UiT The Arctic University of Norway

Faculty of Health Science – The Department of Community Medicine

Self-reported diabetes mellitus and education in NOWAC, a cross-sectional study

MPH student: Jane Skoglund, Supervisor Dr. Kristin Borch Benjaminsen, co-supervisor postdoc Marko Lukic

Master thesis in Public Health (MPH), HEL-3950, May 2022

Preface

It has been a great journey being a student at the master in public health program at UiT. I have extended my knowledge regarding public health intervention and grown as a person. My previous experience as a nurse together with increased interest throughout MPH-program, made med explore the field of diabetes and social inequalities.

I would like to thank my competent supervisor Dr. Kristin Borch Benjaminsen and co-supervisor Marko Lukic for their patience and willingness to contribute after I had to change the topic for my thesis, and my supervisors after 3rdsemester. I am thankful for the support, their eagerness to shear knowledge, their time, feedback and encouragement throughout this process.

My supervisors have guided me towards the thesis, but without the support from my husband Thor Marius Holte Skoglund, and patients from my kids, Casper and Iben, this project wouldn´t have succeeded.

Finally, I would like to thank my fellow classmate and friend forever, Brita Marie Pedersen, I´m forever grateful for our friendship, and for your backing towards my thesis and always in life. The master program wouldn´t have been the same without you.

Table of Contents

List of Tables

List of Figures

Abstract

Introduction: Diabetes mellitus is an increasingly public health concern, causing increased mortality, morbidity and healthcare costs globally. Global increase in overweight and obesity, high calorie diets, sedentary lifestyle, smoking habits and aging population are the main drivers for this epidemic. Social inequalities reveal systematic differences in health and disease. Social inequalities are increasing in Norway, especially among women. Persons with lower socio-economic position are more often exposed to negative lifestyle behavior. Education are one of several measures for SEP used in epidemiological studies. This study aims to investigate the association between educational levels and diabetes among Norwegian female population.

Material and methods: This study used data from the Norwegian Women and Cancer cohort that included participants between 30-70 years of age, recruited in waves in 1996, 1997, 1998, 2003 and 2004 (N=100 347). Questionnaires collected at enrolment was used. Complete-case analysis was performed, where 21.1% were excluded due to missing information, resulting in N=79 221 eligible participants. Diabetes was self-reported, and education level grouped ("primary" ≤9 years, "secondary" 10-12 years of education, and "tertiary" ≥13 years of education). Binary logistic regression analysis was performed with self-reported diabetes as the dependent variable, and education as the main independent variable with "primary as reference category. The logistic regression analyses were performed with three models: A=univariable, B= age-adjusted and C=multivariable (adjusted for: age, BMI, physical activity, smoking status, alcohol consumption, CVD and diet (total intake of fiber, vegetables, fruit and total sugar)).

Results: In the unadjusted model A, the results showed that participants with tertiary education had lower odds of being diagnosed with diabetes compared to those with primary education: OR:0.49 (95% CI: 0.43-0.57), similar is seen for secondary compared to primary education: OR:0.69 (95% CI: 0.6-0.8). After age-adjustments, those with tertiary education had lower odds of being diagnosed with diabetes compared to those with primary education OR:0.65 (95% CI: 0.56-0.76), secondary compared to primary gave an OR of 0.86 (95% CI: 0.74-1.0). After adjusting for covariates, none of the results were significant; tertiary compared to primary OR: 0.91(95% CI:0.78-1.08), and secondary compared to primary OR: 1.02 (95% CI: 0.87-1.19). Thus, the results showed a strong interference of covariates.

Conclusion: This study found an educational gradient in self-reported diabetes within the Norwegian female population, but the association was strongly confounded by different factors, suggesting that educational level does not predict diabetes. Educational is assumed to be a proxy for lifestyle factors, which predict diabetes. The inconsistency regarding diabetes prevalence induces the urgent need for a diabetes register in Norway.

Keywords: Diabetes, education, socio-economic position

Abbreviations

BMI	Body Mass Index
FFQ	Food Frequency questionnaire
GDM	Gestational Diabetes Mellitus
NCD	Non-communicable diseases
NiPH	The Norwegian Institute of Public Health
NOWAC	The Norwegian Woman and Cancer cohort study
NSD	Norwegian Data Inspectorate
OR	Odds Ratio
PPV	Positive Predicted Value
REC	Regional Committee for Medical Research Ethics
SD	Standard Deviation
SEP	Socioeconomic Position
SPSS	Statistical Package for Social Sciences
SR	Systematic review
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
WHO	World Health Organization

1 Introduction

Diabetes mellitus (hereafter diabetes) is a rapidly increasing epidemic causing increased mortality and morbidity worldwide (1). Diabetes is one of the three major non-communicable diseases (NCDs), among cancer and heart diseases, that contributes to 70% of deaths worldwide, and, according to the World Health Organization (WHO), these NCDs are the largest threats to global public health (2). Diabetes is a public health problem, where 1 in 10 adults are living with the condition (3). The main risk factors contributing to these deaths include tobacco use, physical inactivity, binge drinking, unhealthy diets, and air pollution (2). These risk factors have a higher prevalence among persons of lower socioeconomic position (SEP). Thus, SEP is an important predictor for health and well-being (4).

1.1 Diabetes

Diabetes is a disease caused by insufficient insulin production or an abnormal response to insulin(1). Insulin is essential for nutrition uptake; deficiency of insulin causes elevated levels of blood glucose, and long-term elevated blood glucose levels applies damage to the heart, vasculature, kidneys, eyes, and nerves (1). Diabetes is mainly classified by three main types: type 1 diabetes mellitus(T1DM), type 2 diabetes mellitus (T2DM) and gestational diabetes mellitus (GDM;(5). Type 1 diabetes mellitus is characterized by deficient or almost deficient insulin production and requires administration of insulin. Onset is most common in childhood and adolescent years, and it cannot be prevented (1). Gestational diabetes mellitus occurs during pregnancy, and while this state is transient, women experience gestational diabetes might have an increased risk of developing T2DM later in life (1).

Type 2 diabetes mellitus accounts for 90 % of all cases and is, therefore, the most common type (5). It is traditionally diagnosed later in life, and while it is most common among older adults, it is also seen among younger people with obesity and/or other risk factors (1). It is feasible to prevent or delay T2DM (1), and research suggests a possibility for remission, in some cases (6). Family history of diabetes, elevated, body mass index (BMI), and high triglyceride levels are found among individuals at high risk of developing T2DM (7, 8). The risk factors for developing T2DM is often divided into non-modifiable factors, such as ethnicity, family history, and genetic predisposition, and modifiable risk factors, such as obesity, low physical activity levels, and unhealthy diet (9).

Diabetes is a global concern, as it causes premature mortality, decreased utility, and increase healthcare costs worldwide (10, 11). It is estimated that 536.6 million people live with diabetes, and that diabetes contributed to over 6.7 million deaths worldwide in 2021(3). These numbers are predicted to rise rapidly in the future (3). The rapid increase is attributed to individual risk factors, environmental risk factors, detection efficiency, the evolution of disease, and global changes (12). Diabetes is highly prevalent worldwide, independent of the economic situation in the country (5). Global increases in overweight and obesity, high calorie diets, sedentary lifestyles, smoking habits, and aging populations are the main drivers of the T2DM epidemic (9, 13).

Approximately 270.000 people, or five percent of the population, are diagnosed with diabetes in Norway (14). Additionally, it is suggested that around 60.000 lives with undiagnosed diabetes (14). Norway ranks in the top three in diabetes-related health expenditures ($ 11,166 USD per person with diabetes) globally (3).

1.2 Socioeconomic position

Kriger et al. (15) defined SEP as “an aggregate concept that includes both resource-based and prestige-based measures, as linked to both childhood and adult social class position” (15p. 345). SEP describes the social and economic factors influencing the position individuals or groups have within the social structure (16). There are numerous indicators for SEP; education, income, and occupation, or a combination thereof are the most used proxies in epidemiology (17). Educational levels correspond to social status, income corresponds to economic status, and occupation corresponds to work prestige (18).

Investigating differences in socioeconomic position reveals social inequalities in health (19). Social inequalities in health expresses systematic differences in health and disease, when comparing social and economic categories, especially among the most used proxies (19). These inequalities in health is found in every step of the socioeconomic ladder (20). Hence, this occurs when comparing the richest against the poorest, but also when comparing the richest against the next richest. The major determinants of health are material circumstances, psychosocial circumstances, behavioral and biological factors, and the health system (15). These determinants are found to be distributed unevenly along the socioeconomic ladder (21).

People of lower SEP may not have easy access to affordable, nutritious foods in their area nor the cooking skills or time to prepare them, which will affect dietary habits (19). This can be a result of multiple time-depending jobs and/or lack of knowledge. Multiple jobs may also affect leisure time, and thereafter their physical activity level. Physical activity level can also be affected from access to green space, or the fact that people with lower SEP more often live in smaller apartments in the middle of cities (12). Several jobs and economic concerns may affect stress levels, which again can lead to smoking and alcohol abuse (15). Persons with lower SEP are more likely to be exposed to more air-pollution, which can cause numerous health issues like asthma, and again affect their ability to exercise (12). Those with higher BMI may experience weight stigma and bias in healthcare settings, which may affect the level of care they receive and may prevent them from seeking regular checkups, both which can exacerbate medical problems (22). People in lower SEP groups may also need to get a job as students to help support their future or to help support their families. This could affect their ability to study or pay attention at school, along with healthy nutrition intake(19).

There is evidence for an inverse association between SEP and total mortality, which cannot be fully explained by risk factors alone (23). Hence, the social differences and social stratifications are pathogenic in themselves, and relative poverty can be an important determinants of health (19). This social gradient in health is complex and not fully understood. SEP is related to material standard of living. However, low SEP is related to a wide range of psychosocial and material factors, such as poorer housing, poorer working conditions, poorer diet, more exposure to air pollution, exposure to crime, and dangerous environments (15). Parts of the explanation are seen in lifestyle factors. Thus, individuals with lower SEP are most often exposed to negative lifestyle habits including smoking, inactivity, obesity, and low fruit and vegetables intake (24).

Education, income, and occupation are correlated, but measures different phenomena (25). Education is found to be a suitable indicator used in epidemiological studies. It is a rich source of data, because it reflects early life SEP, is usually stable over time, is a strong determinant of employment and income, and affects values, cognitive decision making and risk-taking behavior (17, 26, 27).

