

Derivation and validation of German risk scores predicting substantial gain in weight and waist circumference

vorgelegt von

Mag. rer. nat. in Ernährungswissenschaften, M. Sc. in Epidemiology

Ursula Bachlechner

geb. in Rum

von der Fakultät VII – Wirtschaft und Management

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktorin der Gesundheitswissenschaften / Public Health

– Dr. P.H. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzende: Prof. Dr. Elke Schöffner

Gutachter: Prof. Dr. Reinhard Busse

Gutachter: Prof. Dr. Heiner Boeing

Tag der wissenschaftlichen Aussprache: 23. Juni 2016

Berlin 2016

Table of Contents

Index of Tables	V
Index of Figures	VI
List of Abbreviations	VII
1. Introduction	1
1.1 Background and aim of the thesis	1
1.2 General and abdominal overweight and obesity	3
1.2.1 Energy balance and regulation of body fat mass	4
1.2.2 Risk factors for gain in weight and waist circumference	6
1.3 Prognostic prediction models	14
1.3.1 Measures of model performance	16
1.3.2 Risk scores for gain in weight and waist circumference in adults: state of research	21
1.4 Public health relevance of the thesis	22
1.5 Challenges and research questions of the thesis	23
2. Material and methods	25
2.1 EPI Germany Consortium: the subproject 2	25
2.1.1 Assessment of relevant variables	30
2.1.2 Analytical study population	33
2.2 Statistical approach	35
2.2.1 Definition of study variables	35
2.2.1.1 Case status	35
2.2.1.1 Predictor variables	37
2.2.2 Descriptive statistics	39
2.2.3 Construction and evaluation of the German risk scores	39
3. Results	42
3.1 Description of the study population	42
3.1.1 General characteristics	42
3.1.2 Incidences of SWG and SWCG	45
3.2 Derivation of the German risk scores	47
3.2.1 Description of the German risk scores	47
3.2.2 Predictive performances of the German risk scores	55
4. Discussion	64
4.1 Material and methods	65

Table of Contents

4.1.1 Study design and study population _____	65
4.1.2 Data quality of predictors and outcomes _____	65
4.1.3 Methodological approach _____	67
4.2 Results _____	70
4.2.1 Associations of predictors with substantial gain in weight and waist circumference ____	70
4.2.2 Predictive performances of the German risk scores _____	73
4.3 Strengths and limitations _____	76
4.4 Implications for public health _____	77
4.5 Conclusion and outlook _____	79
Summary _____	80
Zusammenfassung _____	82
References _____	85
Appendix _____	105
Danksagung _____	115
Eidesstattliche Erklärung _____	116

Index of Tables

Table 1. Cohorts involved in subproject 2 of the EPI Germany Consortium of the CNO	25
Table 2. Assessment of study-relevant data across the cohorts	32
Table 3. Predictor variables and their corresponding scales	38
Table 4. General characteristics of the study population	44
Table 5. Follow-up times, cases of SWG and SWCG during total follow-up and within the first five years	46
Table 6. Meta-analytically combined estimates of relative risk for the associations of predictors with SWG and SWCG in the maximum models	50
Table 7. Meta-analytically combined estimates of relative risk for the associations of predictors with SWG and SWCG in the minimum models	51
Table 8. Meta-analytically combined estimates of relative risk for the associations of predictors with SWG and SWCG in the selection models	52
Table 9. Predictive performances of the risk score across the cohorts	56
Table 10. Sensitivity, specificity, PPV, and NPV for various cut-off points for the selection model of SWG	58
Table 11. Sensitivity, specificity, PPV, and NPV for various cut-off points for the selection model of SWCG	59

Index of Figures

Figure 1. Examples of receiver operating characteristic curves ranging from acceptable to outstanding predictive values	18
Figure 2. Example of a calibration plot of observed risk against predicted risk across deciles of predicted risk	20
Figure 3. Flow diagram of participants excluded from the present study	34
Figure 4. Incidence rates of SWG and SWCG (per 10,000 person-years)	47
Figure 5. Associations of age and body weight with risk for SWG (a) and SWCG (b) in the selection models	49
Figure 6. Meta-analytically combined aROCs of the selection models for the prediction of SWG (left) and SWCG (right) over five years	57
Figure 7. Calibration plots for SWG using the selection model	62
Figure 8. Calibration plots for SWCG using the selection model	63

