sett ett kryss pr. linje)

	aldri/ sjelden	1-3 mnd.	1 pr. uke	2-3 pr uke	4-6 pr. uke	1+ pr. dag
Pudding sjokolade/karamell	☐	☐	☐	☐	☐	☐
Riskrem, fromasj	☐	☐	☐	☐	☐	☐
Kompott, fruktgrøt, hermetisk frukt	☐	☐	☐	☐	☐	☐
Jorbær (friske, frosne)	☐	☐	☐	☐	☐	☐
Andre bær (friske, frosne)	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du sjokolade? (Sett ett kryss)

	aldri/ sjelden	1-3 mnd.	1 pr. uke	2-3 pr uke	4-6 pr. uke	1+ pr. dag
Mørk sjokolade	☐	☐	☐	☐	☐	☐
Lys sjokolade	☐	☐	☐	☐	☐	☐

Dersom du spiser sjokolade, hvor mye pleier du vanligvis å spise hver gang? Tenk deg størrelsen på en Kvikk-Lunsj sjokolade, og oppgi hvor mye du spiser i forhold til den.

☐ 1/4 ☐ 1/2 ☐ 3/4 ☐ 1 ☐ 1,5 ☐ 2+

Hvor ofte spiser du snacks? (Sett ett kryss)

	aldri/ sjelden	1-3 pr. mnd.	1 pr. uke	2-3 pr. uke	4-6 pr. uke	7+ pr. uke
Potetchips	☐	☐	☐	☐	☐	☐
Peanøtter	☐	☐	☐	☐	☐	☐
Andre nøtter	☐	☐	☐	☐	☐	☐
Annen snacks	☐	☐	☐	☐	☐	☐

Tran og fiskeoljekapsler

Bruker du tran (flytende)? Ja ☐ Nei ☐

Hvis ja; hvor ofte tar du tran?
Sett ett kryss for hver linje.

	aldri/ sjelden	1-3 pr. mnd.	1 pr. uke	2-6 pr. uke	daglig
Om vinteren	☐	☐	☐	☐	☐
Resten av året	☐	☐	☐	☐	☐

Hvor mye tran pleier du å ta hver gang?

☐ 1 ts. ☐ 1/2 ss. ☐ 1+ ss.

Bruker du tranpiller/kapsler? Ja ☐ Nei ☐

Hvis ja; hvor ofte tar du tranpiller/kapsler?
Sett ett kryss for hver linje.

	aldri/ sjelden	1-3 pr. mnd.	1 pr. uke	2-6 pr. uke	daglig
Om vinteren	☐	☐	☐	☐	☐
Resten av året	☐	☐	☐	☐	☐

Hvilken type tranpiller/kapsler bruker du vanligvis, og hvor mange pleier du å ta hver gang?

Antall

Navn _____ ☐

Bruker du fiskeoljekapsler? (omega-3) Ja ☐ Nei ☐

Hvis ja; hvor ofte tar du fiskeoljekapsler?

aldri/ sjelden	1-3 pr. mnd.	1 pr. uke	2-6 pr. uke	daglig
☐	☐	☐	☐	☐

Hvilken type fiskeoljekapsler bruker du vanligvis, og hvor mange pleier du å ta hver gang?

Navn _____ antall ☐

Varm mat

Hvor mange ganger i løpet av en måned spiser du varm mat?

Antall

Til frokost ☐

Til lunsj ☐

Til middag ☐

Til kvelds ☐

Kosttilskudd

Hvor ofte bruker du kosttilskudd?
(Sett ett kryss pr. linje)

Navn på vitamin/mineraltilskudd:	aldri/ sjelden	1-3 pr. mnd.	1 pr. uke	2-6 pr. uke	daglig
_____	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐

Alkohol

Er du totalavholdskvinne? Ja ☐ Nei ☐

Hvis Nei, hvor ofte og hvor mye drakk du i gjennomsnitt siste året? (Sett ett kryss for hver linje)

	aldri/ sjelden	1 pr. mnd.	2-3 pr. uke	1 pr. uke	2-4 pr. uke	5-6 pr. uke	1+ pr. dag
Øl (1/2 l.)	☐	☐	☐	☐	☐	☐	☐
Vin (glass)	☐	☐	☐	☐	☐	☐	☐
Brennevin (drink)	☐	☐	☐	☐	☐	☐	☐
Likør/Hetvin	☐	☐	☐	☐	☐	☐	☐

Sosiale forhold

Er du: (Sett ett kryss)

☐ gift ☐ samboer ☐ ugift ☐ skilt ☐ enke

Hvor mange års skolegang/yrkesutdannelse har du i alt, ta med folkeskole og ungdomsskole? ☐

Hvor mange personer er det i ditt hushold? ☐

Hvor høy er bruttoinntekten i husholdet pr. år?

under 150.000 kr.	☐	151.000-300.000 kr.	☐
301.000-450.000 kr.	☐	451.000-600.000 kr.	☐
601.000-750.000 kr.	☐	over 750.000 kr.	☐

Hva er din arbeidssituasjon? (sett kryss)

☐ Arbeider heltid ☐ Arbeider deltid ☐ Pensjonist

☐ Hjemmearbeidende ☐ Under utdanning ☐ Uføretrygdet

☐ Under attføring ☐ Arbeidssøkende

Yrke: _____

Hvordan var de økonomiske forhold i oppveksten?

☐ Meget gode ☐ Gode

☐ Dårlige ☐ Meget dårlige

Arbeider du utendørs i yrkessammenheng? Ja ☐ Nei ☐

Hvis Ja;
hvor mange timer pr. uke? .....Sommer .....vinter

Solvaner

Får du fregner når du soler deg? ☐ Ja ☐ Nei

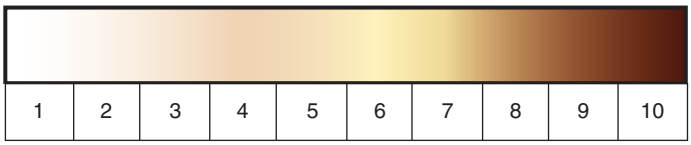
Hvilken øyefarge har du? (sett ett kryss)

☐ brun ☐ grå, grønn eller blanding ☐ blå

Hva er din opprinnelige hårfarge? (sett ett kryss)

☐ mørkbrunt, svart ☐ brun ☐ blond, gul ☐ rød

For å kunne studere effekten av soling på risiko for hudkreft ber vi deg gi opplysninger om hudfarge
Sett ett kryss på det tallet under fargen som best passer din naturlige hudfarge (uten soling)



1	2	3	4	5	6	7	8	9	10

Hvor mange ganger pr. år er du blitt forbrent av solen slik at du har fått svie og blemmer med avflassing etterpå? (ett kryss for hver aldersgruppe)

Alder	Aldri	Høyst 1 gang pr. år	2-3 g. pr. år	4-5 g. pr. år	6 eller flere ganger
Før 10 år	☐	☐	☐	☐	☐
10-19 år	☐	☐	☐	☐	☐
20-29 år	☐	☐	☐	☐	☐
30-44 år	☐	☐	☐	☐	☐
45+ år	☐	☐	☐	☐	☐

Hvor mange uker soler du deg pr. år i syden?

Alder	Aldri	1 uke	2-3 uker	4-5 uker	7 uker eller mer
Før 10 år	☐	☐	☐	☐	☐
10-19 år	☐	☐	☐	☐	☐
20-29 år	☐	☐	☐	☐	☐
30-44 år	☐	☐	☐	☐	☐
45+ år	☐	☐	☐	☐	☐
Siste 12 mnd.	☐	☐	☐	☐	☐

Hvor mange uker pr. år soler du deg i Norge eller utenfor syden?

Alder	Aldri	1 uke	2-3 uker	4-5 uker	7 uker eller mer
Før 10 år	☐	☐	☐	☐	☐
10-19 år	☐	☐	☐	☐	☐
20-29 år	☐	☐	☐	☐	☐
30-44 år	☐	☐	☐	☐	☐
45+ år	☐	☐	☐	☐	☐
Siste 12 mnd.	☐	☐	☐	☐	☐

Hvor ofte dusjer eller bader du?

	mer enn 1 g. dagl.	1 g. dagl.	4-6 g. pr. uke	2-3 g. pr. uke	1 g. pr.	2-3 g. pr.uke	sjel-den/ aldri
Med såpe/shampo	☐	☐	☐	☐	☐	☐	☐
Uten såpe/shampo	☐	☐	☐	☐	☐	☐	☐

Når bruker du krem med solfaktor? (sett evt. flere kryss):

☐ i påsken ☐ i Norge eller utenfor syden ☐ solferie i syden ☐ aldri

Hvilken solfaktor bruker du i disse periodene?

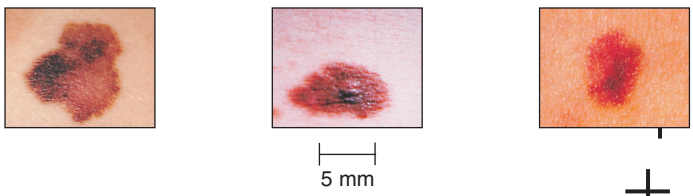
	påsken	i Norge eller utenfor syden	solferie i syden
I dag	☐☐	☐☐	☐☐
For 10 år siden	☐☐	☐☐	☐☐

Hvor ofte har du solt deg i solarium?