1.3 Health inequalities in Norway

Norway is ranked highly in international indexes for measuring welfare and development, such as the human development index (28) and the sustainable development goal index (29). This is due to its wealthy economic situation, comprehensive welfare system, and a trustworthy government (28). Norway provides 13 years of free education, and 82% of the population has completed upper secondary education and 35.3% of the population had attained higher education in 2020 (30). The average household net-adjusted income per capita is $35.725 USD per year, which scores above the Organization for Economic Co-operation and Development average (31). Overall, health in Norway is ranked as good, and life expectancy among women was 84.3 years in 2017 which is greater than The United Nations estimate of the global average of 72.6 years in 2019 (13).

Despite this, social inequalities are increasing in Norway among all genders and at all ages (19), especially among women in Norway (32). The gap between the richest and the poorest is significant, where the top 20% has four times higher income than the poorest 20% (31). Individuals with higher education have the highest life expectancy, and this gap is increasing. Social inequalities are seen in all diseases, injuries and disorders, in addition to lifestyle choices (32). The social inequalities are surprisingly greater in Norway, than other European countries (33).

Reduction in social inequality in health is central to the United Nations Sustainable Development goals, expressing “leave no one behind” (34). As well as for WHO global action plan for prevention and control of NCDs (2). Nonetheless, Norwegian Institute of Public Health (NIoPH) has set their focus on social inequalities. NIoPH published three aims for public health work in 2018, where the second aim states: “the Norwegian population shall experience more years of good health and well-being and reduce social inequalities in health” (13p.9).

1.4 Socioeconomic position and diabetes

Lower SEP is correlated to advanced biological aging, greater risk for poorer health and early death (23, 33, 35). Individuals with lower SEP are more vulnerable to developing NCDs including diabetes (23). Nevertheless, lower SEP is also found to have significant impact of mortality and morbidity caused by NCDs (4). Development of T2DM is a multifactorial process that progresses over lifetime (36). SEP is shown to have a predominant role in development for chronic progressive diseases, such as T2DM (9). The association between SEP and diabetes has been explored in different countries and between different economic situations between countries. A systematic review (SR) and meta-analysis found evidence for an inverse association between SEP and diabetes (37). Prevalence of diabetes is found to have increased among population with lower SEP in high-income countries, and this association within high-income countries was consistently measured by education and occupation (37).

Several other studies not included in the abovementioned SR, that investigated this association between education and diabetes agree with the conclusions in the SR. Hence, there exists evidence of an inverse association between SEP and diabetes, and there exist a social gradient within diabetes prevalence (24, 38, 39). This supports findings from research where lower SEP increases the risk of developing diabetes (37). This association is confirmed for diagnosed and undiagnosed diabetes within the Norwegian population (8). Another SR from China, Hong Kong and Taiwan, found an association between education and diabetes, with an unclear association for other SEP proxies (40). Research found this association to be more pronounced among female participants, compared male participants (41).

Investigating the association between SEP and diabetes is complex, since potential biological mechanisms and lifestyle behaviors are likely to exist on the causal path (4, 42). Several studies have found potential mediating factors that influences the association between educational level and risk of developing diabetes. These factors include BMI(24), depression (43), smoking status (42), physical activity (42), alcohol consumption (44) and psychosocial working condition (39). The relationship between education and diabetes is partly mediated by lifestyle choices. Research from Denmark recently suggested that there are important social inequalities in lifestyle and motivation among participants with diabetes (45). Persons in the lower SEP groups are associated with a physically inactive lifestyle, and the lowest motivation to be more active (45).

Differences in access to healthcare and health information are found to be attributed to social inequalities in health (46). Findings from a SR reveals that there exists socioeconomic inequalities within diabetes health care (46). Results revealed that individuals with lower SEP are more at risk for worse process indicators and worse intermediate outcomes (22). Other research suggests that there exists an educational gradient within digital coping of equipment for search engines and apps for patients with diabetes, which may affect health outcomes and further increase social inequalities in health (47).

1.5 Objectives for this study

The differences in health outcomes in Norway are comparable to those in the Nordic countries and the continental European countries but higher than those in southern Europe. Educational inequalities among woman highlights this pattern significantly (35). The increase in social inequalities among woman is significant, hence it is of public interest to investigate the association between diabetes and education in a representative population of Norwegian women. More research is needed regarding the drivers of increased risk in individuals with lower SEP, especially due to globalization and the normalization of sedentary lifestyles, increased obesity, and increased comorbidities (9).

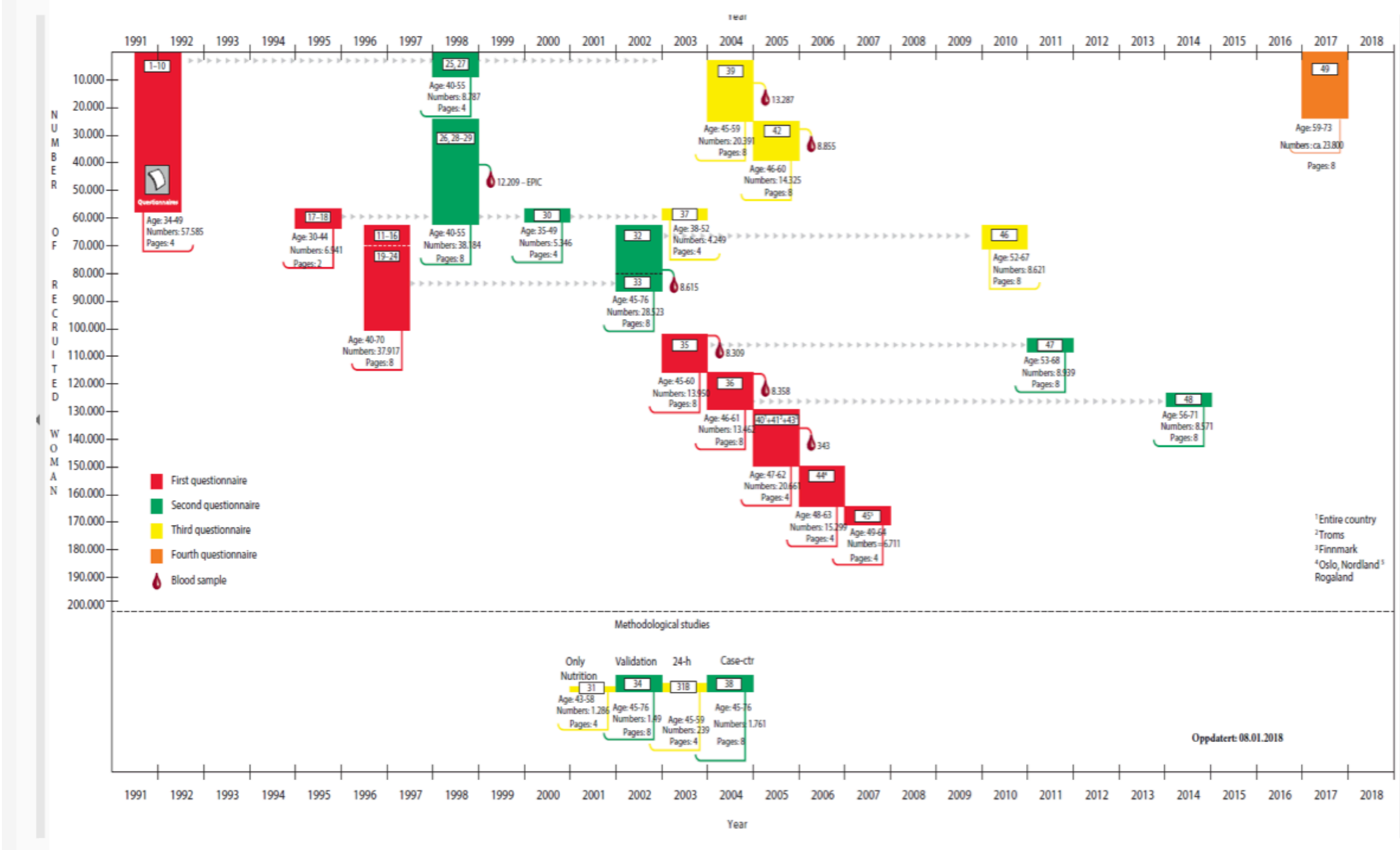
The present study will examine the educational gradient in self-reported diabetes in a large national study of Norwegian women. Findings will provide more knowledge on a public health topic of interest. Hence, this study will contribute important knowledge regarding diabetes mellitus and education level among Norwegian women. These findings will add to the existing knowledge of the relationship between diabetes and education level.

The objective for this study is to investigate the association between diabetes and education. Thus, the research question is: What is the association between self-reported diabetes and educational level in Norwegian women? The hypothesis of this study is that there exists an educational gradient within diabetes in Norwegian women.

2 Materials and methods

2.1 The Norwegian Woman and Cancer study (NOWAC)

The Norwegian Women and Cancer (NOWAC) study was a nationwide cohort established in 1991, with the initial purpose of examining the association between breast cancer risk and oral contraceptive use among Norwegian women. Detailed information of the design and characteristics of NOWAC cohort is described elsewhere (48). Briefly, this national representative cohort study, included approximately 172 000 Norwegian women aged 30-70 years at enrollment. In segments of the first wave (1991-1992) participants did not answer questions regarding dietary habits and, therefore, this information is not included in the present study. Participants answered questionnaires concerning their health situation and lifestyle habits. These questionnaires were distributed in waves. The first wave was performed from 1991 to 1997, and an expansion was performed between 2003 and 2006. Additionally, follow-up questionnaires were performed from 1998 to 2002 for the first wave and in 2011 for the second wave (48). Figure 1 visually describes enrollment and collection of study participants in NOWAC cohort. The study includes data regarding self-reported diseases, various lifestyle factors, such as diet, physical activity, and smoking status, social background, and reproductive health.