List of Abbreviations

BGS98	<i>Bundes-Gesundheitssurvey 1998</i> (German National Health Interview and Examination Survey 1998)
BMI	Body mass index
BMR	Basal metabolic rate
CARLA	Cardiovascular disease, living and ageing in Halle
CI	Confidence interval
CNO	Competence Network Obesity
DEGS	<i>Deutsche Erwachsenen-Gesundheits-Studie</i> (National Health Interview and Examination Survey for Adults)
EPIC	European Prospective Investigation into Cancer and Nutrition
FFQ	Food frequency questionnaire
HLT	Hosmer-Lemeshow test
KORA	<i>Kooperative Gesundheitsforschung in der Region Augsburg</i> (Cooperative Health Research in the Region of Augsburg)
MONICA	Monitoring of Trends and Determinants of Cardiovascular Disease
PH	Proportional hazard
RSF	Random survival forest
SHIP	Study of Health in Pomerania
SWCG	Substantial waist circumference gain
SWG	Substantial weight gain
TEE	Total energy expenditure
WC	Waist circumference
WC _{BMI}	Residual waist circumference
WHO	World Health Organization

1. Introduction

1.1 Background and aim of the thesis

According to the most recent representative survey, more than 60 % of German adults are overweight or obese (1). During the last decade the prevalence of overweight remained stable at a high level, while the prevalence of obesity increased further (1). Overweight and obesity are unequivocally linked to elevated risks of morbidity and premature mortality (2, 3). Excessive body fat accumulation has become a well-established risk factor for many non-communicable chronic health disorders such as cardiovascular diseases, type-II-diabetes, chronic respiratory diseases, musculoskeletal disorders, and several types of cancer (3, 4). Owing to its associations with metabolic abnormalities, accumulation of abdominal visceral fat, in particular, has attracted scientific interest in the development of non-communicable diseases (5). In 2014 the entirety of non-communicable diseases accounted for 91 % of total deaths in Germany (6). As a consequence, excessive body fat accumulation is accompanied by large social and economic burdens (7). According to current estimations, excess body fat costs nearly 16.8 billion euros in total expenditure that is distributed in more or less equal parts as direct costs (including costs for treatment, rehabilitation, and non-medical costs) and indirect costs (including costs for sickness absence, early retirement and mortality) (8).

In view of a high prevalence of excessive body fat accumulation in the German population, the related burden of morbidity and mortality as well as the associated social and economic impacts, further development of effective prevention measures is of major public health interest. Besides the fact that changes in body fat mass principally result from a long-term imbalance of energy intake and energy expenditure, it is well accepted that susceptibility and extent of body fat accumulation result from a complex interaction between genetic predisposition, hormonal influences, and environmental, sociocultural, psychosocial and behavioral factors (7). For the development of successful prevention strategies it is essential to consider the multifactorial nature of body fat accumulation. Risk scores combine multiple factors (predictors) to estimate the probability of the occurrence of a future event. They generally aim to predict the individual risk of an event as precisely as possible on the basis of limited information. Such tools have been developed in the medical field to identify high-risk

individuals of various chronic health disorders (including cardiovascular diseases, several types of cancer, and type-II-diabetes) to assign targeted measures of health promotion (9-16).

In 2013 the first risk score predicting weight gain in European middle-aged adults, which used the large-scale data of the multi-center European Prospective Investigation into Cancer and Nutrition (EPIC) study, was published (17). It is a simple public health instrument based on 13 easily obtainable predictor variables; it comprises socio-demographic and anthropometric characteristics as well as dietary and lifestyle factors. Against the background of a varying prevalence of overweight and obesity as well as variable effects of environmental, sociocultural, and behavioral factors on weight gain between European countries (7), country-specific risk scores were proposed to be more promising than the universal ones for large heterogeneous populations (9). Considering this, the European risk score was adapted to German population using the data of the two German EPIC cohorts in the run-up to the present thesis. The adaptation procedure revealed that the predictive performance of the risk score, adapted to Germany-specific characteristics and behaviors, was not superior to the performance of the universal, transnational European risk score. In addition to the adaptation procedure, an extension of the risk score by supplementary weight gain-related factors was performed. Although the extension was able to improve the scores' predictive performance in a cohort-specific manner, the improvements were not reproducible in the independent German sample. Concerning localization of weight gain, such as specific abdominal fat accumulation, no risk prediction tool was ever derived among adults. In view of this preliminary state of research, the aim of the present thesis is to derive pragmatic, informative, and Germany-wide valid risk scores predicting substantial gain in overall body fat mass and specific abdominal body fat accumulation. Therefore, it will define complementary anthropometric measures, including meaningful thresholds, and overcome the problem of over-adaptation relating to a pre-defined derivation sample. Furthermore, it will investigate appropriate, preferably minimized predictor sets leading to the best possible predictive performances in the German population.