Alder	Aldri	Sjelden	1 gang pr. mnd.	2 ganger pr. mnd.	3-4 ganger pr. mnd	oftere enn1 gang pr. uke
Før 10 år	☐	☐	☐	☐	☐	☐
10-19 år	☐	☐	☐	☐	☐	☐
20-29 år	☐	☐	☐	☐	☐	☐
30-44 år	☐	☐	☐	☐	☐	☐
45+ år	☐	☐	☐	☐	☐	☐
Siste 12 mnd.	☐	☐	☐	☐	☐	☐

Hvor mange uregelmessige føflekker større enn 5 mm har du sammenlagt på begge beina (fra tærne til lysken)? Tre eksempler på føflekker større enn 5 mm med uregelmessig form er vist i nedenfor.

☐ 0 ☐ 1 ☐ 2-3 ☐ 4-6 ☐ 7-12 ☐ 13-24 ☐ 25+



Hvor ofte bruker du følgende hudpleiemidler?

(Sett ett kryss pr. linje)

	aldri/ sjelden	1-3 pr.mnd.	1 pr.uke	2-4 pr.uke	5-6 pr.uke	1 pr.dag	2+ pr. dag
Ansiktskrem	☐	☐	☐	☐	☐	☐	☐
Håndkrem	☐	☐	☐	☐	☐	☐	☐
Body lotion	☐	☐	☐	☐	☐	☐	☐
Parfyme	☐	☐	☐	☐	☐	☐	☐

Til slutt vil vi spørre deg om ditt samtykke til å kontakte deg på nytt pr. post. Vi vil hente adressen fra det sentrale personregister.

Ja ☐ Nei ☐

Er du villig til å avgi en blodprøve?

Ja ☐ Nei ☐

Takk for at du ville delta i undersøkelsen

Additional File 2.

KVINNER OG KREFT

Høst 2011

KONFIDENSIELT

Hvis du samtykker i å være med, sett kryss for JA i ruten ved siden av. Dersom du ikke ønsker å delta kan du unngå purring ved å sette kryss for NEI og returnere skjemaet i vedlagte svarkonvolutt. **Vi ber deg fylle ut spørreskjemaet så nøye som mulig.**

Skjemaet skal leses optisk. Vennligst bruk blå eller sort penn. Du kan ikke bruke komma, forhøy 0,5 til 1. Bruk blokkbokstaver.

Med vennlig hilsen
Eiliv Lund
Professor dr. med

47 KK/2011
12384 invit.
300001-330000
8siders - 2.gang

FH

Jeg samtykker i å delta i spørreskjemaundersøkelsen JA ☐ NEI ☐

Menstruasjon og overgangsalder

Har du regelmessig menstruasjon fremdeles?

☐ Ja
☐ Har uregelmessig menstruasjon
☐ Vet ikke (menstruasjon uteblitt pga. sykdom o.l.)
☐ Vet ikke (bruker hormonpreparat med østrogen)
☐ Nei

Hvis Nei;

har den stoppet av seg selv? ..... ☐
har du operert vekk eggstokkene? ..... ☐
har du operert vekk livmoren? ..... ☐
annet? ..... ☐

Alder da menstruasjonen opphørte ..... ☐☐

Har du eller har du hatt smerter eller ømhet i brystene av minst fem dagers varighet før menstruasjonen? ..... ☐ Ja ☐ Nei

Hvis Ja; i begge brystene? ☐ Ja ☐ Nei
er/var smerten eller ømheten mindre under og etter menstruasjonen? ☐ Ja ☐ Nei
forstyrret plagene ditt sosiale liv, yrkesaktivitet eller privatlivet? ☐ Ja ☐ Nei

Hvor mange år har du hatt slike plager? ☐☐

Bruk av hormonpreparater mot plager i overgangsalderen

Har du noen gang brukt østrogentabletter/plaster? ..... ☐ Ja ☐ Nei
(Gjelder også progestagen/ Tibolon)

Hvis Ja; hvor mange år i alt? ..... ☐☐

Hvor gammel var du første gang du brukte østrogentabletter/plaster? ..... ☐☐

Bruker du tabletter/plaster nå? ..... ☐ Ja ☐ Nei

Utfyllende spørsmål til alle som har brukt preparater med østrogen i form av tabletter eller plaster fra 2003 og frem til i dag

Har du svart «ja», ber vi deg utdype dette nærmere ved å svare på spørsmålene nedenfor. For hver periode med sammenhengende bruk av samme hormonpreparat håper vi du kan si oss hvor gammel du var da du startet, hvor lenge du brukte det samme hormonpreparatet og navnet på dette. Dersom du har hatt opphold eller skiftet merke skal du besvare spørsmålene for en ny periode. Dersom du ikke husker navnet på hormonpreparatet, sett «usikker». For å hjelpe deg til å huske navnet på hormonpreparatene ber vi deg bruke vedlagte brosjyre som viser bilder av hormonpreparater som har vært solgt i Norge. Vennligst oppgi også nummer på hormontabletten/plasteret som står i brosjyren.

Periode	Alder ved start	Brukt samme hormontablett/plaster/ sammenhengende fra 2003 antall år	antall mnd.	nr.	Navn på hormontablett/plaster (se brosjyre)
1.	☐☐	☐☐	☐☐	☐☐☐	
2.	☐☐	☐☐	☐☐	☐☐☐	
3.	☐☐	☐☐	☐☐	☐☐☐	
4.	☐☐	☐☐	☐☐	☐☐☐	
5.	☐☐	☐☐	☐☐	☐☐☐	

Har du eller har du hatt smerter eller ømhet i brystene ved hormonbehandling av plager i overgangsalderen? ..... ☐ Ja ☐ Nei

Hvis Ja;

er/var smerten eller ømheten i begge bryst? ☐ Ja ☐ Nei
forsvinner/forsvant plagene ved stopp av hormonbehandlingen? ☐ Ja ☐ Nei
forstyrrer/forstyrret plagene ditt sosiale liv, yrkesaktivitet eller privatlivet? ☐ Ja ☐ Nei

Byttet du legemiddel? ☐ Ja ☐ Nei
Hvis Ja, ble du bedre? ☐ Ja ☐ Nei

Østrogenpreparat til lokal bruk i skjeden

Har du noen gang brukt østrogen-krem/stikkpille? ☐ Ja ☐ Nei

Hvis Ja; bruker du krem/stikkpille nå? ☐ Ja ☐ Nei

Alternativer til hormonbehandling mot plager i overgangsalderen

Har du noen gang brukt alternativer til hormonbehandling mot plager i overgangsalderen? ☐ Ja ☐ Nei

Hvis Ja; har du brukt noen av følgende:

Soyatilskudd ☐ Ja ☐ Nei

Preparatnavn _____

Andre tilskudd for overgangsplager ☐ Ja ☐ Nei

Preparatnavn _____

Andre legemidler enn hormoner ☐ Ja ☐ Nei

Preparatnavn _____

Akupunktur ☐ Ja ☐ Nei

Homeopati ☐ Ja ☐ Nei

Avspenningsteknikk/trening ☐ Ja ☐ Nei

Andre alternativer, spesifiser: ☐ Ja ☐ Nei

_____

Sykdom

Har du eller har du hatt noen av følgende sykdommer? (sett ett eller flere kryss)

	Ja	Nei	Hvis ja: Alder ved start
Kreft	☐	☐	
Høyt blodtrykk	☐	☐	
Hjertesvikt/hjertekrampe	☐	☐	
Hjerteinfarkt	☐	☐	
Slag	☐	☐	
Depresjon (oppsøkt lege)	☐	☐	
Hypothyreose/lavt stoffskifte	☐	☐	
Sukkersyke (diabetes)	☐	☐	

Hvis ja på sukkersyke, hvilken type:

☐ Type 1 ☐ Aldersdiabetes ☐ Svangerskap

Behandles du i dag med (sett ett eller flere kryss):

☐ Insulin ☐ Legemidler ☐ Kost

For følgende tilstander ber vi deg krysse av for hvilket år tilstanden oppsto første gang

	før 04	04	05	06	07	08	09	10	11
Muskelsmerter (myalgi)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Fibromyalgi/Fibrositt	☐	☐	☐	☐	☐	☐	☐	☐	☐
Kronisk tretthetssyndrom	☐	☐	☐	☐	☐	☐	☐	☐	☐
Ryggsmerter ukjent årsak	☐	☐	☐	☐	☐	☐	☐	☐	☐
Nakkeslengskade	☐	☐	☐	☐	☐	☐	☐	☐	☐
Osteoporose (benskjørhet)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Brudd									
Underarmen (håndledd)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Lårhalsen	☐	☐	☐	☐	☐	☐	☐	☐	☐
Ryggvirvel (kompresjon)	☐	☐	☐	☐	☐	☐	☐	☐	☐

Selvopplevd helse

Oppfatter du din egen helse som? (Sett ett kryss)

☐ Meget god ☐ God ☐ Dårlig ☐ Meget dårlig

Høyde og vekt

Hvor høy er du i dag? (i hele cm) ____

Hvor mye veier du i dag? (i hele kg) ____

Røykevaner

Har du i løpet av livet røykt mer enn 100 sigaretter til sammen? ☐ Ja ☐ Nei

Hvis Ja, ber vi deg fylle ut for perioden 2003–2011 hvor mange sigaretter du i gjennomsnitt røykte pr. dag.