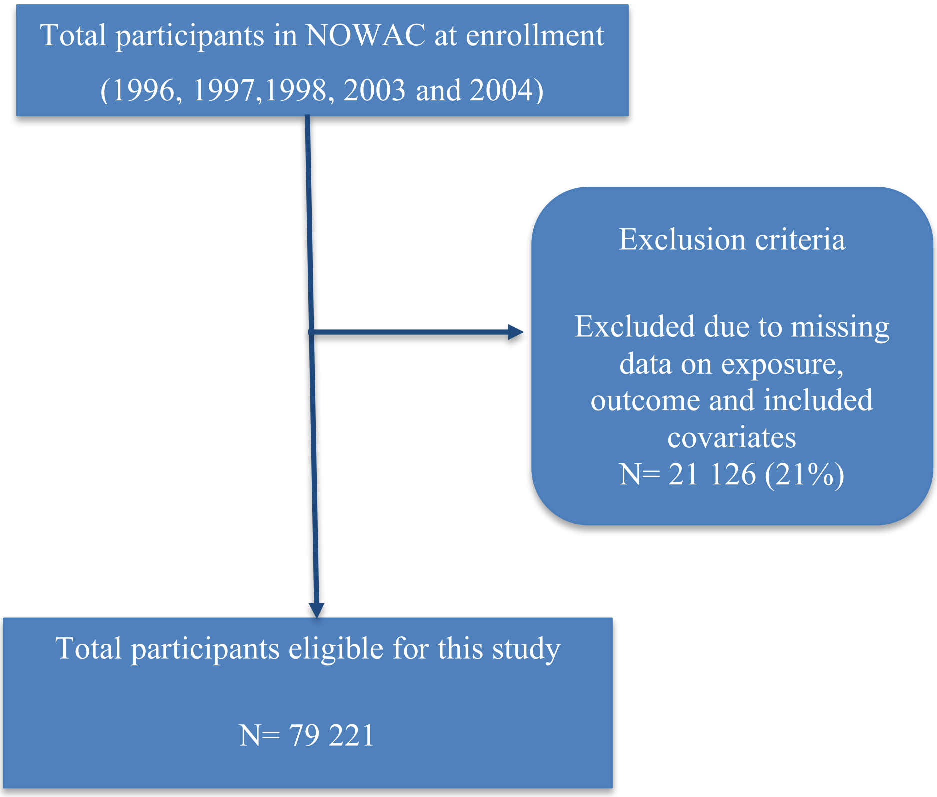


Figur 1 NOWAC enrollment

2.2 Study population and exclusion criteria

Participants included in this study participated in the first and second waves conducted in 1996, 1997, 1998, 2003 and 2004 and answered questions on dietary information (N=100 347). Participants were between 30-70 years old at enrollment and had answered the baseline questionnaires. These questionnaires were sent out to a random selection of potential participants at the point of data collection. The response-rate at enrollment for NOWAC was 52.7% (48). The flow chart describes the selection of study participants (Figure 2). A complete case analysis was performed for all analytical statistics. Thus, participants with missing data regarding self-reported diabetes or educational level were excluded from this study. Participants were also excluded for missing data regarding other covariates, such as age, BMI, physical activity, alcohol consumption, self-reported CVD, and diet variables (total intake of sugar, fruit, fiber and vegetables). This resulted in a total of 21.1% (n= 21 126) of participants being excluded due to missing data on physical activity (8.9%), alcohol consumption (6.5%), education (4.9%), and BMI (3.8%) (supplementary table in Appendix 1). There were a few outliers in the study population that were included in this study because they were assumed not to change the estimates.

Figure 2: Flow diagram: for selection of participants



2.3 Data description

Participants answered questionnaires regarding health status and lifestyle habits. An example of a NOWAC questionnaire can be found in the appendix (appendix 2). The variables of interest in this study are described here.

2.3.1 Dependent variable

The dependent variable was diabetes. Diabetes was self-reported and dichotomized. Participants answered the binary question: "Do you have or have you had the following illness: Diabetes?". The age of diagnosis was reported, where participants answered the following question "If Yes, age when first discovered." Participants with missing answers regarding their diabetes status were recoded as not having diabetes after validation studies of diabetes incidence in NOWAC were performed (49).

2.3.2 Independent variable

Educational level was used as a socioeconomic gradient. Participants answered questions regarding their highest level of completed education in years. Educational level was grouped into three levels based on the Norwegian educational program(50). Primary education is considered nine or fewer years of education, secondary is between 10–12 years of education, and tertiary is ≥13 years of education. Participants with a primary education level were used as a reference in the logistic regression.

2.3.3 Covariates

Age was collected at the time of inclusion in the NOWAC cohort. Age was used as a continuous variable and age-adjustments were performed.

BMI was calculated using self-reported heights ("How tall are you in centimeters?") and weights ("What is your current weight in kilogram?") using the following formula: BMI = weight (kg)/height(meters)^2. BMI was divided into underweight, normal weight, overweight, and obese based on the following WHO classifications: underweight is >18.5 kg/m^2, normal weight is 19–24.9 kg/m^2, overweight is 25–29.9 kg/m^2, and obese is >30 kg/m^2. Relatively few individuals were categorized as underweight (N = 2,134 or 2.2%) without diabetes and (N = 14 or 1%) with diabetes, underweight category was therefore merged with the normal weight category. Final grouping of BMI resulted in the following categories: normal weight >25 kg/m^2, overweight 25–29.9 kg/m^2, and obese >30 kg/m^2. The normal weight category was used as a reference category in the logistic regression analyses.

Physical activity was self-reported in NOWAC on a scale ranging from 1 as the lowest to 10 as the highest. Participants chose their physical activity level based on the following prompt: "Please indicate the level of your physical activity on a scale from very low to very high today. The scale goes from 1–10. By physical activity, we mean both work in and outside the home, as well as training/exercise and other physical activity, such as walking, etc." Physical activity was grouped into "low", "moderate," and "high", where "low" accounted for values between 1–3, "moderate" between 4–7, and "high" between 8–10. The low activity category was used as a reference in the logistic regression analyses.

Smoking status was collected at baseline using several questions that asked about smoking habits from adolescence to mid-life and the frequency of use. This information was combined into one variable for smoking status that was divided into three groups: "never," "former," and "current" smokers. "Never" was used as a reference in the logistic regression analyses.

Alcohol consumption was measured using the following questions: "Are you a teetotaler? Yes/No. If No, how often and how much have you drunk on average in the last twelve months? (beer, wine, sprits and liqueurs).” Alcohol intake was calculated as grams of alcohol per day. The median alcohol consumption in this sample was 2.7 grams/day. The continuous variable was translated into a categorical variable as “never” as non-consumer, 0.1-2.7 grams/day as “>median” and >2.7 grams/day as “<median”.

Self-reported CVD was merged into one category by combining the following five variables: high blood pressure, heart attack, heart failure, heart cramps and stroke. Participants answered five binary questions “Do you have, or have you had any of the following illnesses: high blood pressure, heart failure, heart cramps, heart attack, stroke?”. The answer “yes” on any of the five conditions, resulted in positive CVD case, while “no” on every of the five conditions resulted in a negative CVD case. Missing values were recoded as not having the illness based on the assumption that participants not answering did not have the condition.

Dietary variables were collected from the questionnaires. Participants answered several questions regarding their dietary habits and last consumption. These questions were calculated into five continuous variables of interest in this study: total intake of kilojoules, sugar, vegetables, fruit, and fiber in grams per day. After analyzing for a correlation between dietary variables (total intake of kilojoules, sugar, fruit, vegetables and fiber) using Pearson’s correlation, a strong correlation was found between total fiber and total kilojoules (0.749). Total fiber was, therefore, used in the analysis, and total kilojoules were excluded from the analysis.

2.4 Statistical analysis

All the statistical analyses were performed using the software IBM SPSS Statistics software version 26 for Mac OS (51). Characteristics from the study population were investigated with descriptive statistics, and are provided in tables. Mean age, BMI, smoking status, educational level, physical activity level and dietary habits are presented with a standard deviation (SD) and proportion (%) of self-reported diabetes and not reported diabetes. The dataset was then cross tabulated and chi-square tests were performed to investigate associations between variables. Analytic statistics were then performed. Covariates included in the analysis were chosen based on the literature and the selection was based on data-driven statistical approaches for confounder selection where a change-in-estimates of more than a 10% change were assessed. Descriptive statistics based on self-reported diabetes also gave inductions for a linear association for covariates selection for the multivariable model.

To investigate characteristics of the participants that were excluded (21.1%, N=21 126), a binary variable was constructed. A descriptive table was created for included and excluded cases, including the following variables: diabetes, education, BMI, physical activity, smoking status, CVD cases, alcohol consumption, and dietary variables. The table was visually investigated, and statistically analyzed. Independent *t*-tests were performed to investigate continuous variables between the group's mean values, and are presented with *p*-values. Categorical variables were cross tabulated, and chi-square tests were performed, and presented with *p*-values.

2.4.1 Binary Logistic regression

Binary logistic regression was performed to calculate the odds ratio (OR) and 95% confidence intervals (CI) with *p*-values for the association between self-reported diabetes and educational level. To answer the research question, self-reported diabetes was used as the dependent variable, and educational level was the main independent variable. Three models were investigated: a univariable regression analysis (Model A) was conducted for diabetes and education, an age-adjusted analysis was performed (Model B), and a multivariable model (Model C) was performed that explored possible confounding covariates, such as BMI, physical activity, smoking status, alcohol intake, self-reported CVD, and dietary variables (total intake of fiber, vegetables, fruit and sugar).

Analysis of variance was performed to investigate, and evaluate if including covariates improved the model. The OR was presented as a result of the statistical logistic regression analysis within the corresponding 95% CI. Results were presented with STROBE structure for cross-sectional studies, and *p*-values below 0.05 were considered statistically significant. Multicollinearity was also tested.

2.5 Ethical considerations

Participants have given a written informed consent, when participating in the NOWAC study. The NOWAC study was approved by the Regional Committee for Medical Research Ethics (REC) and the Norwegian Data Inspectorate (NSD). Additional ethical approval from REC and NSD was not required to conduct this study.

3 Results

3.1 Study sample by diabetes

A total of 79,221 participants were eligible for this study. Of these, 1,082 (1.4%) reported having diabetes and 78,136 (98.6%) reported not having diabetes (Table 1). Those with diabetes had a higher age, compared to those who did not have diabetes (52.6 years and 47.6 years). There was a higher proportion of primary education among those who reported having diabetes compared to those who reported not having diabetes (39.2% and 24.2 %, respectively). There was also a lower proportion of tertiary education among those who reported having diabetes compared to those who reported not having diabetes (29.3% and 41.1%, respectively). Those who reported having diabetes had higher BMIs (overweight: 32.1% and obese: 28.5%) compared to those who reported not having diabetes (overweight: 23.1% and obese: 5.9%). Those with diabetes also had a higher proportion of low physical activity than those who reported not having diabetes (20.2% and 12.6%, respectively). The opposite trend was seen in relation to alcohol consumption; those who reported having diabetes had a higher proportion of never or >median alcohol consumption (36.6% and 41.5%, respectively) compared to those who reported not having diabetes (23.1% and 39%, respectively). Participants who reported having diabetes also reported lower intake of sugar, fiber, vegetables, and fruit compared to those who reported not having diabetes. Those with diabetes also had a higher proportion of CVD cases compared to those who reported not having diabetes. Additionally, those with diabetes had a slightly lower proportion of current smokers compared to those who reported not having diabetes (28.3% and 30.7%, respectively).