1.2 General and abdominal overweight and obesity

Overweight and obesity are defined as conditions of excessive accumulation of body fat mass, which can be detrimental to human health (4). Although accurate, direct methods for the assessment of body fat mass are available, the measurement of body fat mass by these methods is often time-consuming and expensive. Therefore, surrogate measures are widely used. The two most commonly used anthropometric measures for the assessment of body fat are body mass index (BMI) and waist circumference (WC). BMI is an indicator of overall body fat mass and is calculated as body weight in kilograms (kg) divided by the square of body height in meters (m^2). According to the World Health Organization (WHO), general overweight and obesity are classified by BMI equaling to or greater than 25 kg/m^2 and 30 kg/m^2 , respectively (4). These cut-offs are equal to men and women regardless of their age (18). Even though BMI cannot distinguish between weight associated with muscle and weight associated with fat, research has shown that there is a strong correlation between BMI and gold standard methods for measuring body fat mass such as underwater weighing and dual energy X-ray absorptiometry (19-21). Hence, BMI is generally described as an easily measured, inexpensive, non-invasive, and valid anthropometric indicator of overall body fat mass. However, BMI does not provide any information on localization of body fat. The anthropometric assessment of body fat localization is commonly measured by body circumferences. Owing to its close correlation to total abdominal fat mass measured by computer tomography, WC is often used as a surrogate indicator of abdominal (visceral) fat accumulation (22), which was suggested to play a prominent role regarding the risk of metabolic disorders associated with body fat (5, 7, 18). WC is usually measured at the natural waist (between the lowest rib and the top of the hip bone), the umbilicus (belly button), or at the narrowest point of the midsection (18). Based on epidemiological data, cut-offs for WC were defined by the risk of metabolic complications associated with excessive abdominal fat mass (23). It was supposed that Caucasian men with WC equal or greater than 92 cm and Caucasian women with WC equal or greater than 80 cm are at increased risk (referred to as abdominal overweight), while Caucasian men with WC equal or greater than 102 cm and Caucasian women with WC equal or greater than 88 cm are at substantially increased risk of metabolic complications (referred to as abdominal obese) (23, 24). Usually, WC and BMI are strongly correlated anthropometric measures (25). Regarding

the prediction of diseases associated with body fat, research suggests that WC may provide additional information beyond BMI, particularly when WC disproportionally increases (26). Consequently, it has become a common practice to measure both BMI and WC as well as to combine the gathered information on overall body fat mass and abdominal body fat accumulation to estimate health risks.

1.2.1 Energy balance and regulation of body fat mass

Body weight is determined by the balance between energy intake and energy expenditure. Energy intake is derived from ingestion of macronutrients, i.e., fats (37 kJ/g), carbohydrates (17 kJ/g), proteins (17 kJ/g), and alcohol (29 kJ/g), all provided by food and drinks (27). The energy in those compounds is converted by a biochemical cascade into adenosine-triphosphate that enables human cells to meet energy expenditure to maintain structural and functional integrity of the body. Total energy expenditure (TEE) comprises basal metabolic rate (BMR), non-resting energy expenditure (= physical activity), and (postprandial) thermogenesis. BMR corresponds to the energy required for the maintenance of cell-renewing processes, osmotic regulation, and the muscular functions of the internal organs of an awake individual being in dorsal position, 10 to 12 hours after last food intake, eight hours after physical rest, in the thermos-neutral environment, and in a relaxed mental state (27). In adults BMR accounts for 45 to 70 % of TEE and principally depends on age, sex, body height, and body composition (27). Another 20 to 45 % of TEE is allocated to physical activity, the most variable component of daily energy expenditure (27). Approximately 10 % of TEE is assigned to thermogenesis comprising energy-consuming processes during food digestion, nutrient absorption and storage, and the energy used for regulatory processes to keep body temperature constant (27). When energy intake equals energy expenditure in the long-term, body weight remains stable. Changes in body weight only occur in case of positive energy balance (favoring energy intake) or negative energy balance (favoring energy expenditure), resulting in weight gain or weight loss, respectively.