	Antall sigaretter pr. dag						
	0	1–4	5–9	10–14	15–19	20–24	25+
2003–2007	☐	☐	☐	☐	☐	☐	☐
2008–2011	☐	☐	☐	☐	☐	☐	☐

Hvor gammel var du da du tok din første sigarett? ____

Røyker du daglig nå? ☐ Ja ☐ Nei

Hvis Nei, hvor gammel var du da du sluttet? ____

Røyker du av og til nå? ☐ Ja ☐ Nei

Røykte noen av dine foreldre da du var barn? ☐ Ja ☐ Nei

Hvis Ja, hvor mange sigaretter røykte de til sammen pr. dag? (Antall) ____

Fysisk aktivitet

Vi ber deg angi din fysiske aktivitet etter en skala fra svært lite til svært mye ved 14 års alder, ved 50 års alder og i dag. Skalaen nedenfor går fra 1–10. Med fysisk aktivitet mener vi både arbeid i hjemmet og i yrkeslivet samt trening og annen fysisk aktivitet som turgåing ol.

Alder	Svært lite									Svært mye
14 år	1	2	3	4	5	6	7	8	9	10
50 år	1	2	3	4	5	6	7	8	9	10
I dag	1	2	3	4	5	6	7	8	9	10

Vi er interessert i informasjon om ulike former for fysisk aktivitet i dagliglivet. Spørsmålene gjelder tiden du har brukt på fysisk aktivitet de siste 7 dagene. Vennligst svar på alle spørsmålene uansett hvor fysisk aktiv du er. Tenk på aktiviteter du gjør på jobb, som en del av hus- og hagearbeid, for å komme deg fra et sted til et annet og aktiviteter på fritiden (rekreasjon, mosjon og sport).

Tenk på alle svært anstrengende aktiviteter du har drevet med de siste 7 dagene. Svært anstrengende aktivitet er aktivitet som krever hard innsats og får deg til å puste mye mer enn vanlig. Ta bare med aktiviteter som varer minst 10 minutter i strekk.

1. Hvor mange dager i løpet av de siste 7 dager har du drevet med meget anstrengende aktivitet som tunge løft, gravearbeid, aerobics, løp eller rask sykling?

   Dager i uken ..... ☐

   ☐ Ingen meget anstrengende aktivitet *Gå til spørsmål 3*

2. Hvor lang tid brukte du vanligvis på svært anstrengende aktivitet en av disse dagene? ..... timer pr. dag ☐ minutter pr. dag ☐ Vet ikke ☐

Tenk på all middels anstrengende aktivitet du har drevet med de siste 7 dagene. Middels anstrengende aktivitet er aktivitet som krever moderat innsats og får deg til å puste litt mer enn vanlig. Ta bare med aktiviteter som varer minst 10 minutter i strekk.

3. Hvor mange dager i løpet av de siste 7 dagene har du drevet med middels anstrengende fysisk aktivitet som å bære lette ting, jogge eller sykle i moderat tempo? Ikke ta med gange.

   Dager i uken ..... ☐

   ☐ Ingen *Gå til spørsmål 5*

4. Hvor lang tid brukte du vanligvis på middels anstrengende fysisk aktivitet på en av disse dagene? ..... timer pr. dag ☐ minutter pr. dag ☐ Vet ikke ☐

Tenk på tiden du har brukt på å gå de siste 7 dagene. Dette inkluderer gange på jobb og hjemme, gange fra ett sted til et annet eller gange som du gjør på tur eller som trening på fritiden.

5. Hvor mange dager i løpet av de siste 7 dagene gikk du i minst 10 minutter i strekk?

   Dager i uken ..... ☐

   ☐ Ingen *Gå til spørsmål 7*

6. Hvor lang tid brukte du vanligvis på å gå en av disse dagene? ..... timer pr. dag ☐ minutter pr. dag ☐ Vet ikke ☐

Tenk på all tid du har tilbrakt sittende på ukedagene i løpet av de siste 7 dagene. Inkluder tid du har brukt på å sitte på jobb, hjemme, på kurs og på fritiden. Dette kan tilsvare tiden du sitter ved et arbeidsbord, hos venner, mens du leser eller sitter eller ligger for å se på TV.

7. Hvor lang tid brukte du på å sitte en vanlig hverdag i løpet av de siste 7 dagene? ..... timer pr. dag ☐ minutter pr. dag ☐ Vet ikke ☐

Brystkreft i nærmeste familie

Har noen nære slektninger hatt brystkreft?

	Ja	Nei	Vet ikke	Alder ved start
Datter	☐	☐	☐	☐
Mor	☐	☐	☐	☐
Søster	☐	☐	☐	☐

Mammografiundersøkelse

Har du vært til undersøkelse av brystene med mammografi ..... ☐ Ja ☐ Nei

Hvis Ja, hvor gammel var du? (hele år) ..... første gang ☐ siste gang ☐

Hvor mange ganger har du vært undersøkt?

-etter invitasjon fra Kreftregisteret/ Det nasjonale mammografiprogrammet ..... ☐

-etter henvisning fra lege ..... ☐

-uten henvisning fra lege ..... ☐

- som del av egen forsikring/ gjennom arbeidsplass ..... ☐

- gjennom frivillige organisasjoner ..... ☐

Kosthold

Vi er interessert i å få kjennskap til hvordan kostholdet ditt er vanligvis. Kryss av for hvert spørsmål om hvor ofte du i gjennomsnitt siste året har brukt den aktuelle matvaren, og hvor mye du pleier å spise/drikke hver gang.

Drikke

Hvor mange glass melk drikker du vanligvis av hver type? (Sett ett kryss pr. linje)

	aldri/ sjelden	1–4 pr. uke	5–6 pr. uke	1 pr. dag	2–3 pr. dag	4+ pr.dag
Helmelk (søt, sur)	☐	☐	☐	☐	☐	☐
Lettmelk (søt, sur)	☐	☐	☐	☐	☐	☐
Ekstra lettmelk	☐	☐	☐	☐	☐	☐
Skummet (søt, sur)	☐	☐	☐	☐	☐	☐

Hvor mange kopper kaffe/te drikker du vanligvis av hver sort? (Sett ett kryss for hver linje)

	aldri/ sjelden	1–6 pr. uke	1 pr. dag	2–3 pr. dag	4–5 pr.dag	6–7 pr.dag	8+ pr.dag
Kokekaffe, presskanne	☐	☐	☐	☐	☐	☐	☐
Traktekaffe	☐	☐	☐	☐	☐	☐	☐
Espresso	☐	☐	☐	☐	☐	☐	☐
Latte	☐	☐	☐	☐	☐	☐	☐
Pulverkaffe	☐	☐	☐	☐	☐	☐	☐
Svart te	☐	☐	☐	☐	☐	☐	☐
Grønn te	☐	☐	☐	☐	☐	☐	☐

Bruker du følgende i kaffe

Sukker *(ikke kunstig søtstoff)* ☐ Ja ☐ Nei

Melk eller fløte ☐ Ja ☐ Nei

Bruker du følgende i te

Sukker *(ikke kunstig søtstoff)* ☐ Ja ☐ Nei

Melk eller fløte ☐ Ja ☐ Nei

Hvor mange glass vann drikker du vanligvis?
(Sett ett kryss for hver linje)

	aldri/ sjelden	1–6 pr. uke	1 pr. dag	2–3 pr. dag	4–5 pr.dag	6–7 pr.dag	8+ pr.dag
Springvann/ flaskevann	☐	☐	☐	☐	☐	☐	☐

Hvor mange glass juice, saft og brus drikker du vanligvis? (Sett ett kryss for hver linje)

	aldri/ sjelden	1–3 pr. uke	4–6 pr. uke	1 pr. dag	2–3 pr. dag	4+ pr.dag
Appelsinjuice	☐	☐	☐	☐	☐	☐
Annen juice	☐	☐	☐	☐	☐	☐
Saft/brus med sukker	☐	☐	☐	☐	☐	☐
Saft/brus sukkerfri	☐	☐	☐	☐	☐	☐

Yoghurt/kornblanding

Hvor ofte spiser du yoghurt (1 beger)? (Sett ett kryss)

☐ Aldri/sjelden ☐ 1–3 pr. uke

☐ 4–6 pr. uke ☐ 1 + pr. dag

Hvor ofte spiser du kornblanding, havregryn eller müsli? (Sett ett kryss)

☐ Aldri/sjelden ☐ 1–3 pr. uke

☐ 4–6 pr. uke ☐ 1 + pr. dag

Brødmat

Hvor mange skiver brød/rundstykker og knekkebrød/skonrokker spiser du vanligvis?
(½ rundstykke = 1 brødskive) (Sett ett kryss for hver linje)

	aldri/ sjelden	1–4 pr. uke	5–7 pr. uke	2–3 pr. dag	4–5 pr. dag	6+ pr.dag
Grovt brød	☐	☐	☐	☐	☐	☐
Kneipp/halvfint	☐	☐	☐	☐	☐	☐
Fint brød/baguett	☐	☐	☐	☐	☐	☐
Knekkebrød o.l.	☐	☐	☐	☐	☐	☐

Nedenfor er det spørsmål om bruk av ulike påleggstyper. Vi spør om hvor mange brødskiver med det aktuelle pålegget du pleier å spise. Dersom du også bruker matvarene i andre sammenhenger enn til brød (f. eks. til vafler, frokostblandinger, grøt), ber vi om at du tar med dette når du besvarer spørsmålene.