Tabell 1 Baseline characteristics by diabetes status in the Norwegian Women and Cancer study (N= 79 221)

	Diabetes NO		Diabetes YES	
	Number	Mean (SD/%)	Number	Mean (SD/%)
Age	78139	47.6 (8.1)	1082	52.6 (8.2)
Education[1]	78139	2.2 (0.7)	1082	1.9 (0.8)
- Primary	16992	21.7	350	32.3
- Secondary	27303	34.9	388	35.9
- Tertiary	33844	43.3	344	31.8
Smoking status	78139	1.9 (0.8)	1082	1.9 (0.8)
- Never	27921	35.7	381	35.2
- Former	26274	33.6	381	35.2
- Current	23944	30.6	320	29.6
Physical activity[2]	78139	1.9 (0.6)	1082	1.9 (0.5)
- Low	9758	12.5	214	19.8
- Moderate	55019	70.4	730	67.5
- High	13362	17.1	138	12.8
Alcohol consume[3]	78139	1.9 (0.7)	1082	1.7 (0.7)
- Never	17969	23.1	394	36.6
- >median	30368	39	447	41.5
- <median	29445	37.9	235	21.8
BMI (kg/m^2)[4]	78139	1.3 (0.6)	1082	1.9 (0.8)
- Normal	53414	68.4	398	36.8
- Overweight	192341	24.8	341	31.5
- Obese	5384	6.9	343	31.7
CVD Cases [5]	78139	0.1 (0.3)	1082	0.4 (0.5)
- NO	68636	87.8	569	52.6
- YES	9503	12.2	513	47.4
Total fiber*	78139	21.5 (7)	1082	16.3 (7.7)
Total sugar*	78139	24.2 (16)	1082	16.3 (16.7)
Total vegetables*	78139	164.7 (108.7)	1082	169.4 (121)
Total fruit*	78139	202.6 (149.9)	1082	202.2 (155.6)

[1]Educational level: Primary ≤9 years, secondary 10-13 years and tertiary ≥13 years.
[2]Physical activity on a 10-scale: low 1-3, moderate 4-7 and high 8-10.
[3]Alcohol consumption: non-consumer, <median: 0.1-2.7 grams/day and >median: >2.7 grams/day.
[4]BMI kg/m^2: Normal >25 kg/m^2, overweight 25-29.9 kg/m^2 and obese <30 kg/m^2.
[5]CVD cases consist of self-reported illness: high blood pressure, heart attack, heart failure, heart cramps and stroke.
*Measured as grams/day

3.1.1 Study sample by education

Supplementary table (appendix 3) describes characteristics of the study sample by education level (primary, secondary and tertiary). The descriptive statistic illustrates a higher mean age among lower levels of education (50.8 years, 47 years and 46.9 years respectively). The descriptive statistics reveals higher proportion of reported diabetes among those with lower education (primary: 2.1%, secondary: 1.4% and tertiary: 1%) along with higher proportion of self-reported CVD among lower education (primary: 18.1%, secondary: 13.4% and tertiary: 10.2%). Those with lower levels of education had a higher proportion of overweight and obesity (primary: overweight 30.3% and obesity 9.2%, secondary: overweight 22.5% and obesity 5.8%: and tertiary: overweight 17.7% and obesity 3.9%). Among those with lower education a higher proportion of smokers were found (primary: 38.2%, secondary: 34.8% and tertiary: 22.6%). There is also seen that those with higher education eat more fruit, vegetables and fiber, while those with lower education eat more sugar. Additionally, those with higher education had a higher alcohol consumption (> median: primary: 24.9%, secondary: 36.6% and tertiary: 44.9%). Lastly, those with lower levels of education have a higher proportion of low physical activity primary: 15.6%, secondary: 12.5% and tertiary: 11.3%.

3.2 Association between education and diabetes

Binary logistic regression is presented in table 2. A univariate logistic regression model (Model A) resulted in a 31% decrease in odds for a woman with secondary education to self-report having diabetes compared to a woman with primary education (OR = 0.69, 95% CI: 0.5–0.8). There was a significant 51% decrease in odds for a woman with tertiary education to self-report having diabetes compared to a woman with primary education (OR = 0.49, 95% CI: 0.43–0.57).

The age-adjusted model resulted in a 14% decrease in odds for a woman with secondary education compared to a woman with primary education to report having diabetes (OR=0.86, 95% CI: 0.74-1.0), this result was not statistically significant. There was a significant 35% decrease in odds for a woman with tertiary education, compared to a woman with primary education for self-reporting diabetes (OR=0.65 95% CI: 0.56-0.76).

The multivariable model (Model C) adjusted for the following covariates: age, BMI, physical activity, alcohol consumption, smoking status, CVD disease and diet (total grams of fruit, total gams of vegetables, total grams of sugar, and total grams of fiber, per day). This model resulted in 2% increased odds for secondary education compared to primary education (OR = 1.02, 95% CI: 0.87–1.19). The association for tertiary education resulted in a 9 % decrease in odds for self-reporting diabetes among woman with tertiary education compared to primary education (OR = 0.91, 95% CI: 0.78–1.08). None of the results from the multivariate models were statistically significant.

Table 2: Binary logistic regression summary of the association between educational level and diabetes in the Norwegian Women and Cancer study (n=79 221)

	N	OR (95% CI)	*p*-value
Model A[1]	79221		
Primary education		1.00	<0.001
Secondary education		0.69 (0.6-0.8)	<0.001
Tertiary education		0.49 (0.43-0.57)	<0.001
Model B[2]	79221		
Primary education		1.00	<0.001
Secondary education		0.86 (0.74-1.0)	0.053
Tertiary education		0.65 (0.56-0.76)	<0.001
Model C[3]	79221		
Primary education		1.00	0.344
Secondary education		1.02 (0.87-1.19)	0.829
Tertiary education		0.91 (0.78-1.08)	0.279

[1]Univariable model
[2] Age-adjusted model
[3]Multivariable model, adjusted for: age, BMI (normal >25 kg/m^2, overweight 25-29.9 kg/m^2, obese >30 kg/m^2), physical activity level (low: 1-3, moderate 4-7, high 8-10), alcohol consumption (never, >median (0.1-2.7 grams/day) and <median (>2.7 grams/day)), smoking status (never, former, current), CVD cases (consist of self-reported illness: high blood pressure, heart attack, heart failure, heart cramps and stroke) and diet (total intake of fruit (grams/day), vegetables (grams/day), sugar(grams/day) and fiber (grams/day).

3.3 Characteristics of those included and those not included

Participants included in the study were compared with those excluded (Table 4); included participants had a lower mean age, 47.7 and 51 ($p < 0.001$), a lower share of participants reporting diabetes, 1.4% and 1.7% ($p < 0.001$), and a higher level of education (primary 21.9%, secondary 35%, tertiary 43.2% and primary 36.8%, secondary 33.3%, tertiary 29.8%, respectively). Further, those included had lower BMIs (normal 67.9%, overweight 24.8%, and obese 7.2%) compared to those excluded (normal 62%, overweight 28.8%, and obese 9.2%; $p < 0.001$), higher levels of physical activity (low: 12.6%, medium: 70.4%, high: 17% and low: 13.5%, medium: 69.9%, and high: 16.6%; $p < 0.351$), a higher proportion of current smokers ($p < 0.001$), a lower proportion of CVD cases ($p < 0.001$), a lower level of alcohol consumption ($p < 0.001$), higher grams of sugar per day ($p < 0.002$), lower grams of fruit ($p < 0.001$), lower grams of vegetables per day ($p < 0.724$), and higher intake of fiber in grams per day ($p <0.001$).

Table 2: Characteristics included and excluded

	Attendance The Norwegian Woman and Cancer study at baseline N=100347							
	Included N= 79 221			**Excluded N= 21126**			**Independent T**	**Chi-square**
	Valid	**Missing**	**Mean (SD/%)**	**Valid**	**Missing**	**Mean (SD/%)**	***p*-value**	***p*-value**
Age (years)	79221	0	47.7 (8.1)	21126		51 (8.7)	<.001	
Diabetes	79221	0	0.01 (0.1)	21126	0	0.02 (0.1)		<.001
- Yes	1082		1.4	365		1.7		
- No	78139		98.6	20761		98.3		
Diabetes age	930		41.8 (14.7)	255	114549	46.2 (14.2)	<.001	
Education[1]	79221	0	2.21 (0.8)	16121 (88.7)	5005 (11.3)	1.93 (0.8)		.100
- Primary	17342		21.9	5937		36.8		
- Secondary	27691		35	5373		33.3		
- Tertiary	34188		43.2	4811		29.8		
BMI (kg/m^2)[2]	79221	0	1.39 (0.6)	17287	3839	1.47 (0.7)		<.001
- Normal	53812		67.9	10726		62		
- Overweight	19682		24.8	4975		28.8		
- Obese	5727		7.2	1586		9.2		
Physical activity[3]	79221	0	2.04 (0.5)	12152	8974	2.03 (0.5)		.351
- Low	9972		12.6	1641		13.5		
- Moderate	55749		70.4	8494		69.9		
- High	13500		17	2017		16.6		
Smoking status	79221	0	1.96 (0.8)	19586 (92.7)	1540 (7.3)	1.96 (0.8)		<.001
- Never	28302		35.7	7471		38.1		
- Former	26655		33.6	6088		31.1		
- Current	24264		30.6	6027		30.8		
CVD[4]	79221	0	0.12 (0.3)	21126	0	0.16 (0.4)		<.001
- Yes	10016		12.6	3496		16.5		
- No	69205		87.4	17630		83.5		
Alcohol consume[5]	79221	0	1.99 (0.7)	14635 (85.4)	6491 (14.6)	1.82 (0.7)		<.001
- Never	18363		23.2	4819		32.9		

- <median	43075		54.4	7559		51.7	
- >median	17783		22.4	2257		15.4	
Total sugar[*]	79221	0	24.1 (15.9)	21126	136	22.93 (16.4)	.002
Total fiber [*]	79221	0	21.4 (7.1)	21126	82	20.14 (7.6)	<.001
Total fruit[*]	79221	0	202.63 (149.9)	21126	1254	199.7 (157.1)	<.001
Total vegetables [*]	79221	0	164.75 (108.9)	21126	461	147.1 (109.2)	.724

[1]Education level: Primary ≤9 years, secondary 10-12 years and tertiary ≥13 years.
[2]BMI (kg/m)^2: Normal >25 kg/m^2, overweight 25-29.9 kg/m^2 and obese <30 kg/m^2.
[3]Physical activity on a 10-scale: low 1-3, moderate 4-7 and high 8-10.
[4]CVD cases consist of self-reported illness: high blood pressure, heart attack, heart failure, heart cramps and stroke.
[5]Alcohol consumption: non-consumer, <median: 0.1-2.7 grams/day and >median: >2.7 grams/day.
[*]Measured in grams/day

4 Discussion

4.1 Summary of findings

This cross-sectional study investigated the association between self-reported diabetes and educational levels, among Norwegian women. The results showed an inverse association between educational levels and self-reported diabetes. The findings remained when the analyses were adjusted for age with a slightly lower OR, and only result for the highest level of education compared to lowest level of education remained significant. However, when adjusting for potential confounders in the multivariable models there was no statistically significant association between level of education and diabetes. Thus, the hypothesis is partly rejected.