Despite considerable random fluctuations in dietary intake on a daily basis, body weight (body fat mass) in healthy individuals is remarkably stable over time. This stability is explained by a complex regulatory system, referred to as energy homeostasis (28). The

entirety of functional mechanisms of this regulatory system is still incompletely understood. Till date, a number of theories have been postulated to explain the matching of energy intake and energy expenditure over prolonged periods of time - e.g., the set point model, the settling point model, and the general intake model. The set point model hypothesizes that there is an active negative-feedback mechanism, which connects energy storages (body fat mass) with energy intake and expenditure by a target set point (29). According to the model, energy homeostasis is based on circulating signals that communicate information about the present state of energy storages to brain areas comprising hypothalamus and brainstem. In response to this input, the brain compares the received signals with a target level of body fatness. In case of any discrepancy, adaptive adjustments would be induced by changes in energy intake or energy expenditure (29). The set point model is in line with many physiological aspects of energy regulation including the high genetic contribution to the variation in BMI (30-32) and the phenomenon of weight regain after a period of dieting (33). The model, however, does not explain the enormous increase in the prevalence of overweight and obesity taking place since the 1980s (4). Furthermore, the model does not take into account the impact of environmental, behavioral, socioeconomic, or psychosocial factors in the development of overweight and obesity (34). In view of these aspects, the settling point model has been suggested. This model is based on the idea of a passive feedback system. The main characteristic of the model is that a parameter of interest (energy stores) has both inputs (energy intake) and outputs (energy expenditure) (34). Importantly, one of these parameters (in this case energy intake) must be independent of the size of the parameter of interest and the other must change in direct relation to the size of this parameter (in this case, energy expenditure). So the level of energy storages settles to a natural equilibrium, which is determined by the energy intake that matches the energy expenditure because the latter is passively related to the level of energy storages. In other words, as body fat mass grows, the rate of energy expenditure also grows, due to a growth in lean body mass that is necessary to support the moving of the resulting larger overall body mass (34). Under the settling point model the rapid increase in the prevalence of overweight and obesity during the last few decades is mainly explained as a consequence of current environmental changes including increased availability of and easier access to palatable energy-rich food and drinks as well as the less physical energy needed to obtain them, also known as the “obesogenic environment” (7, 34). However, an environmentally

determined settling point model cannot adequately explain the variability of individual susceptibility of weight gain in a common environment (34). With respect to the shortcomings of the set point model and the settling point model, the general model of intake regulation has been proposed (35). This model combines elements of both models to be a comprehensive model of energy intake and body weight regulation (34). Under the general model of intake regulation intake of food and drinks is influenced by a large variety of physiological, environmental, social, psychological, and behavioral factors. The model divides these factors mainly into compensated (primarily physiological) and uncompensated (primarily environmental) factors. The main difference between these kinds of factors is that compensated factors simultaneously affect and are affected by intake of food and drinks, while uncompensated factors affect but are not affected by such intake. Moreover, each factor is thought to contribute only to a small proportion of total variance in intake and the impact of the factors may also vary on individual basis, which implies that they are affected by genes. The model states that intake of food and drinks corresponds to the sum of all compensated and uncompensated factors, without assuming the presence of any set point for energy intake or body fatness. However, if all influencing factors are quite stable over a longer period of time, the model would act as if there is something like a set point (34).

1.2.2 Risk factors for gain in weight and waist circumference

To comprehensively study the etiology of overweight and obesity, multiple factors directly or indirectly affecting the balance between energy intake and energy expenditure have to be considered. Even if the effect of any single factor may be limited, the combined effect of a wide range of factors appears to be crucial (36). Besides genetic predisposition and hormonal changes, social factors including socio-demographic characteristics and psychosocial influences - which are in part highly interrelated with behavioral factors such as diet and lifestyle - may have substantial impact on the development of excess body fat. In view of further development of strategies relating to primary prevention based on non-invasive and easily obtainable information, a magnitude of influencing factors involved in the complex interplay leading to general and abdominal overweight and obesity can be considered. The spectrum of social and behavioral risk factors for gain in weight and WC,

included in the German risk scores derived and validated in the present thesis, will be briefly discussed in the following.

Social factors

Social factors, including socio-demographic characteristics (e.g., age and educational level), and psychosocial factors (e.g., living in partnership and life satisfaction) may play a considerable role in the development of excess body fat mass. While some of these factors mainly result in increased body fat mass through their strong interrelations with dietary and lifestyle habits, others have primarily an inherent physiological background, or involve both.

Socio-demographic characteristics

A well-established physiological mechanism contributing to **age**-related gain in body fat mass is a decline in BMR. BMR declines as a result of age-related decreases in lean body mass, including muscles and bones, leading to decreased TEE (37, 38). Besides the physiological reduction of energy expenditure during aging, also age-related changes in physical activity patterns towards a more sedentary lifestyle may further promote this development. Moreover, during the aging process body fat mass accumulates specifically in the abdominal region (37), presumably due to the underlying hormonal changes. For instance, studies show that decrease in testosterone and increases in body fat mass reinforce each other in aging men, leading to a spiral of gain in body fat mass and hormonal disturbance (39, 40). Additionally, low testosterone levels have been specifically associated with abdominal obesity (41). In middle-aged women decrease in circulating estradiol has been related to decreased energy expenditure and fat oxidation as well as increased abdominal fat mass in the course of menopausal transition (42, 43).