På hvor mange brødskiver bruker du?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. uke	4–6 pr. uke	1 pr. dag	2–3 pr. dag	4+ pr. dag
Syltetøy	☐	☐	☐	☐	☐	☐
Brunost, helfet	☐	☐	☐	☐	☐	☐
Brunost, halvfet/mager	☐	☐	☐	☐	☐	☐
Hvitost, helfet	☐	☐	☐	☐	☐	☐
Hvitost, halvfet/mager	☐	☐	☐	☐	☐	☐
Rekesalat, italiensk o.l.	☐	☐	☐	☐	☐	☐

Kjøttpålegg
(Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. uke	4–6 pr. uke	1 pr. dag	2–3 pr. dag	4+ pr. dag
Leverpostei	☐	☐	☐	☐	☐	☐
Magert (kokt skinke o.l.)	☐	☐	☐	☐	☐	☐
Fett (salami, fenalår o.l.)	☐	☐	☐	☐	☐	☐

På hvor mange brødskiver pr. uke har du i gjennomsnitt siste året spist? (Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. uke	2–3 pr. uke	4–6 pr. uke	7–9 pr. uke	10+ pr. uke
Makrell i tomat, røkt makrell	☐	☐	☐	☐	☐	☐
Kaviar	☐	☐	☐	☐	☐	☐
Sild/Ansjos	☐	☐	☐	☐	☐	☐
Laks *(gravet/røkt)*	☐	☐	☐	☐	☐	☐
Annet fiskepålegg	☐	☐	☐	☐	☐	☐

Dersom du bruker fett på brødet, hvor tykt lag pleier du å smøre på? (En kuvertpakke med margarin veier 12 gram). (Sett ett kryss)

☐ Skrapet *(3 g)* ☐ Tynt lag *(5 g)*

☐ Godt dekket *(8 g)* ☐ Tykt lag *(12 g)*

Hva slags fett bruker du vanligvis på brødet?
(Sett gjerne flere kryss)

- ☐ Bruker ikke fett på brødet
- ☐ Smør
- ☐ Hard margarin *(f. eks. Melange)*
- ☐ Myk margarin *(f. eks. Soft, Vita)*
- ☐ Smørblandet margarin *(f.eks. Bremyk)*
- ☐ Brelett
- ☐ Lettmargarin *(f. eks. Soft light, Vita Lett)*
- ☐ Margarin med olivenolje *(f. eks. Brelett oliven, Soft oliven)*

Frukt og grønnsaker

Hvor ofte spiser du frukt? (Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–4 pr. uke	5–6 pr. uke	1 pr.dag	2+ pr.dag
Epler/pærer	☐	☐	☐	☐	☐	☐	☐
Appelsiner o.l.	☐	☐	☐	☐	☐	☐	☐
Bananer	☐	☐	☐	☐	☐	☐	☐
Annen frukt	☐	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du kokt potet? (Sett ett kryss pr. linje)

	aldri/ sjelden	1–4 pr. mnd	5–6 pr. uke	1 pr. dag	2 pr. dag
Kokt	☐	☐	☐	☐	☐

Hvor mange poteter spiser du hver gang?
(Sett ett kryss)

☐ 0 ☐ 1 ☐ 2 ☐ 3–4 ☐ 5–6 ☐ 7+

Hvor ofte spiser du stekt, fritert eller most potet
(Sett ett kryss pr. linje)

	aldri/ sjelden	1–4 pr. mnd	5–6 pr. uke	1 pr. dag	2 pr. dag
Stekt, fritert, most	☐	☐	☐	☐	☐

Hvor ofte spiser du ulike typer grønnsaker?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. mnd.	1 pr. uke	2 pr. uke	3 pr. uke	4–5 pr. uke	6–7 pr. uke
Gulrøtter	☐	☐	☐	☐	☐	☐	☐
Kål	☐	☐	☐	☐	☐	☐	☐
Kålrot	☐	☐	☐	☐	☐	☐	☐
Brokkoli/blomkål	☐	☐	☐	☐	☐	☐	☐
Blandet salat	☐	☐	☐	☐	☐	☐	☐
Tomat	☐	☐	☐	☐	☐	☐	☐
Grønnsakblanding	☐	☐	☐	☐	☐	☐	☐
Løk	☐	☐	☐	☐	☐	☐	☐
Bønner	☐	☐	☐	☐	☐	☐	☐
Erter	☐	☐	☐	☐	☐	☐	☐
Andre grønnsaker	☐	☐	☐	☐	☐	☐	☐

For de grønnsakene du spiser, kryss av for hvor mye du spiser hver gang. (Sett ett kryss for hver sort)

Gulrøtter	☐ ½ stk	☐ 1 stk	☐ 1 ½ stk	☐ 2+ stk.
Kål	☐ ½ dl	☐ 1 dl	☐ 1 ½ dl	☐ 2+ dl
Kålrot	☐ ½ dl	☐ 1 dl	☐ 1 ½ dl	☐ 2+ dl
Brokkoli/blomkål	☐ 1–2 buketter	☐ 3–4 buketter	☐ 5+ buketter	
Blandet salat	☐ 1 dl	☐ 2 dl	☐ 3dl	☐ 4+ dl
Tomat	☐ 1/4 stk	☐ ½ stk	☐ 1 stk	☐ 2+ stk.
Grønnsakblanding	☐ ½ dl	☐ 1 dl	☐ 2 dl	☐ 3+ dl
Bønner	☐ 1–2 ss	☐ 3–4 ss	☐ 5–6 ss	☐ 7+ ss
Erter	☐ 1–2 ss	☐ 3–4 ss	☐ 5–6 ss	☐ 7+ ss

Ris, spaghetti, grøt, suppe

Hvor ofte bruker du ris og spaghetti/makaroni?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2 pr. uke	3+ pr. uke
Ris	☐	☐	☐	☐	☐
Spagetti, makaroni, nudler	☐	☐	☐	☐	☐

Hvor ofte spiser du grøt ? (Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd	2–3 pr. mnd	1 pr. uke	2–6 pr. uke	1+ pr. dag
Risengrynsgrøt	☐	☐	☐	☐	☐	☐
Annen grøt (havre o.l.)	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du suppe? (Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2 pr. uke	3+ pr. uke
Som hovedrett	☐	☐	☐	☐	☐
Som forrett, lunsj eller kveldsmat	☐	☐	☐	☐	☐

Fisk

Vi vil gjerne vite hvor ofte du pleier å spise fisk, og ber deg fylle ut spørsmålene om fiskeforbruk så godt du kan. Tilgangen på fisk kan variere gjennom året. Vær vennlig å markere i hvilke årstider du spiser de ulike fiskeslagene.

	aldri/ sjelden	like mye hele året	vinter	vår	sommer	høst
Torsk, sei, hyse, lyr	☐	☐	☐	☐	☐	☐
Steinbit, flyndre, uer	☐	☐	☐	☐	☐	☐
Laks, ørret	☐	☐	☐	☐	☐	☐
Makrell	☐	☐	☐	☐	☐	☐
Sild	☐	☐	☐	☐	☐	☐
Annen fisk	☐	☐	☐	☐	☐	☐

Med tanke på de periodene av året der du spiser fisk, hvor ofte pleier du å spise følgende til middag?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd.	2–3 pr. mnd	1 pr. uke	2+ pr. uke
Kokt torsk, sei, hyse, lyr	☐	☐	☐	☐	☐
Stekt torsk, sei, hyse, lyr	☐	☐	☐	☐	☐
Steinbit, flyndre, uer	☐	☐	☐	☐	☐
Laks, ørret	☐	☐	☐	☐	☐
Makrell	☐	☐	☐	☐	☐
Sild	☐	☐	☐	☐	☐
Annen fisk	☐	☐	☐	☐	☐

Dersom du spiser fisk, hvor mye spiser du vanligvis pr. gang? (1 skive/stykke = 150 gram)

Kokt fisk *(skive)*	☐ 1	☐ 1,5	☐ 2	☐ 3+
Stekt fisk *(stykke)*	☐ 1	☐ 1,5	☐ 2	☐ 3+

Hvor mange ganger pr. år spiser du fiskeinnmat?
(Sett ett kryss pr. linje)

	0	1–3	4–6	7–9	10+
Rogn	☐	☐	☐	☐	☐
Fiskelever	☐	☐	☐	☐	☐

Dersom du spiser fiskelever, hvor mange spiseskjeer pleier du å spise hver gang? (Sett ett kryss)

☐ 1 ☐ 2 ☐ 3–4 ☐ 5–6 ☐ 7+

Hvor ofte bruker du følgende typer fiskemat?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd.	2–3 pr. mnd	1 pr. uke	2+ pr. uke
Fiskekaker/pudding/boller	☐	☐	☐	☐	☐
Plukkfisk/fiskegrateng	☐	☐	☐	☐	☐
Frityrfisk/fiskepinner	☐	☐	☐	☐	☐
Andre fiskeretter	☐	☐	☐	☐	☐

Hvor stor mengde pleier du vanligvis å spise av de ulike rettene? (Sett ett kryss for hver linje)

Fiskekaker/pudding/boller *(stk.) (2 fiskeboller=1 fiskekake)*	☐ 1	☐ 2	☐ 3	☐ 4+
Plukkfisk, fiskegrateng *(dl)*	☐ 1–2	☐ 3–4	☐ 5+	
Frityrfisk, fiskepinner *(stk.)*	☐ 1–2	☐ 3–4	☐ 5–6	☐ 7+

I tillegg til informasjon om fiskeforbruk er det viktig å få kartlagt hvilket tilbehør som blir servert til fisk. Hvor ofte bruker du følgende til fisk?

	aldri/ sjelden	1 pr. mnd.	2–3 pr. mnd	1 pr. uke	2+ pr. uke
Smeltet/fast smør	☐	☐	☐	☐	☐
Smeltet/fast margarin	☐	☐	☐	☐	☐
Seterrømme *(35%)*	☐	☐	☐	☐	☐
Lettrømme *(20%)*	☐	☐	☐	☐	☐
Saus med fett *(hvit/brun)*	☐	☐	☐	☐	☐
Saus uten fett *(hvit/brun)*	☐	☐	☐	☐	☐

For de ulike typene tilbehør du bruker til fisk, vær vennlig å kryss av for hvor mye du vanligvis pleier å spise.