4.2 Comparisons to other literature

Studies investigating the association between educational levels and diabetes have found an inverse association between education and diabetes in crude-, or age-adjusted models (37, 42, 52-58). A systematic review, found consistent results for the strength of the association in high income countries after adjustments (37). Maty et at. (58) found similar results where education predicted diabetes, but the association was no longer statistical significant after adjustment (58). Van Zon et al. (53) similarly found an inverse association, but adjustments attenuated the effect, even though the risk remained increased. This is comparable to Lee et al. (52) and Demakakos et al. (55) where adjustments also attenuated the association. Maty et al. (58) suggested that overweight and obesity appeared to mediate the association in their study. Mediation analysis was not performed in this study.

In contrast to this study, Steele et al. (54) found that educational level was inversely associated with diabetes incidence among older (50-75 years) after adjusting for possible confounders. This association was found to be mediated by BMI, alcohol consumption, hypertension, fasting triglycerides, high-density lipoprotein cholesterol, physical activity, and smoking status, which all together explained 31.7% of the association (54). Comparable to Steele et al. (54), Espelt et al. (56) found significant inequalities in prevalence and incidence of T2DM among women in Europe, and these inequalities were found to be mediated by BMI. Furthermore, Demakakos et al. (55) found that the inverse association was mediated by psychosocial factors, unhealthy behaviors, and obesity.

A systematic review exploring the association between SEP (education, income and occupation) and T2DM found an inverse association independent of the SEP indicator used (37). Demakakos et al. (55) investigated the association between SEP and incidence of diabetes using several SEP measures; childhood SEP, education, occupational class, income, wealth and participant's social status. The study found an inverse association between diabetes and all SEP measures. However, after adjustments for confounder only childhood SEP, education and participant´s social status remained significant (55). Geyer et al. (25) investigated education, income, and occupational class in relation to four types of health outcomes in Sweden and Germany (25). The study found the SEP measures (education, income and occupational class) to represent different underlying phenomena, and causal mechanisms. Further, education was found to be the most significant predictor for diabetes (25). Similar to Geyer et al., Williams et al. (42) investigated the association between SEP (education, income, and occupation) and glucose metabolism and T2DM, with emphasis on mediating role of health behavior in Australian population. They confirmed that education was the strongest predictor, but after adjustments for age, sex, and health behavior only the highest compared to the lowest educational group remained significant. Mediation analysis found that physical activity and smoking explained 27% of the association between SEP and T2DM(42).

4.3 Diabetes prevalence

Lack of a national registry for diabetes combined with the progressive nature of the disease makes it challenging to assess prevalence and incidence of diabetes in Norway. In the present study, the prevalence of diabetes was 1.4%. Stene et. al (59) estimated diabetes prevalence in Norway to be between 260 000-280 000 diagnosed cases and approximately 60 000 undiagnosed, which indicated that 5% of the Norwegian population between 30-89 years had either diagnosed or undiagnosed T2DM in 2020 (59). This estimate is based on discretionary assessments and is therefore uncertain. The prevalence from Stene et al. (2020) included diagnosed and undiagnosed diabetes for both male and female population.

Diabetes is a collective designation for variants of the disease affecting blood glucose levels. The three main types T1DM, T2DM and GDM differ between each other (1). The present study does not distinguish between the different variants of diabetes, as the participants were asked to report diabetes yes/no without specifications. However, the mean age at the diagnosis of diabetes was calculated to be 39.14 years in this study, and T2DM is most commonly diagnosed among elderly. Additionally, 90% of diabetes cases are presumed to be T2DM in Norway. Thus, it is assumed that most of the diabetes cases in the study population are T2DM. A validation of self-reported diabetes in NOWAC has been done previously, where participants and their general practitioners confirmed diabetes diagnosis and which type of diabetes (60). The findings from the validation study revealed a positive predicted value (PPV) of 93% for all types of diabetes, and corresponding PPV of 84% for T2DM. Hence, there is a 7% risk of overestimating diabetes in the NOWAC cohort(60). Additionally, a test-retest reliability validation was performed of self-reported diabetes diagnosis in NOWAC in 2016 (49). This validation study found that self-reported information about diabetes was reliable, with no clear pattern of inconsistency. The kappa agreement, however, was stronger among responders with higher education, compared to responders in lower educational group(49).

The prevalence of diabetes and the proportion of undiagnosed diabetes was investigated in the Tromsø study from 1994-2016 (61). The findings from this study showed an increasing prevalence among women from 2.3% to 3.9% between 1994 to 2016, and a decreasing proportion of undiagnosed diabetes from 33% to 20% (61). Thus, these findings might be a reasonable explanation for a low prevalence of T2DM in this study.

4.4 Educational gradient

It is well established that education is one of the determinants of health (32). This gradient is seen within life expectancy across all age groups, and life style behavior, health, and disease outcomes (32). Education is the basis for, and contributor to, several processes that support health throughout the adult life cycle (62). These processes are complex and versatile and consist of multiple causal relationships. The Norwegian directorate of health's report highlighting that causal pathway between education and health can go through occupation, standard of living, working conditions, lifestyle or behavior habits, problem-solving skills, sense of control, stressful life events among others (62).

Highly educated people generally have better health than lower educated people (62). The WHO suggests that nine in ten T2DM cases are associated with health behaviors, and can be prevented with changes in diet, physical activity, BMI and smoking (1). Unhealthy lifestyle behavior such as obesity, smoking, heavy alcohol consumption, and physical inactivity are more common among those with lower educational level (63). Similar is found in this study, where participants with primary education had increased BMI, lower levels of physical activity, were more likely to be current smokers, had lower intake of fruit, vegetables and fiber, and higher levels of sugar intake compared to those with tertiary education On the other hand, alcohol consumption was higher in tertiary educated participants. The educational gradient is also seen within self-reported diseases (diabetes and CVD) in this study sample, where those with primary education have higher proportion of diabetes and CVD, compared to those with tertiary education.

Unfavorable lifestyle behavior such as increased BMI, sedentary lifestyle, smoking habits, poor diet, binge drinking habits, increased psychosocial stress, and sleeping problems in life increase risk of being diagnosed with diabetes (1, 64, 65). These lifestyle behaviors exacerbate risk factors concerning ethnicity, family history and genetic predisposition to development of diabetes (66). Thus, it is reasonable to assume that education is a proxy for lifestyle behaviors, and that the levels of education operate as a proxy for risk factors that can predicts diabetes.

4.5 Possible explanation of the educational gradient

Dahl et al. (32) highlights skills and competence, better jobs and safer income, and intellectual tools as three possible explanations for the educational gradient seen in health (32). Skills and competence provide people with the ability to adjust their social environments, either through adaptation or change, and they provide the basis for lifelong learning (32). Education can provide time and resources for leisure activities and training through factors such as environmental factors, and psychological factors among others(32). Even though physical activity was found not to explain the increasing prevalence of diabetes among adults in Canada, the urbanization was suggested as one of several possible factors explaining the increase (12). Urbanization is a big concern for public health, where unhealthy living and working conditions, inadequate green space, air- and noise pollution, water and soil contamination, along with lack of space for physical activity are seen in a greater amount among people with lower education (67).

Along with increased educational levels comes the probability of better employment opportunities and financial security, which increase opportunities for better living conditions, nutrition, and recreation (32). These bring more stability into people's life cycle, and contribute to improved health (62). Researchers have found an association between higher occupational social class and increased food expenditure, which resulted in healthier food intake (68). The study was conducted in the United Kingdom and, therefore, the data used may not be comparable with Norwegian prices, food choices, and cultural differences between people in Norway and the United Kingdom. Nonetheless, increased educational level and more secure employment can lead to healthier diet. Poor diet is a risk factor for developing diabetes, and a healthier diet is recommended as a preventive measure (3). Intake of fruit and vegetables are another possible factor explaining the increase in prevalence of diabetes (12). Further, financial security can affect stress levels. The experience of the feeling of control over own life can reduce harmful stress reactions, when exposed to stressful and health-threatening situations or conditions (32). Stress is found to be a risk factor for diabetes, and lower levels of education are related to higher stress levels. Long-term stress affects the endocrine system, which may lead to diabetes (69).

The third explanation acknowledges that education provides intellectual tools that help people to gain control of their own life (32). Education provides better access to health literacy and strengthens the ability to transform this knowledge into healthy behaviors, such as healthy diet, increased physical activity and none-smoking behavior (32, 70). In addition, education can contribute to being more receptive to health information and the use of health services (71). Further, education can give knowledge capital, and strengthens competence and the ability to acquire knowledge (70). Preventive measures can therefore paradoxically reinforce educational differences in health behaviors, because more educated people increasingly absorb information about healthy habits. Finally, education entails access to social relationships and networks that possess health-promoting resources, such as social support, information and various material goods (32).

The educational gradient is complex and multifactorial. Explaining the underlying mechanisms is beyond the scope of this master thesis, due to the research question and the fact that observational designs are not compatible for investigating causal relations.

4.6 Strengths and limitations

The present study uses data from a nationally representative cohort, with a reasonably high response rate (52.7%) similar to other population based studies(48). The study had a large number of participants that were randomly selected from the general population of women in Norway. Additionally, information about diabetes (49, 60), BMI (72), and physical activity (73), has been validated in NOWAC previously. Validation has moreover been done with the entire NOWAC cohort (74). Nonetheless, this study has several limitations.

4.6.1 Internal validity

Internal validity is the magnitude to which we can be assured that “a cause and effect” – relationship recognized in this study, cannot be explained by other factors. Selection bias along with external validity, information bias, and residual confounding the items discussed here.

4.6.1.1 Selection bias and external validity

Participation in population-based studies is voluntary. Those who volunteer in cohorts tend to have better health profile and health outcomes, compared to none-responders(75). The healthy volunteer effect might threaten the external validity (75). However, Lund et al. examined the NOWAC cohort for the disparities between responders and non-responders (74). This validation study found negligible differences between responders and non-responders, except from a somewhat higher education for woman participating in NOWAC cohort compared to the source population (74). Thus, the inverse association between education and diabetes, and the somewhat higher proportion of educated participants might be a part of the explanation of the lower diabetes prevalence in this present study. Lund et al. concluded that the somewhat higher proportion of participants with higher education was not a source for selection bias in NOWAC cohort (74). This study is therefore generalizable for the female Norwegian population.

The present study excluded 21 126 (21.05%) participants due to missing values on included covariates. Excluding missing cases might lead to selection bias. Excluded participants had significantly higher age, higher diabetes prevalence, higher BMI, more never smokers, lower CVD cases, lower alcohol consumption, and lower sugar, fruit, and fiber intake.