Educational level represents an indicator of socio-economic status going along with the acquisition of knowledge and competences that enable individuals to integrate a health-promoting lifestyle. According to a comprehensive literature overview, years of schooling have been referred to as the most important factor relating to good health (44). In line with this, available evidence regarding higher educational level and excess body fat show strong

inverse associations (45-47); studies investigating a causal effect, however, are limited and report more mixed results (45, 48-50). Cutler and Lleras-Muney found that highly educated individuals are less likely to be overweight or obese; furthermore, it has been observed that they are also less likely to smoke, but more likely to exercise (51). A German cross-sectional survey confirms that individuals having a lower level of school education are more likely affected by excess body weight (47). Also, Webbink et al. demonstrate in their cross-sectional twin study a negative relationship between education and the probability of being overweight (52). Regarding abdominal body fat accumulation, evidence supporting an inverse relationship with education has been mainly found for women (46, 53, 54).

Psychosocial factors

There has been much research on the investigation of benefits of social support, as a general term covering resources provided by others (e.g., **partner**, children, friends, colleagues) (55) on health outcomes (56, 57). But only a few studies report independent associations with outcomes related to excessive body fat mass. These studies mainly explore the importance of interpersonal relationships in connection with the maintenance of weight management efforts (58, 59). Nevertheless, there is evidence suggesting an inverse relationship between social support and risk of excess body fat, which appears to be particularly present in women (60, 61). A Finnish study, based on middle-aged adults, observes an inverse association between social support and 12-month weight gain only among women (60). Furthermore, such inverse association has also been seen with gain in WC among postmenopausal women (62). Presumably, lower levels of social support are also highly interrelated with less favorable behavioral factors such as unhealthier food purchasing behavior and lower access to physical activity facilities. In addition, lower levels of social support may further be related to stress-related hormonal imbalances (63), favoring an increase in body fat mass.

Life satisfaction is another meaningful psychosocial construct related to health and well-being, which reflects the subjective grading of an individual's feelings, attitudes and behaviors ranging from positive to negative (64). To some differing extent, a wide range of factors may be involved in determining an individual's life satisfaction such as individual needs and desires, existing interpersonal relationships, physical and mental health, and

professional success (65). Increased life satisfaction has been positively related with self-esteem and psychological well-being while negatively related with chronic stress, anxiety and depression (66-68). Regarding overweight and obesity, most research in this field focuses on decreased life satisfaction as influenced by chronic stress and depression. Dysregulation of the glucocorticoid hormone cortisol has been suggested as the underlying biological mechanisms for the association of chronic stress, and depression with excess body fat accumulation, especially in the abdominal region (69-74). Additionally, behavioral factors are found to have been involved in the association. Both chronic stress and depression may induce increased consumption of energy-dense, palatable food as well as reduced physical activity by elevated secretion of glucocorticoids (63, 70), which in turn, may stimulate pleasure centers in the brain which regulate mood (70).

Behavioral factors

Dietary and lifestyle behaviors are thought to be the most important contributors to the substantial rise in excessive fat accumulation during the last few decades (36, 75-77). Increased portion sizes, higher availability, and facilitated access to inexpensive, palatable, and energy-dense food as well as reduced physical activity at work and during leisure time shifting towards reduced energy expenditure (78, 79), the so-called “obesogenic environment”, lay the ground for behavioral patterns that promote gain in body fat mass. At the same time, numerous psychological, social, and cultural factors have important underlying influences on dietary and lifestyle behaviors, finally determining or preventing the development of overweight and obesity.

Dietary factors

Fruits and **vegetables** are rich in water and low in energy; additionally, they contain high amounts of dietary fiber, which has been shown to promote satiety and to modulate hormonal responses, resulting in fewer total energy intake consumed (80-82). Because of the displacement of energy-dense food, higher intake of fruits and vegetables has been proposed as safeguard against excessive gain in body weight (83). Evidence from observational and experimental studies mainly supports the protective effect of a diet with

lots of fruits and vegetables, but such effects are generally weak and not completely consistent (84-88).