Smeltet/fast smør *(ss)*	☐ ½	☐ 1	☐ 2	☐ 3	☐ 4+
Smeltet/fast margarin *(ss)*	☐ ½	☐ 1	☐ 2	☐ 3	☐ 4+
Seterrømme *(ss)*	☐ ½	☐ 1	☐ 2	☐ 3	☐ 4+
Lettrømme *(ss)*	☐ ½	☐ 1	☐ 2	☐ 3	☐ 4+
Saus med fett *(dl)*	☐ ¼	☐ ½	☐ ¾	☐ 1	☐ 2+
Saus uten fett *(dl)*	☐ ¼	☐ ½	☐ ¾	☐ 1	☐ 2+

Hvor ofte spiser du skalldyr (f. eks. reker, krabbe og skjell)? (Sett ett kryss)

☐ Aldri/sjelden ☐ 1 pr. mnd
☐ 2–3 pr. mnd ☐ 1+ pr. uke

Hva bruker du vanligvis å steke i når du steker fisk og/eller tilbehør til fisk: (sett ett kryss)

☐ Steker uten fett ☐ Soyaolje
☐ Smør ☐ Rapsolje
☐ Fast margarin ☐ Olivenolje
☐ Flytende margarin ☐ Solsikkeolje
☐ Annen olje (spesifiser) _____

Kjøtt

Hvor ofte spiser du reinkjøtt?

☐ Aldri/sjelden ☐ 1 pr. mnd. ☐ 2–3 pr. mnd.
☐ 1 pr. uke ☐ 2–3 pr. uke ☐ 4+ pr. uke

Hvor ofte spiser du følgende kjøtt- og fjærkreretter?
(Sett ett kryss for hver rett)

	aldri/ sjelden	1 pr. mnd.	2–3 pr. mnd	1 pr. uke	2+ pr. uke
Steik *(okse, svin, får)*	☐	☐	☐	☐	☐
Koteletter	☐	☐	☐	☐	☐
Biff	☐	☐	☐	☐	☐
Kjøttkaker, karbonader	☐	☐	☐	☐	☐
Pølser	☐	☐	☐	☐	☐
Gryterett, lapskaus	☐	☐	☐	☐	☐
Pizza med kjøtt	☐	☐	☐	☐	☐
Kylling med skinn	☐	☐	☐	☐	☐
Kylling uten skinn	☐	☐	☐	☐	☐
Bacon, flesk	☐	☐	☐	☐	☐
Andre kjøttretter	☐	☐	☐	☐	☐

Dersom du spiser følgende retter, oppgi mengden du vanligvis spiser: (Sett ett kryss for hver linje)

Steik *(skiver)*	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5+
Koteletter *(stk.)*	☐ ½	☐ 1	☐ 1½	☐ 2+	
Kjøttkaker, karbonader *(stk.)*	☐ 1	☐ 2	☐ 3	☐ 4+	
Pølser *(stk. à 150g)*	☐ ½	☐ 1	☐ 1½	☐ 2+	
Gryterett, lapskaus *(dl)*	☐ 1–2	☐ 3	☐ 4	☐ 5+	
Pizza m/kjøtt *(stykke à 100 g)*	☐ 1	☐ 2	☐ 3	☐ 4+	

Hvilke sauser bruker du til kjøttretter og pastaretter?
(Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd.	2–3 pr. mnd	1 pr. uke	2+ pr. uke
Brun saus	☐	☐	☐	☐	☐
Sjysaus	☐	☐	☐	☐	☐
Tomatsaus	☐	☐	☐	☐	☐
Saus med fløte/rømme	☐	☐	☐	☐	☐

Hvor mye bruker du vanligvis av disse sausene?

Brun saus *(dl)*	☐ ¼	☐ ½	☐ ¾	☐ 1	☐ 2+
Sjysaus *(dl)*	☐ ¼	☐ ½	☐ ¾	☐ 1	☐ 2+
Tomatsaus *(dl)*	☐ ¼	☐ ½	☐ ¾	☐ 1	☐ 2+
Saus med fløte/rømme *(dl)*	☐ ¼	☐ ½	☐ ¾	☐ 1	☐ 2+

Andre matvarer

Hvor mange egg spiser du vanligvis i løpet av en uke? (stekte, kokte, eggerøre, omelett) (Sett ett kryss)

☐ 0 ☐ 1 ☐ 2 ☐ 3–4 ☐ 5–6 ☐ 7+

Hvor ofte spiser du iskrem? (til dessert, Krone-is osv.)
Sett ett kryss for hvor ofte du spiser iskrem om sommeren, og ett kryss for resten av året

	aldri/ sjelden	1 pr. mnd.	2–3 pr. mnd	1 pr. uke	2+ pr. uke
Om sommeren	☐	☐	☐	☐	☐
Resten av året	☐	☐	☐	☐	☐

Hvor mye is spiser du vanligvis pr. gang?
(Sett ett kryss)

☐ 1 dl ☐ 2 dl ☐ 3 dl ☐ 4+ dl

Hvor ofte spiser du bakevarer som boller, kaker, wienerbrød eller småkaker (Sett ett kryss pr. linje)

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–3 pr. uke	4–6 pr. uke	1+ pr. dag
Gjærbakst (boller o.l.)	☐	☐	☐	☐	☐	☐
Wienerbrød, kringle	☐	☐	☐	☐	☐	☐
Kaker	☐	☐	☐	☐	☐	☐
Pannekaker	☐	☐	☐	☐	☐	☐
Vafler	☐	☐	☐	☐	☐	☐
Småkaker, kjeks	☐	☐	☐	☐	☐	☐
Lefser, lomper	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du dessert? (Sett ett kryss pr. linje)

	aldri/ sjelden	1 pr. mnd	2–3 pr. mnd	1 pr. uke	2–3 pr. uke	4+ pr.uke
Pudding sjokolade/karamell	☐	☐	☐	☐	☐	☐
Riskrem, fromasj	☐	☐	☐	☐	☐	☐
Kompott, fruktgrøt, hermetisk frukt	☐	☐	☐	☐	☐	☐
Jordbær *(friske, frosne)*	☐	☐	☐	☐	☐	☐
Andre bær *(friske, frosne)*	☐	☐	☐	☐	☐	☐

Hvor ofte spiser du sjokolade? (Sett ett kryss)

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–3 pr. uke	4–6 pr. uke	1+ pr.dag
Mørk sjokolade	☐	☐	☐	☐	☐	☐
Lys sjokolade	☐	☐	☐	☐	☐	☐

Dersom du spiser sjokolade, hvor mye pleier du vanligvis å spise hver gang? Tenk deg størrelsen på en Kvikk-Lunsj sjokolade, og oppgi hvor mye du spiser i forhold til den.

☐ ¼ ☐ ½ ☐ ¾ ☐ 1 ☐ 1½ ☐ 2+

Hvor ofte spiser du snacks? (Sett ett kryss)

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–3 pr. uke	4–6 pr. uke	1+ pr.dag
Potetchips	☐	☐	☐	☐	☐	☐
Peanøtter	☐	☐	☐	☐	☐	☐
Andre nøtter	☐	☐	☐	☐	☐	☐
Annen snacks	☐	☐	☐	☐	☐	☐

Tran og fiskeoljekapsler

Bruker du tran (flytende)?

☐ Ja ☐ Nei

Hvis ja; hvor ofte tar du tran? Sett ett kryss for hver linje.

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–6 pr. uke	daglig
Om vinteren	☐	☐	☐	☐	☐
Resten av året	☐	☐	☐	☐	☐

Hvor mye tran pleier du å ta hver gang?

☐ 1 ts ☐ ½ ss ☐ 1+ss

Bruker du tranpiller/fiskeoljekapsler?

☐ Ja ☐ Nei

Hvis ja; hvor ofte tar du tranpiller/fiskeoljekapsler?
Sett ett kryss for hver linje.

	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–6 pr. uke	daglig
Om vinteren	☐	☐	☐	☐	☐
Resten av året	☐	☐	☐	☐	☐

Hvilken type tranpiller/fiskeoljekapsler bruker du vanligvis, og hvor mange pleier du å ta hver gang?