4.6.1.2 Education as SEP indicator

Education was used as an SEP indicator in this study. Some advantages of using education as an indicator are that education is stable over time and is well studied in epidemiological studies. A systematic review investigating the relationship between SEP and diabetes found an inverse relationship irrespective of the SEP indicator used (37). However, education can be problematic to interpret in this study since these data were collected from 1996–2003 from women who were 30–70 years old at the point of collection, including participants born between 1926 to 1973. Thus, the study included participants born in very different birth cohorts. Research has already established that 85% of women born before 1938 completed up to 10–11 years of education (76). Comparing levels of education based on years of education might introduce bias in the estimates because education levels are not weighted similarly for the older birth cohorts as for the younger birth cohorts.

Another limitation when using education as a measure of SEP for women is that education level may not always reflect their SEP, as their SEP can be impacted by the SEP of their partner (23). Nevertheless, income and occupation are also well-researched gradients. Baseline information in NOWAC provides information on education and income. A combination of education and income might have been better SEP indicator, but as the research question specified education, income was therefore not used. Even though findings might have differed slightly if income was also used, we assume that the main conclusion would be rather similar, since the most used SEP measures (income, education, and occupation) correlate (25). In addition, other studies using education, occupation, and income as measures of SEP found education to be the most significant predictor for diabetes (25, 58).

4.6.1.3 Information bias

This study uses several self-reported data: diabetes, education, BMI (height and weight), physical activity, CVD, smoking habits, alcohol consumption and dietary variables. With self-reported data comes the risk of misclassification, through recall bias or social desirability bias (77, 78). Thus, participants could potentially report incorrect or imprecise data. Social desirability bias occurs when participants over- or underreport their harmful behavior (e.g., underreporting smoking habits as a result of social stigma; (78). This can occur when participants answer sensitive question, and thereof chooses the most socially accepted answer.

Self-reported variables can lead to misclassification, which can over- or underestimate studied association. Research indicates that participants tend to over-report their height and physical activity while they underreport their weight, food intake, and alcohol consumption (79). Validation study of self-reported BMI in NOWAC found a slight tendency of participants to underreport weight, but there was a substantial agreement between self-reported and measured BMI (72). Additionally, self-reported physical activity has been validated in NOWAC. This validation study suggested that self-reported physical activity in a 10-category scale was reliable in ranking the physical activity levels among participants, but couldn´t quantify frequency, duration, intensity or type of activities (73). Dietary variables have similarly been validated in NOWAC. Dietary variables were collected from the food frequency questionnaire (FFQ), which consisted of 90 questions regarding food intake the preceding year. The NOWAC FFQ was validated through a 24-hour dietary recall (80) and a test-retest reproducibility (81), that reported good recall for food eaten frequently and for macro-instructions. Intake of energy, fat, added sugar, and alcohol were lower in the FFQ than in the 24-hour dietary recall (80).

4.6.1.4 Residual confounding

While research has found no change in the prevalence of diabetes among Sami and non-Sami inhabitants (82), it is also reported that people with Asian and South African backgrounds are at a higher risk of developing diabetes (1, 3). Similar to ethnicity, family history is a well-recognized risk factor for T2DM (1, 3, 7). Research provided evidence that low education and a family history of T2DM is associated with the development of T2DM (53). Additional, low education and family history exacerbate each other's impact on T2DM development for women (53). Information about family history of diabetes was not collected in the NOWAC cohort, which made adjustment for this factor not possible. Further, research found evidence for psychosocial stress as a potential risk factor for diabetes in Swedish women (39). Perceived stress is found to be a significant risk factor for T2DM among women, and this risk is found not to be mediated by other known risk factors (55, 83). However, information regarding psychosocial stress was not collected.

4.6.2 Methodological considerations

The limitation of a cross-sectional design is reverse causation, which hinders causal interpretation. This study investigated an association between educational levels and prevalence of diabetes and found statistically significant inverse association in the crude model. However, based on simultaneous data-collection, we cannot be certain that the associations direction goes from education to diabetes, i.e. that educational level predicts diabetes. With reverse causation arises the question whether diabetes diagnosis may cause changes in educational level. However, reverse causation was expected not to be present here, because the participants were 30–70 years old, and we can assume that most of these women already had completed the principal of their education before being included in NOWAC and before being diagnosed with diabetes. The mean age of self-reported diabetes diagnosis was 39.2 years, which supports this assumption.

4.6.2.1 Confounder selection

Investigating the association between education level and health is complex and multifactorial. Confounding is a concern in observational studies. A confounder is a factor that is associated with both exposure and outcomes, and is not on the causal pathway between the exposure and the outcome. If not accounted for in a regression model, it can bias the association between said exposure and outcome (84). Adjusting for a potential confounding factor will make it possible to separate the impact of the exposure of interest on an outcome. In this study education and diabetes can be affected by a third variable, that contributes to explain parts of the association. When investigating the relationship between the exposure and the outcome of interest, there is a possibility that the effect of the exposure (i.e, educational level) acts through other factors (i.e., physical activity or smoking) that mediates the association with an outcome (i.e, diabetes). Mediator is a factor that is on the causal pathway between exposure and an outcome, and is therefore explaining, either partly or fully, the observed association between the exposure and the outcome (84). Mediating factors should not be adjusted for in a regression model. In the association between educational level and diabetes, factors like BMI, physical activity, smoking status, alcohol consumption, and dietary may act like mediation factors.

Previous studies have found that obesity is a strong risk factor for diabetes and other metabolic diseases (85-87). It addition, there is an educational gradient within different categories of BMI, where increased education is related to lower BMI (19). In this study the results revealed this gradient; the proportion of obese participants are higher among primary education compared to tertiary (9.2% and 3.9%). Thus, BMI is related to the exposure and outcome, and could operate as a confounder. However, several studies have observed that overweight and obesity mediate the association between education and diabetes (24, 52, 54, 55, 58). This means that the effect of education on diabetes partly acts through BMI and therefore operates as a mediator, rather than as a confounder in the association. In cross-sectional design the exposure and the outcome are measured at the same time and therefore it is possible that BMI is on the causal pathway between education and diabetes in this study. Similar to BMI, other lifestyle factors adjusted for CVD, alcohol consumption, physical activity, smoking status and dietary variables, can operate as mediators, rather than confounders in the study.

The major challenge with this study was to differentiate between confounder and mediator. Thus, multivariable analysis was performed including possible factors that could act as both confounders and mediators. The univariate model found a significant inverse association between education and diabetes, similar to the age-adjusted model, while the multivariable model found no significant association. If the previously mentioned covariates operated as confounders in this present study, the fully adjusted model would present the true estimate between education and diabetes. Based on previous studies, mentioned covariates can mediate the effect. Mediation analysis is therefore required, but it is outside the scope of this master thesis program. Additionally, mediation analysis would be inappropriate without a prospective design, and are therefore not performed here. Nevertheless, it is reasonable to assume that with a prospective design, lifestyle behavior would mediate at least to some extent the association between education and diabetes.

4.7 Clinical relevance and further research

The conclusion from this study is that education is not a predictor for diabetes in the study sample, after adjusting for lifestyle behavior. Lifestyle behavior interfered strongly with the estimate, and education is expected to be a proxy for these behaviors. The findings from this study suggest that preventive measures should focus on lifestyle behaviors identified as risk factors for diabetes, towards all educational levels, and with an additional focus on the lower educational level, due to the fact that higher educated persons have better knowledge capital to convert preventive measures into actions. Preventive measures for diabetes should aim to adhere to a better lifestyle behavior- with increasing physical activity, normalizing BMI, decreasing or even stopping smoking, healthier diet, moderate alcohol consumption, and decreased stress levels. The modifiable risk factors are previously found to exacerbating the non-modifiable. Preventive measures with broad covering can therefore affect participants with modifiable risk factors for diabetes. Additionally, preventive measures towards lifestyle behavior would in fact have a preventive effect on other health threatening diseases or conditions as well, especially other NCDs.

The study design used in this thesis is not appropriate to draw causal relationships between education and diabetes, and further research with a prospective design is required.

5 Conclusion

This study found an educational gradient in self-reported diabetes within the female population in Norway, but the association was strongly confounded by different factors, suggesting that educational level does not predict diabetes. Education is assumed to be a proxy for lifestyle factors, which predicts diabetes. The inconsistency regarding diabetes prevalence induces the urgent need for a diabetes register in Norway.

References

1. WHO. Global report on Diabetes. WHO.int; 2016. Contract No.: ISBN 978 92 4 156525 7.
2. WHO. Ten threats to global health in 2019 who.int2019 [cited 2021 13.december]. Available from: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019.
3. International Diabetes Federation. IDF Diabetes Atlas, 10th edition.; 2021.
4. Lago - Peñas S, Rivera B, Cantarero D, Casal B, Pascual M, Blázquez-Fernández C, et al. The impact of socioeconomic position on non-communicable diseases: what do we know about it? Perspect Public Health. 2021;141(3):158-76.
5. Verma AK, Goyal Y, Bhatt D, Dev K, Alsahli MA, Rahmani AH, et al. A Compendium of Perspectives on Diabetes: A Challenge for Sustainable Health in the Modern Era. Diabetes, metabolic syndrome and obesity. 2021;14:2775-87.
6. Buse JB, Caprio S, Cefalu WT, Ceriello A, Del Prato S, Inzucchi SE, et al. How Do We Define Cure of Diabetes? Diabetes Care. 2009;32(11):2133-5.
7. Joseph J, Svartberg J, Njølstad I, Schirmer H. Incidence of and risk factors for type-2 diabetes in a general population: The Tromsø Study. Scand J Public Health. 2010;38(7):768-76.
8. Ruiz PL-D, Hopstock LA, Eggen AE, Njølstad I, Grimnes G, Stene LC, et al. Undiagnosed diabetes based on HbA1c by socioeconomic status and healthcare consumption in the Tromsø Study 1994–2016. BMJ open diabetes research & care. 2021;9(2):e002423.
9. Galicia-Garcia U, Benito-Vicente A, Jebari S, Larrea-Sebal A, Siddiqi H, Uribe KB, et al. Pathophysiology of Type 2 Diabetes Mellitus. Int J Mol Sci. 2020;21(17):6275.
10. Zhou B, Stevens GA, Riley LM, Fouad HM, Pilav A, Agyemang C, et al. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. The Lancet (British edition). 2016;387(10027):1513-30.
11. Roth GA, Geleijnse JM. 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. The Lancet (British edition). 2018;392(10159):1736-88.
12. Thibault V, Belanger M, LeBlanc E, Babin L, Halpine S, Greene B, et al. Factors that could explain the increasing prevalence of type 2 diabetes among adults in a Canadian province: a critical review and analysis. Diabetol Metab Syndr. 2016;8(1):71-.
13. Norwegian Institute of Public Health. Public Health Report: Health status in Norway 2018 [Report]. 2018 [cited 2022 10. January]. Available from: https://www.fhi.no/contentassets/d021a759c5ed48ae85fffc94e35785cf/health_status_in_norway_2018.pdf.
14. Stene LC, Ruiz PL-D, Gulseth HL, Berg C. Diabetes i Norge (Diabetes in Norway) [Public health-report]. fhi.no2014 [updated 31.05.2021; cited 2021 9.December]. Available from: https://www.fhi.no/nettpub/hin/ikke-smittsomme/diabetes/.
15. Smelser NJ, Baltes PB. International Encyclopedia of Social and Behavioral Sciences. Saint Louis: Saint Louis: Elsevier Science & Technology; 2001.
16. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: Concepts, methodologies, and guidelines. Annu Rev Public Health. 1997;18(1):341-78.
17. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). J Epidemiol Community Health. 2006;60(1):7-12.
18. Morroni C, Susser E, Link B, Myer L. Social Epidemiology. 2008. p. 74-87.