There is growing evidence that **whole grain products** protect against the development of overweight and obesity (82, 89). Various nutritional components of whole grains have been proposed to be involved in physiological mechanisms promoting weight management due to their effects on satiety, satiation, and hormonal responses (82). Besides valuable micronutrients, such as diverse minerals, trace-elements, and vitamins, whole grain products are particularly characterized by high contents of dietary fiber and generally low values of glycemic index (82). Dietary fiber has been shown to increase satiety and satiation, to reduce transit time, and to stimulate gut hormones (82, 90), whereas low glycemic index results in lower postprandial glucose responses and insulin demand that may have a regulative impact on appetite in further consequence (82). In intervention and observational studies the intake of whole-grain products has been inversely related to plasma levels of obesity biomarkers including insulin, C-peptide, and leptin (91, 92). There is limited epidemiological evidence explicitly investigating whole grain consumption and gain in body fat mass. Existing studies, however, generally show a decreased risk of excessive body fat mass for higher intake of whole grain products (89, 93).

Foods of animal origin, such as **meat** and **fish**, are rich sources of dietary protein, but they frequently also contain high amounts of fat resulting in high energy density. Protein-rich diet has been hypothesized to have preventive effects on weight gain and beneficial effects for weight management due to increased protein-induced thermogenesis and satiety (94, 95). According to the conclusion of a comprehensive review on the epidemiological evidence comprising eight cohort studies, however, no associations between dietary protein intake and subsequent changes in body weight or WC have been found, even though the results have been not fully consistent (96). One study found higher gain in weight and waist-to-hip ratio, while another study found lower gain at the waist to be associated with higher protein intakes. Observational studies differentiating between protein sources predominantly found positive associations with excess body fat, which have mainly been attributed to the intake of red and processed meat as well as poultry than to the intake of fish and dairy sources (97-101). The results of intervention studies regarding protein sources, however, are more mixed (102-104).

Sweets, such as **chocolate**, **cakes**, and **cookies**, are usually rich in fat and sugar, thus resulting in high energy-density. Consequently, higher consumption of sweets has been hypothesized to promote gain in body fat mass. Available evidence, however, is limited and reported findings were inconsistent. Cross-sectional studies find inverse and no associations regarding higher intake of chocolate and candy (105, 106), whereas prospective studies find no associations for higher consumption of sweets (107), positive association for higher chocolate consumption with long-term weight gain (108), and inverse associations of higher desserts and candy consumption with four-year weight gain (109).

There is cumulative evidence showing positive associations of greater consumption of sugar-sweetened beverages (**soft drinks**) such as colas, fruit drinks, and lemonades, with excess body fat mass (110, 111). Numerous cross-sectional studies observe positive trends in the relation between sugar-sweetened beverage consumption and overweight or obesity (112-115). Additionally, prospective studies carrying out repeated measurements of sugar-sweetened beverages consumption and weight observe that increased intakes of sugar-sweetened beverages are significantly associated with greater weight gain and greater obesity risk over time (116, 117). Large amounts of sugar, low satiety of liquid carbohydrates, and the resulting incomplete compensation for total energy intake have been proposed as the underlying physiological mechanisms by experimental studies (118, 119).

Lifestyle factors

Similar to sugar-sweetened beverages, energy provided by **alcoholic beverages** appears to act additively to energy from solid foods. Several studies suggest that energy provided by alcohol before or during a meal is not compensated by that meal (120), which results in excess energy intake. In addition, alcohol consumption seems to have little impact on satiety but further stimulates food intake (120). Biological as well as psychosocial mechanisms have been proposed to highlight the stimulating effect of alcoholic beverages on food intake, including reduced levels of plasma glucose resulting from reduced gluconeogenesis (121, 122), enhanced socialization, psycho-emotional disinhibition, and prolonged duration of food intake (123, 124). Besides influences on food intake, alcohol consumption also affects

energy storage. Alcohol suppresses fat oxidation, which suggests that frequent intake of alcoholic beverages could lead to higher body fat storages in the long-term (125). However, alcohol may also affect energy expenditure due to its high thermogenic effect (126) and as a function of alcohol degradation (127). Most recent evidence from prospective studies suggests no associations between light-to-moderate alcohol intake and gain in body weight or WC (128), while associations have been consistently found between heavy alcohol consumption and body weight gain (129-131). Also, there is mixed evidence from experimental studies, but it has been suggested that moderate alcohol consumption does not lead to weight gain over the short-term (132). Multiple factors may contribute to the mixed evidence of alcohol in promoting body weight gain, including variations in frequency, type and amount of alcohol consumption, different drinking pattern, genetic predisposition to alcohol-related weight gain, composition of diet, and also individual sleeping habits, level of physical activity, and education (133-135).

In several cross-sectional studies current **smoking** has been associated with lower body weight or BMI in comparison with non-smoking (136-138). These associations have been suggested as a result of nicotine-induced increased energy expenditure and inhibited appetite (139), which may also be involved in weight gain following smoking cessation (140, 141). Regarding smoking intensity, however, it has been shown that individuals smoking higher numbers of cigarettes daily tend to have higher body weight than individuals who smoke less (142-144); this association has been hypothesized to be attributable to a clustering of unfavorable behavior of heavy smokers, including less physical activity (145), higher alcohol consumption (146), and unhealthier diet (147, 148). Additionally, smoking has been shown to be particularly associated with fat accumulation in the abdominal region (149, 150).