Navn: _____

Antall: ☐ 1 ☐ 2 ☐ 3+

Kosttilskudd

Bruker du kosttilskudd (vitaminer/mineraler)?
☐ Ja ☐ Nei

Hvis ja, hvor ofte bruker du kosttilskudd?
(Sett ett kryss pr. linje)

Navn på kosttilskudd	aldri/ sjelden	1–3 pr. mnd	1 pr. uke	2–6 pr. uke	daglig
_____	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐

Alkohol

Er du totalavholdskvinne?
☐ Ja ☐ Nei

Hvis Nei; hvor ofte og hvor mye drakk du i gjennomsnitt siste året? (Sett ett kryss for hver linje)

	aldri/ sjelden	1 pr. mnd	2–3 pr. mnd.	1 pr. uke	2–4 pr. uke	5–6 pr. uke	1 pr. dag	2+ pr. dag
Øl *(½ l.)*	☐	☐	☐	☐	☐	☐	☐	☐
Vin *(glass)*	☐	☐	☐	☐	☐	☐	☐	☐
Brennevin *(drink)*	☐	☐	☐	☐	☐	☐	☐	☐
Likør/Hetvin *(glass)*	☐	☐	☐	☐	☐	☐	☐	☐

Sosiale forhold

Hvor mange personer er det i ditt hushold?
☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5+

Hvor høy er bruttoinntekten i husholdet pr. år?
☐ inntil 150.000 kr.
☐ 151.000–300.000 kr.
☐ 301.000–450.000 kr
☐ 451.000–600.000 kr.
☐ 601.000–750.000 kr.
☐ 751.000–900.000 kr.
☐ over 900.000 kr.

Solvaner

Hvor mange ganger pr. år er du blitt forbrent av solen slik at du har fått svie eller blemmer med avflassing etterpå?

Årstall	Aldri	Høyst 1 g. pr. år	2–3 g. pr. år	4–5 g. pr. år	6 eller flere g. pr. år
2003–2011	☐	☐	☐	☐	☐

Hvor mange uker i gjennomsnitt pr. år har du vært på badeferie 2003–2011?

Årstall	Aldri	1 uke	2–3 uker	4–5 uker	7 uker eller mer
I syden	☐	☐	☐	☐	☐
I Norge	☐	☐	☐	☐	☐

Hvor ofte har du solt deg i solarium?

Årstall	Aldri	Sjelden	1 g. pr. mnd.	2 g. pr. mnd.	3–4 g. pr. mnd.	Oftere enn 1 g. pr. mnd.
2003–2011	☐	☐	☐	☐	☐	☐

Hvor ofte dusjer eller bader du?

	mer enn 1 g. dagl	1 g. dagl.	4–6 g. pr. uke	2–3 g. pr. uke	1 g. pr. uke	2–3 g. pr. mnd.	sjelden/ aldri
Med såpe/ shampo	☐	☐	☐	☐	☐	☐	☐
Uten såpe/ shampo	☐	☐	☐	☐	☐	☐	☐

Når bruker du krem med solfaktor? (sett evt. flere kryss):
☐ i påsken
☐ i Norge eller utenfor syden
☐ solferie i syden
☐ aldri

Hvilken solfaktor bruker du i disse periodene?

Faktor	Ingen	1–4	5–9	10–14	15–29	30+
Påsken	☐	☐	☐	☐	☐	☐
I Norge eller utenfor syden	☐	☐	☐	☐	☐	☐
Solferie i syden	☐	☐	☐	☐	☐	☐

Hvor ofte bruker du følgende hudpleiemidler?
(Sett ett kryss pr. linje)

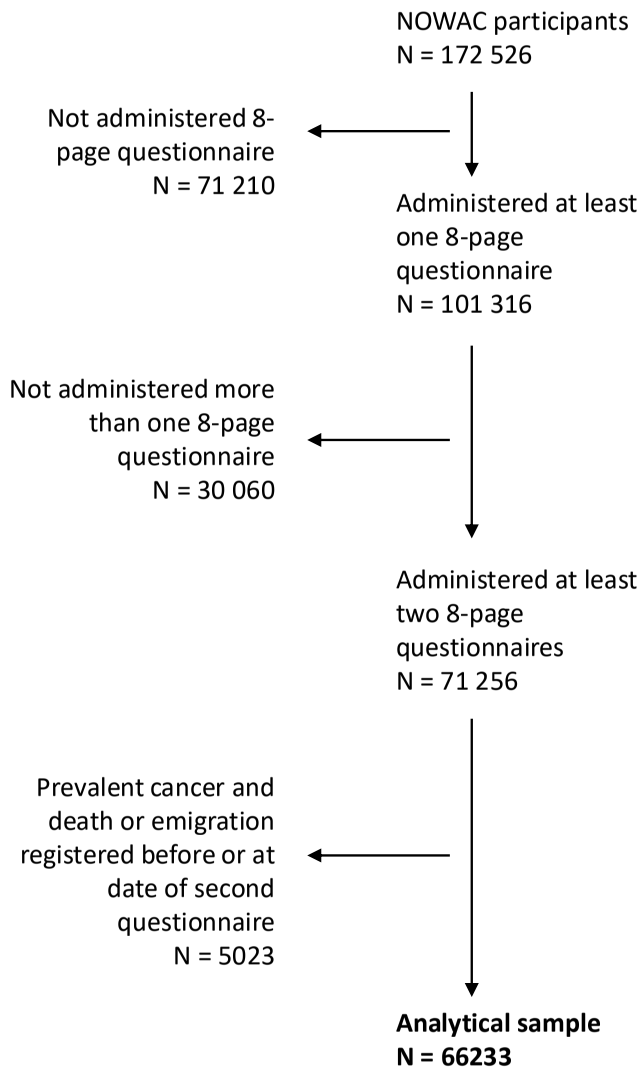
	aldri/ sjelden	1–3 pr. mnd.	1 pr. uke	2–4 pr. uke	5–6 pr. uke	1 pr. dag	2+ pr. dag
Ansiktskrem	☐	☐	☐	☐	☐	☐	☐
Håndkrem	☐	☐	☐	☐	☐	☐	☐
Body lotion	☐	☐	☐	☐	☐	☐	☐
Parfyme	☐	☐	☐	☐	☐	☐	☐

Til slutt vil vi spørre deg om ditt samtykke til å kontakte deg på nytt pr. post. Vi vil hente adressen fra det sentrale personregister ☐ Ja ☐ Nei

Er du villig til å avgi en blodprøve? ☐ Ja ☐ Nei

Takk for at du ville delta i undersøkelsen

Additional File 3.
Sample flow chart, Norwegian Women and Cancer Study (NOWAC)



Additional File 4

Description of healthy lifestyle index (HLI) construction in the Norwegian Women and Cancer study (NOWAC)

Physical activity level was reported by participants on a 10-point scale ranging from very little activity to very active, where participants were asked to consider the entirety of activity at work, outside work, at home, exercise, and other forms of physical activity. Since this measure could not be categorized according to physical activity guidelines or other measures of the dose of physical activity, physical activity level was scored by quintile based on the percentile distribution at Q1 (physical activity scale 7-10 = 4, 6 = 3, 5 = 2, 4 = 1, 1-3 = 0). Body fatness was assessed through self-reported height (centimeters) and weight (kilograms) to calculate body mass index (kg/m^2) (BMI <23 = 4, 23 to <25 = 3, 25 to <27 = 2, 27 to <30= 1, ≥30 = 0), smoking was scored considering smoking status, smoking intensity and time since cessation (never smoker = 4, former smoker >10 years since cessation = 3, former smoker ≤10 years since cessation = 2, smoker <15 cigarettes/day = 1, current smoker ≥15 cigarettes/day = 0), and alcohol (ethanol) consumption was recorded in grams/day (none = 4, >0 to <5 = 3, 5 to <10 = 2, 10 to <20 = 1, >20 = 0). A diet score ranging from 0 to 18 (healthiest) was generated, comprising six food groups: whole grains, fruit, vegetables, dairy, red meat, and processed meat. Using an analysis program developed at the Institute of Community Medicine, UiT The Arctic University of Norway, daily intake of food groups and energy were computed based on the frequencies and portions of food items reported in the FFQ according to the food composition table for Norway (1). Each food group was adjusted for energy intake, by dividing grams of intake by daily energy intake, in millijoules (MJ). The energy-adjusted food groups were categorized into quartiles and scored from 0 (lowest quartile) to 3 (highest quartile). Red and processed meat were scored in reverse order. The 18-point diet score was then divided into quintiles to produce a score ranging from 0 to 4 for inclusion in the HLI (11-18 = 4, 10 = 3, 8-9 = 2, 7 = 1, 0-6 = 0).