19. The Norwegian Directorate of Health. Gradientutfordring. Social- og helsedirektoratets handlingsplan mot sosiale ulikheter i helse. (The gradient challange. The Norwegian Directorate for Health and Social Affairs´action plan against social inequalities in health.). Report. 2005. Contract No.: IS-1229.
20. Lynch J, Smith GD, Harper SAM, Hillemeier M, Ross N, Kaplan GA, et al. Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review. Milbank Q. 2004;82(1):5-99.
21. Irwin A, Vega J, Solar O. Social Determinants of Health, the United Nations Commission of. 2008. p. 64-9.
22. Grintsova O, Maier W, Mielck A. Inequalities in health care among patients with type 2 diabetes by individual socio-economic status (SES) and regional deprivation: a systematic literature review. Int J Equity Health. 2014;13(1):43-.
23. Gallo V, Mackenbach JP, Ezzati M, Menvielle G, Kunst AE, Rohrmann S, et al. Social Inequalities and Mortality in Europe - Results from a Large Multi-National Cohort. PLoS One. 2012;7(7):e39013-e.
24. Smith PMPMPH, Smith BTM, Mustard CAS, Lu HP, Glazier RHMDMPH. Estimating the direct and indirect pathways between education and diabetes incidence among Canadian men and women: a mediation analysis. Ann Epidemiol. 2013;23(3):143-9.
25. Geyer S, Hemström Ö, Peter R, Vågerö D. Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. J Epidemiol Community Health. 2006;60(9):804-10.
26. Krieger N. A glossary for social epidemiology. J Epidemiol Community Health. 2001;55(10):693-700.
27. Conway DI MA, Brown D, et al. . Measuring socioeconomic status and inequalities. In: Vaccarella S L-TJ, Saracci R, et al., editors., editor. Reducing social inequalities in cancer: evidence and priorities for research. Lyon (FR): International Agency for Research on Cancer; 2019.
28. United Nations Development Programme. Human Developemt Report 2020: The frontier - Human Developemt and the Anthropocene. 2020 15. December
29. Sachs J, Kroll, C., Lafortune, G., Fuller, G., Woelm, F. The Decade of Action for the Sustainable Developement Goals: Substainable Development Report 2021. Cambridge: Cambridge University Press; 2021.
30. SSB. Befolkningens utdanningsnivå (The educational level of the population) 2021 [updated 18. june 2021; cited 2022 10. january]. Available from: https://www.ssb.no/utdanning/utdanningsniva/statistikk/befolkningens-utdanningsniva.
31. OECD. OECD Economic Surveys: Norway 2019 2019 [cited 2022 10. January]. Available from: https://www.oecd-ilibrary.org/content/publication/c217a266-en.
32. Dahl E, Bergsli H, van der Wel K. Sosial ulikhet i helse: en norsk kunnskapsoversikt (Social inequalities in health: a Norwegian knowledge overview). Høgskolen i Oslo og Akershus; 2014.
33. de Gelder R, Menvielle G, Costa G, Kovács K, Martikainen P, Strand BH, et al. Long-term trends of inequalities in mortality in 6 European countries. Int J Public Health. 2016;62(1):127-41.
34. United Nations. World social report 2020: inequality in a rapidly changing world [Report]. 2020 [cited 2022 14. january]. Available from: https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2020/01/World-Social-Report-2020-FullReport.pdf.
35. Mackenbach JP, Stirbu I, Roskam A-JR, Schaap MM, Menvielle G, Leinsalu M, et al. Socioeconomic Inequalities in Health in 22 European Countries. N Engl J Med. 2008;358(23):2468-81.

36. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of In Utero and Early-Life Conditions on Adult Health and Disease. N Engl J Med. 2008;359(1):61-73.
37. Agardh E, Allebeck P, Hallqvist J, Moradi T, Sidorchuk A. Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. Int J Epidemiol. 2011;40(3):804-18.
38. Qi Y, Koster A, van Boxtel M, Köhler S, Schram M, Schaper N, et al. Adulthood Socioeconomic Position and Type 2 Diabetes Mellitus-A Comparison of Education, Occupation, Income, and Material Deprivation: The Maastricht Study. Int J Environ Res Public Health. 2019;16(8):1435.
39. Agardh EE, Ahlbom A, Andersson T, Efendic S, Grill V, Hallqvist J, et al. Socio-economic position at three points in life in association with type 2 diabetes and impaired glucose tolerance in middle-aged swedish men and women : Socioeconomic position and health. International journal of epidemiology. 2007;36(1):84-92.
40. Wu H, Meng X, Wild SH, Gasevic D, Jackson CA. Socioeconomic status and prevalence of type 2 diabetes in mainland China, Hong Kong and Taiwan: a systematic review. J Glob Health. 2017;7(1):011103-.
41. Hsu C-C, Lee C-H, Wahlqvist ML, Huang H-L, Chang H-Y, Chen L, et al. Poverty Increases Type 2 Diabetes Incidence and Inequality of Care Despite Universal Health Coverage. Diabetes Care. 2012;35(11):2286-92.
42. Williams ED, Tapp RJ, Magliano DJ, Shaw JE, Zimmet PZ, Oldenburg BF. Health behaviours, socioeconomic status and diabetes incidence: the Australian Diabetes Obesity and Lifestyle Study (AusDiab). Diabetologia. 2010;53(12):2538-45.
43. Mezuk B, Eaton WW, Albrecht S, Golden SH. Depression and Type 2 Diabetes Over the Lifespan A meta-analysis. Diabetes Care. 2008;31(12):2383-90.
44. Koppes LLJ, Dekker JM, Heniks HFJ, Bouter LM, Heine RJ. Moderate alcohol consumption lowers the risk of type 2 diabetes: a meta-analysis of prospective observational studies. Diabetes Care. 2005;28(3):719-25.
45. Andersen MB, Bjørkman A-SD, Pedersen M, Ekholm O, Molsted S. Social inequality in lifestyle, motivation to change lifestyle and received health advice in individuals with diabetes: A nationwide study. Scand J Public Health. 2020;48(8):847-54.
46. Nunes BP, Thume E, Tomasi E, Silva Duro SM, Facchini LA. Socioeconomic inequalities in the access to and quality of health care services. Rev Saude Publica. 2014;48(6):968-76.
47. Hansen AH, Bradway M, Broz J, Claudi T, Henriksen O, Wangberg SC, et al. Inequalities in the Use of eHealth Between Socioeconomic Groups Among Patients With Type 1 and Type 2 Diabetes: Cross-Sectional Study. J Med Internet Res. 2019;21(5):e13615-e.
48. 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. Int J Epidemiol. 2008;37(1):36-41.
49. Sheikh MA, Lund E, Braaten T. Test–retest reliability of self-reported diabetes diagnosis in the Norwegian Women and Cancer Study: A population-based longitudinal study (n =33,919). SAGE Open Med. 2016;4:2050312115622857-.
50. NOKUT. General information about education in Norway 2022 [cited 2022 1. may]. Available from: https://www.nokut.no/en/norwegian-education/general-information-about-education-in-norway/.
51. IBM. IBM SPSS Statistics 2022 [cited 2022 27. april]. Available from: https://www.ibm.com/products/spss-statistics.

52. Lee TC, Glynn RJ, Pena JM, Paynter NP, Conen D, Ridker PM, et al. Socioeconomic Status and Incident Type 2 Diabetes Mellitus: Data from the Women's Health Study. PLoS One. 2011;6(12):e27670-e.
53. van Zon SKR, Snieder H, Bültmann U, Reijneveld SA. The interaction of socioeconomic position and type 2 diabetes mellitus family history: a cross-sectional analysis of the Lifelines Cohort and Biobank Study. BMJ Open. 2017;7(4):e015275-e.
54. Steele CJ, Schöttker B, Marshall AH, Kouvonen A, O'Doherty MG, Mons U, et al. Education achievement and type 2 diabetes—what mediates the relationship in older adults? Data from the ESTHER study: a population-based cohort study. BMJ Open. 2017;7(4):e013569-e.
55. Demakakos P, Marmot M, Steptoe A. Socioeconomic position and the incidence of type 2 diabetes: the ELSA study. Eur J Epidemiol. 2012;27(5):367-78.
56. Espelt A, Borrell C, Palència L, Goday A, Spadea T, Gnavi R, et al. Socioeconomic inequalities in the incidence and prevalence of type 2 diabetes mellitus in Europe. Gac Sanit. 2012;27(6):494-501.
57. Espelt A, Borrell C, Roskam AJ, Roíguez-Sanz M, Stirbu I, Dalmau-Bueno A, et al. Socioeconomic inequalities in diabetes mellitus across Europe at the beginning of the 21st century. Diabetologia. 2008;51(11):1971-9.
58. Maty SC, Everson-Rose SA, Haan MN, Raghunathan TE, Kaplan GA. Education, income, occupation, and the 34-year incidence (1965–99) of Type 2 diabetes in the Alameda County Study. Int J Epidemiol. 2005;34(6):1274-81.
59. Stene LC, Ruiz PL-D, Åsvold BO, Bjarkø VV, Sørgjerd EP, Njølstad I, et al. Hvor mange har diabetes i Norge i 2020? (how many have diabetes in Norway in 2020?). Tidsskrift Norske Legeforening 2020. 2020(Utgave 17, 24. november 2020).
60. Rylander C, Sandanger TM, Engeset D, Lund E. Consumption of Lean Fish Reduces the Risk of Type 2 Diabetes Mellitus: A Prospective Population Based Cohort Study of Norwegian Women. PLoS One. 2014;9(2):e89845.
61. Langholz PL, Wilsgaard T, Njolstad I, Jorde R, Hopstock LA. Trends in known and undiagnosed diabetes, HbA1c levels, cardiometabolic risk factors and diabetes treatment target achievement in repeated cross-sectional surveys: the population-based Tromso Study 1994-2016. BMJ Open. 2021;11(3):e041846-e.
62. Elstad JI, Helsedirektoratet. Utdanning og helseulikheter : problemstillinger og forskningsfunn (education and health inequalities : issues and research findings). Oslo: Helsedirektoratet; 2008.
63. Cutler DM, Lleras-Muney A. Understanding differences in health behaviors by education. J Health Econ. 2010;29(1):1-28.
64. Ardisson Korat AV, Willett WC, Hu FB. Diet, Lifestyle, and Genetic Risk Factors for Type 2 Diabetes: A Review from the Nurses' Health Study, Nurses' Health Study 2, and Health Professionals' Follow-Up Study. Curr Nutr Rep. 2014;3(4):345-54.
65. Norberg M, Stenlund H, Lindahl B, Andersson C, Eriksson JW, Weinehall L. Work stress and low emotional support is associated with increased risk of future type 2 diabetes in women. Diabetes Res Clin Pract. 2006;76(3):368-77.
66. Grundy SM, Benjamin IJ, Burke GL, Chait A, Eckel RH, Howard BV, et al. Diabetes and cardiovascular disease : A statement for healthcare professionals from the American heart association. Circulation. 1999;100(10):1134-46.
67. WHO. Urban Health 2021 [updated 29 october 2021; cited 2022 15 march]. Available from: https://www.who.int/news-room/fact-sheets/detail/urban-health.
68. Pechey R, Monsivais P. Socioeconomic inequalities in the healthiness of food choices: Exploring the contributions of food expenditures. Prev Med. 2015;88:203-9.