The substantial rise in the prevalence of overweight and obesity over the last few decades has been paralleled by a concurrent rise in chronic **sleep deprivation** (151-153). Accumulating evidence supports the role of short sleep duration as a risk factor for weight gain and obesity (151, 152, 154, 155). Many cross-sectional studies report short sleep duration to be associated with higher body weight (151, 156). It has been further observed that this association lessened with age (154, 157). Some studies also find associations between long sleep duration and elevated body weight, resulting in a U-shaped sleep-weight

relation, but available evidence is more mixed (151). Prospective studies generally support the role of short sleep duration to be associated with weight gain and obesity, whereas long sleep duration is not reported to be associated with obesity risk (151, 155). Several biological mechanisms have been hypothesized to link short sleep duration and weight gain, which include both increased energy intake and reduced energy expenditure. Increased energy intake has been proposed as a consequence of multiple hormonal imbalances (e.g., increased ghrelin, reduced leptin, elevated levels of cortisol, and growth hormones) (158-161) as well as increased opportunity to eat due to prolonged waking time, a factor that may further become more important with respect to the current “obesogenic environment” (162). Reduced energy expenditure has been suggested as a consequence of restricted sleep duration on thermoregulation as well as resulting from reduced physical activity due to increased fatigue (163-165). Furthermore, it has been assumed that biological effects on body weight going along with restricted sleep depend on the cause of short sleep duration (166). For instance, some individuals feel completely recovered with less than seven hours sleep, while some voluntarily restrict their sleep duration to spend more time on, e.g., leisure time or work, and others wish to sleep longer but are not able to because of **sleep disorders**. By comparing these groups, it is observed that individuals suffering from sleep disorders have hormonal imbalances, including increased secretion of obesity-related stress hormones (e.g., adrenocorticotrophic hormone, cortisol), which do not appear to be present in the other groups (167, 168).

Increasing evidence from prospective studies and randomized intervention trials suggests the importance of **physical activity** in body weight management - in particular, regarding age-related body weight changes (169, 170). Also, some prospective studies focusing on age-related changes in abdominal fat accumulation indicate protective effects related to increased physical activity (171, 172). Multiple underlying biological mechanisms have been proposed, including maintenance of lean body mass and BMR, decrease of body fat mass, as well as better psychosocial health condition (173). The amount of physical activity needed to counteract age-related weight changes is not clear but presumably depends on age, sex, and energy intake (174). However, determined effects of physical activity are generally moderate, while increasing physical activity alone does not seem to be able to completely prevent age-related increases in overall body fatness and abdominal fat accumulation (173).

In addition, there is growing evidence that increased **sedentary behavior**, such as prolonged time watching TV and other screen-based activities, distinctly affects energy balance, independent of physical activity (175). Besides low energy expenditure during time spent sedentarily, typical snacking behavior during sedentary behaviors has been proposed as the major weight gain promoting factor (176, 177).

1.3 Prognostic prediction models

Risk prediction models combine individual characteristics, referred to as predictors, to estimate the absolute risk that an outcome of interest is present (diagnostic prediction model) or will occur within a certain time period (prognostic prediction model) in an individual with a specific predictor profile (178-180). Prognostic prediction models have been developed in the medical field to predict the individual risk of future events and to stratify individuals into risk categories (9, 11, 13). Such models are valuable tools aiming at more beneficial individual outcomes and improved cost effectiveness of care by supporting health care professionals in objective decision-making, e.g., regarding encouraging lifestyle changes as well as initiating and withholding health-care interventions (180, 181). Prediction models are generally characterized by a pre-defined set of predictors (risk factors) and a mathematical function (risk equation), relating the predictor set to the occurrence of the outcome of interest (182).