References

1. The Norwegian Food Composition Database 2021 [Internet]. 2021. Available from: https://www.matvaretabellen.no/.

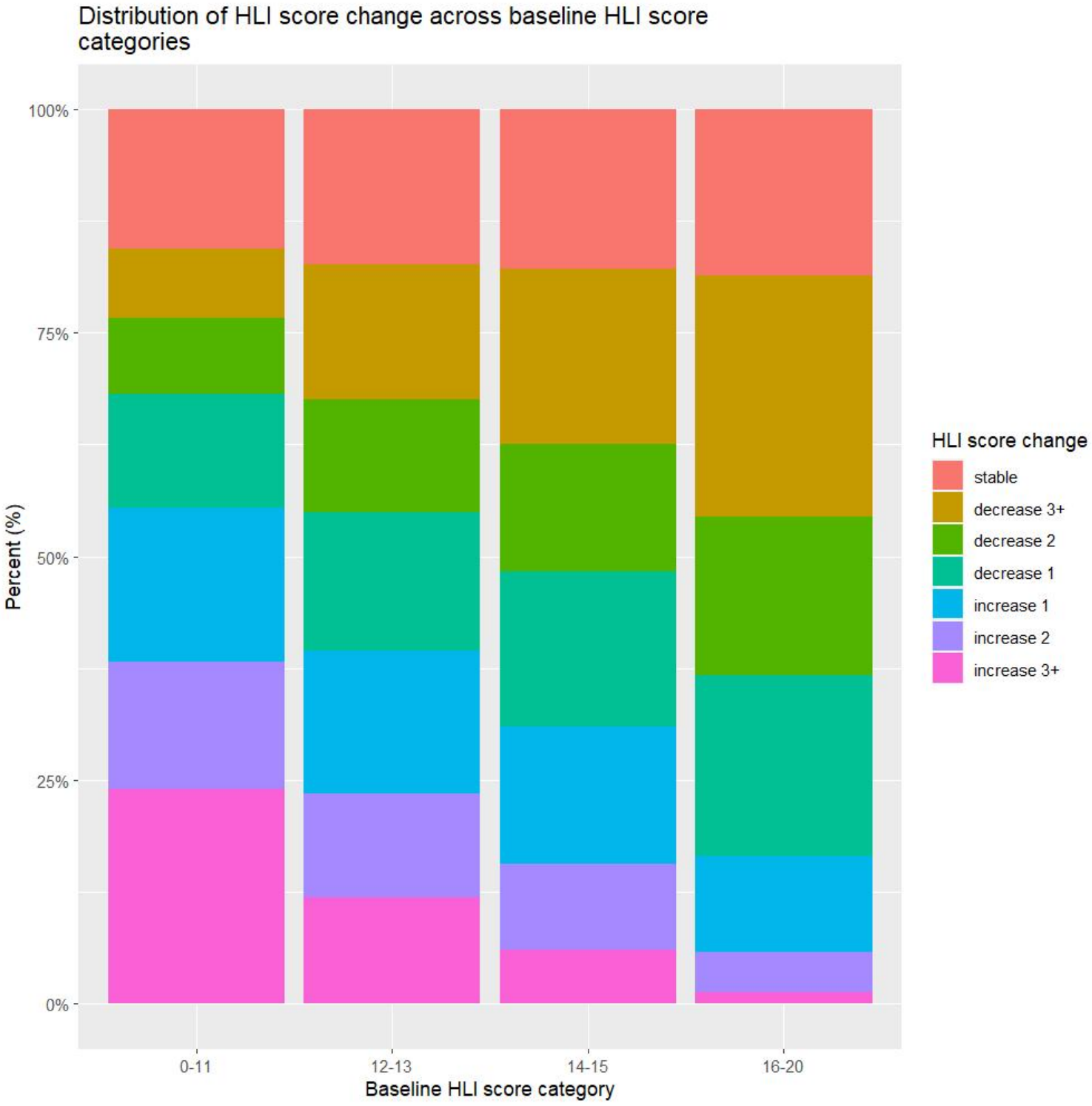
Additional File 5. Cancer types and associated International Classification of Disease, tenth revision (ICD-10) codes included in the study

ICD-10 code	N cases	Site	Exposure-related cancer subgroups				
			Lifestyle	Alcohol	Tobacco	Obesity	Reproductive
C01	12	Upper aerodigestive	■	■	■		
C02	18		■	■	■		
C03	6		■	■	■		
C04	4		■	■	■		
C05	5		■	■	■		
C06	0		■	■	■		
C07	11		■	■	■		
C09	24		■	■	■		
C10	1		■	■	■		
C11	4	Pharynx	■	■	■		
C12	1		■	■	■		
C13	0		■	■	■		
C14	1		■	■	■		
C15	33	Esophagus	■	■	■	■	
C16	113	Stomach	■		■	■	
C18	858	Colorectum	■	■	■	■	
C19	53		■	■	■	■	
C20	301		■	■	■	■	
C22	43	Liver	■	■	■	■	
C23	16		■	■	■	■	
C24	31		■	■	■	■	
C25	214	Pancreas	■		■	■	
C31	6	Accessory sinus	■		■		
C32	9	Larynx	■	■	■		
C33	1	Trachea	■		■		
C34	763	Lung	■		■		
C50	2384	Breast	■	■	■	■	

C51	32	Vulva					
C52	10	Vagina					
C53	79	Cervix					
C54	607	Uterine					
C55	0						
C56	268	Ovarian					
C57	75	Other female genital organs					
C58	0						
C64	189	Kidney					
C65	19						
C66	5						
C73	116	Thyroid					
C67	64	Bladder					
C90	145	Multiple myeloma					
C92	60	Acute myeloid leukemia					

Shaded regions denote cancer diagnosis belonging to the exposure-related cancer subgroup

Additional File 6



Additional File 7.

Associations between healthy lifestyle index score change and lifestyle-related, alcohol-related, tobacco-related, obesity-related, reproductive-related, breast, and colorectal cancer incidence in the Norwegian Women and Cancer Study (n=44404), complete-case analysis

		Lifestyle-related cancer incidence[a]	**Alcohol-related cancer incidence[a]**	**Tobacco-related cancer incidence**	**Obesity-related cancer incidence[a]**	**Reproductive-related cancer incidence[a]**	**Breast cancer incidence[a]**	**Colorectal cancer incidence**
Cases		3926	2348	1805	3309	611	1567	468
Continuous HLI score change	1-SD (2.4 HLI points) increase	0.94(0.92-0.97)	0.97(0.94-1.01)	0.94(0.90-0.98)	0.96(0.93-0.99)	0.91(0.85-0.98)	0.97(0.83-1.01)	1.00(0.93-1.09)
Categorical HLI score change	<= -3	1.18(1.06-1.32)	1.07(0.93-1.24)	1.27(1.08-1.50)	1.12(0.99-1.26)	1.31(0.99-1.72)	1.06(0.89-1.26)	1.25(0.91-1.73)
	-2	1.12(1.00-1.27)	1.16(1.00-1.34)	1.16(0.97-1.38)	1.12(0.98-1.27)	1.14(0.85-1.53)	1.14(0.95-1.36)	1.14(0.80-1.61)
	-1	1.03(0.92-1.14)	1.04(0.90-1.20)	1.07(0.91-1.27)	0.99(0.88-1.12)	0.96(0.72-1.28)	1.03(0.87-1.23)	1.10(0.79-1.53)
	0	1.00 (ref)						
	1	0.98(0.87-1.09)	0.96(0.83-1.11)	1.09(0.93-1.29)	0.96(0.85-1.08)	0.98(0.74-1.30)	0.91(0.76-1.09)	1.19(0.86-1.65)
	2	0.92(0.81-1.04)	0.98(0.83-1.15)	0.97(0.80-1.16)	0.95(0.83-1.08)	0.84(0.60-1.16)	0.96(0.79-1.16)	1.07(0.74-1.54)
	>=3	0.97(0.86-1.09)	0.98(0.84-1.15)	1.01(0.85-1.21)	0.96(0.84-1.09)	1.05(0.78-1.41)	0.94(0.78-1.14)	1.21(0.86-1.72)
HLI score change excluding one factor[b]	1-SD increase							
Excluding physical activity	1.9	0.96(0.93-0.99)	0.94(0.93-1.02)	0.96(0.91-1.01)	0.96(0.93-1.00)	0.94(0.86-1.02)	0.97(0.92-1.03)	0.99(0.90-1.09)
Excluding BMI	2.2	0.93(0.90-0.96)	0.97(0.93-1.01)	0.88(0.83-0.92)	0.97(0.94-1.01)	0.97(0.89-1.06)	0.97(0.92-1.03)	0.99(0.90-1.10)
Excluding smoking	2.4	0.95(0.92-0.98)	0.97(0.93-1.02)	0.98(0.94-1.04)	0.95(0.91-0.98)	0.87(0.79-0.94)	0.96(0.90-1.01)	1.02(0.92-1.13)
Excluding alcohol	2.3	0.93(0.90-0.96)	0.98(0.94-1.02)	0.93(0.88-0.97)	0.95(0.91-0.98)	0.86(0.79-0.94)	0.98(0.92-1.03)	1.01(0.92-1.11)
Excluding diet	1.9	0.93(0.90-0.96)	0.96(0.92-1.00)	0.93(0.89-0.98)	0.95(0.91-0.98)	0.92(0.84-1.00)	0.95(0.90-1.00)	1.01(0.91-1.11)
Single HLI factors[c]	1-unit increase (score 0-4)							
Physical activity score change		0.96(0.93-0.98)	0.98(0.95-1.02)	0.96(0.93-1.00)	0.98(0.95-1.01)	0.93(0.87-1.00)	0.97(0.93-1.01)	1.03(0.95-1.11)
BMI score change		0.97(0.92-1.01)	0.98(0.92-1.03)	1.03(0.97-1.10)	0.93(0.89-0.98)	0.84(0.76-0.94)	0.95(0.89-1.02)	1.01(0.89-1.14)
Smoking score change		0.98(0.93-1.04)	1.00(0.92-1.09)	0.94(0.86-1.02)	1.01(0.95-1.09)	1.08(0.91-1.27)	1.02(0.93-1.13)	1.05(0.87-1.26)
Alcohol score change		1.00(0.95-1.05)	0.93(0.87-0.99)	1.02(0.95-1.10)	0.99(0.93-1.04)	1.18(1.03-1.34)	0.91(0.85-0.99)	0.99(0.85-1.14)
Diet score change		0.99(0.97-1.02)	1.01(0.97-1.04)	1.00(0.96-1.03)	1.00(0.97-1.02)	0.98(0.92-1.04)	1.00(0.96-1.04)	1.01(0.94-1.08)

Footnotes:

All models were adjusted for education (years), height (centimetres), HLI score at Q1 (continuous), and calendar year at Q2 (continuous).
[a]Models additionally adjusted for age at menarche (years), menopausal status (premenopausal/postmenopausal), breastfeeding (cumulative months 0, <=12, >12), hormone replacement therapy use (never/former/current), oral contraceptive use (never/ever), parity (0, 1-2, >2), and

history of breast cancer in a first degree relative (yes/no). [b]Baseline HLI score was adjusted by separately adjusting for HLI score at Q1 excluding the factor in question and the individual factor score at Q1. [c]Mutually adjusted for all single factor HLI score changes and single factor HLI scores at Q1.