69. Pouwer F, Kupper N, Aiaanse MC. Does emotional stress cause Type 2 Diabetes Mellitus? A review from the European Depression in Diabetes (EDID) Research Consortium. Discov Med. 2010;9(45):112-8.
70. Mirowsky J, Ross CE. Education, learned effectiveness and health. London review of education. 2005.
71. Mackenbach JP, Kunst AE. Measuring the magnitude of socio-economic inequalities in health: An overview of available measures illustrated with two examples from Europe. Soc Sci Med. 1997;44(6):757-71.
72. 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.
73. 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(1):2-.
74. Lund E, Kumle M, Braaten T, Hjartåker 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. 2003;14(10):1001-8.
75. Lindsted KD, Fraser GE, Steinkohl M, Beeson WL. Healthy volunteer effect in a cohort study: Temporal resolution in the adventist health study. J Clin Epidemiol. 1996;49(7):783-90.
76. Heffermehl KB. The Status of Women in Norway. The American journal of comparative law. 1972;20(4):630-46.
77. Tripepi G, Jager KJ, Dekker FW, Wanner C, Zoccali C. Bias in clinical research. Kidney Int. 2008;73(2):148-53.
78. Timothy R. Graeff. Response Bias,. ncyclopedia of Social Measurement,. 2005:411-8.
79. Nyholm M, Gullberg B, Merlo J, Lundqvist‐Persson C, Råstam L, Lindblad U. The Validity of Obesity Based on Self-reported Weight and Height: Implications for Population Studies. Obesity (Silver Spring). 2007;15(1):197-.
80. 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.
81. 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(1):4-.
82. Naseribafrouei A, Eliassen B-M, Melhus M, Svartberg J, Broderstad AR. The prevalence of diabetes mellitus among Sami and non-Sami inhabitants of Northern Norway - the SAMINOR 1 Survey (2003-2004) and the SAMINOR 2 Clinical Survey (2012-2014). Rural Remote Health. 2018;18(4):4623-.
83. Harris ML, Oldmeadow C, Hure A, Luu J, Loxton D, Attia J. Stress increases the risk of type 2 diabetes onset in women: A 12-year longitudinal study using causal modelling. PLoS One. 2017;12(2):e0172126-e.
84. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019;34(3):211-9.
85. Bays HE, Chapman RH, Grandy S. The relationship of body mass index to diabetes mellitus, hypertension and dyslipidaemia: comparison of data from two national surveys. Int J Clin Pract. 2007;61(5):737-47.
86. Carey VJ, Walters EE, Colditz GA, Solomon CG, Willett WC, Rosner BA, et al. Body fat distribution and risk of non-insulin-dependent diabetes mellitus in women : The Nurses' Health Study. Am J Epidemiol. 1997;145(7):614-9.

87. Yu H-J, Ho M, Liu X, Yang J, Chau PH, Fong DYT. Association of weight status and the risks of diabetes in adults: a systematic review and meta-analysis of prospective cohort studies. Int J Obes (Lond). 2022.

Appendix 1: Table 3 Baseline charateristics by diabetes status in NOWAC (N=100 347)

	Diabetes NO			Diabetes YES		
	Valid	Missing	Mean (SD/%)	Valid	Missing	Mean (SD/%)
Age	98900(98.6)	0	48.31 (8.3)	1447(1.44)	0	53.6 (8.3)
Smoking status	97377(98.5)	1523 (1.5)	1.95(0.8)	1430(98.8)	17 (1.2)	1.9(0.8)
- Never	35231 (36.2)			542 (37.9)		
- Former	32258 (33.1)			485 (33.9)		
- Current	29888 (30.7)			403 (28.2)		
BMI groups[1]	64062 (96.2)	3766(3.8)		1374 (95)	73 (5)	
- Normal	64062 (67.3)			476 (34.6)		
- Overweight	24210 (25.4)			447 (32.5)		
- Obese	6862 (6.9)			451 (32.8)		
Education[2]	93997 (95)	4903 (5)		1345 (93)	102 (7)	
- Primary	22789 (24.2)			490 (39.2)		
- Secondary	32608 (34.7)			456 (31.5)		
- Tertiary	38600 (41.1)			399 (29.3)		
Physical activity[3]	90098 (91.1)	8802 (8.9)		1275 (88.1)	172 (11.9)	
- Low	11356 (12.6)			257 (20.2)		
- Moderate	6338 (70.4)			855 (87.2)		
- High	15354 (17.0)			163 (12.8)		
Alcohol consume[4]	92527 (93.6)	6373(6.4)	2.147 (0.9)	1329(91.8)	118(8.2)	1.83(0.8)
- Never	22651 (24.5)			531 (40)		
- <median	49987 (54.0)			647 (48.7)		
- >median	19889 (21.5)			151 (11.4)		
Total fruit*	95904	2996	202.12 (151.5)	1371	76	195.05(151.1)
Total fiber*	98770	130	21.18 (7.2)	1442	5	20.23(7.8)
Total sugar*	98696	204	23.93 (16.1)	1438	9	15.91(13.2)
Total vegetables*	97490	1410	161.06 (109)	1406	41	159.36 (119.2)
CVD[5]	98900	0	0.14 (0.3)	1447	0	0.52(0.5)

[1]BMI kg/m^2: Normal >25 kg/m^2, overweight 25-29.9 kg/m^2 and obese <30 kg/m^2.
[2]Education level: Primary ≤9 years, secondary 10-12 years and tertiary ≥13 years
[3]Physical activity on a 10-scale: low 1-3, moderate 4-7 and high 8-10.
[4]Alcohol consumption divided into: non-consumer, <median: 0.1-2.7 grams/day and >median: >2.7 grams/day.
[5]CVD cases consist of self-reported illness: high blood pressure, heart attack, heart failure, heart cramps and stroke.
*Measured as grams per/day

Appendix 2: An example of NOWAC questionnaire

Høst 2003

KVINNER OG KREFT

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 ☐

Appendix 3: Table 4: Baseline characteristic on Education level – Supplementary table

Education[1]:	Primary			Second			Tertiary		
	Valid	Missing	Mean (SD)	Valid	Missing	Mean (SD)	Valid	Missing	Mean(SD)
Age	23279	0	50.83 (9)	33064	0	47 (8.1)	38999	0	46.9 (7.7)
Diabetes	23279	0		33064	0		38999	0	
- Yes	490 (2.1)			456 (1.4)			399 (1)		
- No	22789 (97.9)			32608 (98.6)			38600 (99)		
Diabetes age	372	22907	46.5 (13.7)	394	32670	41.24 (14.6)	351	38648	39.71 (15.1)
Smoking	22832	447		32632	432		38516	483	
- Never	7015		30.7	10504		32.2	16476		42.8
- Former	7091		31.1	10768		33	13317		34.6
- Current	8726		38.2	11360		34.8	8723		22.6
Physical activity[2]	19557 (84)	3722 (16)		30506 (92.3)	2558 (7.7)		37314 (95.7)	1685 (4.3)	
- Low	3048		15.6	3810		12.5	4217		11.3
- Medium	12876		65.8	21595		70.8	27000		72.4
- High	3633		18.6	5101		16.7	6097		16.3
BMI[3]	18706	4573		23146	9918		25054	13945	
- Normal	11311		60.5	16595		71.7	19655		78.5
- Overweight	5677		30.3	5204		22.5	4428		17.7
- Obese	1718		9.2	1347		5.8	971		3.9
Alcohol consume[4]	21664	1615		31264	1800		37124	1875	
- Never	7700		33.3	7319		22.7	7054		18.7
- <median	11187		41.8	17547		40.7	19840		36.4
- >median	2777		24.9	6398		36.6	10230		44.9
CVD[5]	23279	0	0.18 (0.4)	33064	0	0.13 (0.3)	38999	0	0.10 (0.3)
- Yes	4218		18.1	4429		13.4	3971		10.2
- No	19061		81.9	28635		86.6	35028		89.8
Total sugar[*]	23279		24.19 (17.4)	33064		24.01 (16.4)	38999		23.74 (15)
Total Fiber [*]	23279		19.95 (7.4)	33064		21.08 (7.1)	38999		22.17 (6.9)

Total Fruit*	23279	182.6 (147.9)	33064	198.58 (150.2)	38999	216.38 (151.9)
Total vegetables*	23279	134.2 (99.9)	33064	157.82 (105.1)	38999	181 (113.3)

[1]Education level: Primary ≤9 years, secondary 10-12 years and tertiary ≥13 years.
[2]Physical activity on a 10-scale: low 1-3, moderate 4-7 and high 8-10.
[3]BMI kg/m^2: Normal >25 kg/m^2, overweight 25-29.9 kg/m^2 and obese <30 kg/m^2.
[4]Alcohol consumption divided into: never- non-consumer, moderate: >0-5 grams/day and high: >5 grams/day.
[5]CVD cases consist of self-reported illness: high blood pressure, heart attack, heart failure, heart cramps and stroke
[*]Measured in grams/day