Alcohol-related cancers including sites: upper aerodigestive [C01-C10], pharynx [C11-C14], esophagus [C15], colorectum [C18-C20], liver [C22-C24], larynx [C32], breast [C50],

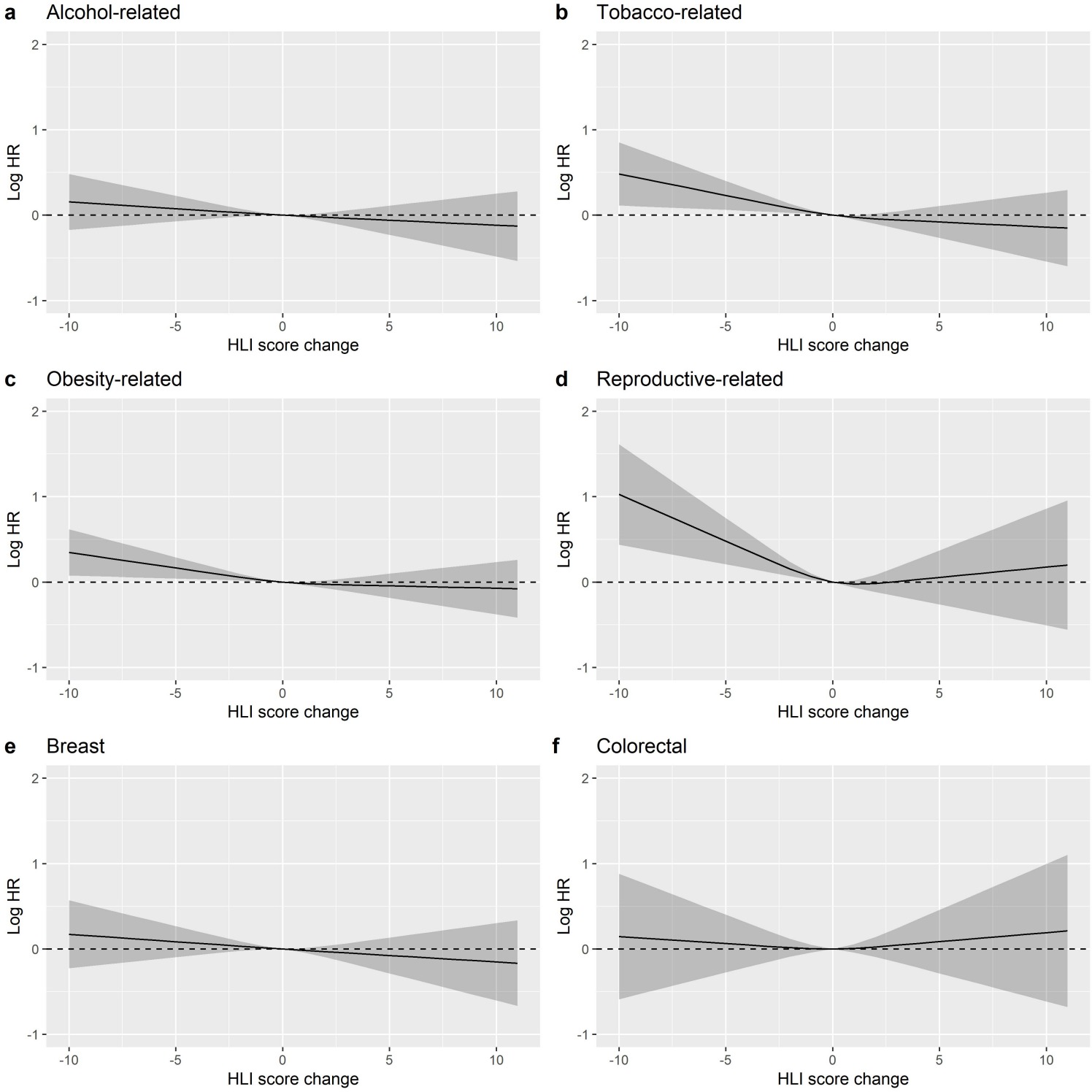
Tobacco-related cancers including sites: upper aerodigestive [C01-C10], pharynx [C11-C14], esophagus [C15], stomach [C16], colorectum [C18-C20], liver [C22-C24], pancreas [C25], accessory sinus [C31], larynx [C32], trachea [C33], lung [C34], breast [C50], cervix [C53], ovarian [C56], kidney [C64-C66], bladder [C67], acute myeloid leukemia [C92]

Obesity-related cancers including sites: esophagus [C15], stomach [C16], colorectum [C18-C20], liver [C22-C24], pancreas [C25], breast [C50], uterine [C54-C55], ovarian [C56], kidney [C64-C66], thyroid [C73], multiple myeloma [C90],

Reproductive-related cancers including sites: vulva [C51] vagina [C52], cervix [C53], uterine [C54-C55], ovarian [C56], other female genital organs [C57-C58],

Additional File 8

Association between HLI score change and incidence of exposure-related cancer subgroups, modelled with restricted cubic splines



Paper III

Appendix

Information letter, Norwegian Women and Cancer study, 2003

INSTITUTT FOR SAMFUNNSMEDISIN
UNIVERSITETET I TROMSØ
9037 TROMSØ
Telefon 77 64 48 16/77 64 66 38

KVINNER OG KREFT

Institutt for samfunnsmedisin ved Universitetet i Tromsø gjennomfører en spørreundersøkelse om levesett og kreft blant norske kvinner. En slik undersøkelse gir et verdifullt grunnlag for å studere mulige sammenhenger mellom f.eks. kosthold, barnefødsler, p-piller, solvaner og utviklingen av kreft. Resultatet vil bli publisert i dagspressen og i internasjonale fagtidsskrifter. Ansvarlig for undersøkelsen er professor Eiliv Lund.

Du forespørres hermed om å delta i undersøkelsen. Alle som blir forespurt er trukket ut tilfeldig. Statistisk Sentralbyrå har trukket utvalget og står for utsending av spørreskjemaene.

Med noen års mellomrom fram til 2033 ønsker vi å sammenholde opplysningene som er gitt i undersøkelsen mot opplysninger fra Kreftregisteret, Mammografiregistrert og Dødsårsaksregisteret. Samtykket fra deg for dette vil være ensbetydende med returnering av spørreskjemaet.
Alle opplysninger fra undersøkelsen og fra registrene vil bli behandlet konfidensielt og etter regler Datatilsynet har gitt i sin tillatelse, samt tillatelse fra Sosial- og helsedirektoratet. På spørreskjemaet er navn og fødselsnummer erstattet med et løpenummer slik at ingen av de som mottar og tar hånd om skjemaene vil kjenne din identitet. Undersøkelsen er tilrådd av Regional komite for medisinsk forskningsetikk i Nord-Norge.

Hvis du vil delta i undersøkelsen, ber vi deg om å besvare det vedlagte spørreskjemaet så riktig som mulig. Dersom ingen av de oppgitte svaralternativ dekker din situasjon, sett kryss for det alternativet som ligger nærmest. Gi eventuelle tilleggsopplysninger i skjemaet.
Du behøver ikke svare på alle spørsmål.

Det vil senere bli aktuelt å samle inn blodprøver fra noen av deltakerne.
Dette vil skje hos nærmeste lege, og vil være gratis. Det vil også bli aktuelt å spørre noen av deltakerne om å være med på et kostholdsintervju over telefon. Bare de av deltakerne som på forhånd har krysset av for at de er villig til å bli kontaktet på nytt og/eller til å bli spurt om å avgi blodprøve, vil få henvendelse om dette. Det vil da bli gitt nærmere informasjon og innhentet samtykke til dette.

Det er frivillig om du vil være med i undersøkelsen. Det er også adgang til å trekke seg senere, hvis du skulle ønske det. Du kan få slettet dine opplysninger hvis du krever det. De innsamlete opplysninger vil bli anonymisert 31.12.2033.

Ditt bidrag til undersøkelsen vil være å svare på spørsmålene i spørreskjemaet.
For spørsmål om hormoner og p-pille bruk finner du bilder i denne brosjyren som skal være et hjelpemiddel til å svare riktig (brosjyren skal ikke returneres). Spørreskjemaet returneres i vedlagte konvolutt med betalt svarporto.

Med vennlig hilsen

Eiliv Lund
professor dr.med.

Bente A. Augdal
prosjektmedarbeider

Du kan finne mer informasjon om "Kvinner og kreft" og om forskningsresultatene på våre nettsider: www.ism.uit.no/kk/