A model's predictor set is defined by considering the complete range of predictor candidates that have been chosen to be studied. In principle, all factors potentially linked to the outcome of interest could be taken into account, whereby causal relations are not necessarily required (180). As a consequence, predictor candidates can encompass a wide range of factors, such as socio-demographic characteristics (e.g., age, educational level), psychosocial aspects (e.g., social support, life satisfaction), behavioral patterns (e.g., dietary, physical activity and smoking habits), as well as components of biological samples (e.g., blood, saliva, urine, feces) and genetic markers. Based on the complete range of predictor candidates, there are different strategies to select the predictor set of the final model. Thereby, two key strategies are mentioned in the literature, including variations within each strategy: the full model and the selection model (180). Using the full model strategy all a

priori chosen predictor candidates are part of the final model and included in the multivariate analysis. It has been argued that this strategy has the advantage to avoid the so-called predictor selection bias, and, with that, over-adaptation to the models' development sample (178, 183). However, the definition of an appropriate predictor set may be challenging, as prior knowledge about the most promising predictors is necessary, particularly when the number of events is limited and inclusion of too many predictors needs to be avoided (178, 183). Following the selection model strategy, an accessory selection procedure is applied to minimize the set of predictors to the most relevant factors. Thereby, selection procedures using univariate and stepwise regression analysis (i.e., backward and forward selection) based on pre-defined levels of statistical significance are most frequently applied. Selection of predictors based on statistical significance testing, however, has been criticized as being arbitrary (178, 183, 184), and some advanced techniques have been proposed, such as least absolute shrinkage and selection operator or random survival forest (RSF), which are independent of the levels of statistical significance. RSF selects most predictive variables of right-censored survival data by generating random survival trees, based on the random procedures of bootstrapping and node splitting (185). Each tree is grown on the basis of a randomly drawn bootstrap sample from the original data, whereby at each node a subset of variables is randomly selected and the node is splitted by using a survival criterion involving information on survival time and censoring status. To determine most important variables, a measure can be used that identifies those variables that most frequently split the branches near the tree trunks (186), referred to as "minimal depth".

To assign weights to the predictors of the final predictor set of a prognostic model, regression coefficients are frequently used. The regression coefficients of each predictor, estimated by multivariate regression models, such as Cox Proportional Hazard (PH) regression, are mutually adjusted to the other predictors (180). It means that each regression coefficient quantifies the predictors' contribution to the risk estimation of the outcome of interest, when the other predictors in the model are kept constant. Besides the predictor weights, the underlying mathematical function of such models includes the background risk or background survival probability that refers to the estimated risk for a hypothetical individual having all predictors at values of zero. In brief, the individual's

absolute risk of the occurrence of an outcome of interest within a certain time period for a specific risk profile can be calculated by combining the assigned predictor weights and the background risk with the individual's predictor characteristics. The underlying prediction rule can be implemented in practice, as per the prediction algorithm of an easy-to-use risk calculator (risk score) after proper evaluation of the model's performance.

1.3.1 Measures of model performance

The key aspects of evaluating the performance of risk prediction models are mainly covered by three measures:

1. Discrimination: The ability of the model to distinguish between the events and non-events of an outcome of interest
2. Calibration: The agreement of observed risk with predicted risk
3. Validation: The model's validity in a derivation-independent study population

Accordingly, the evaluation of the model performance of a newly developed risk prediction model usually involves an evaluation of the performance measures discrimination and calibration in the derivation sample of the model (internal validity), and, additionally, in order to avoid over-optimism, the evaluation of the measures in a derivation-independent sample with comparable individuals (external validity).

Discrimination

Discrimination generally refers to the ability of a risk prediction model to correctly assign individuals into one of two categories (e.g., events and non-events of an outcome of interest) (187). The most frequently used indicator regarding the assessment of a model's discriminatory ability is the concordance statistic (c statistic) or c index. The c index corresponds to the probability that the parameter of predicted risk - i.e., the risk score - is higher for an event compared with a non-event (188). In doing so, the c index reflects the capability of a prediction model to rank individuals of two categories (189). The c index is equivalent to the area under the receiver operating characteristic (aROC) curve (**Figure 1**).

The aROC is a function of the two basic measures of quantifying discriminatory ability: sensitivity and specificity. Sensitivity, also referred to as true-positive rate (TPR), generally represents the probability of a positive test result or a value above a threshold among individuals with the outcome of interest (events), while specificity, also referred to as true-negative rate (TNR), generally represents the probability of a negative test result or a value below a threshold among individuals without the outcome of interest (non-events) (189). Sensitivity and specificity are inversely proportional. The aROC summarizes all combinations of sensitivity and specificity of the full range of a continuous parameter, such as a risk score by plotting sensitivity (TPR) against 1-specificity (false-positive rate, FPR). The aROC or c index covers the range from 0.5 (reflecting no discriminative ability) when the curve lies along the 45° reference line to a theoretical maximum of 1 (reflecting perfect discriminative ability) and when the curve reaches the upper left corner (189). In case of perfect discrimination, all individuals with events of an outcome of interest have higher individual risk scores than all non-events, with no overlap (189). For c indices between 0.5 and 1, the following categorization has been proposed (191):

Acceptable discrimination: $0.7 \leq \text{aROC} < 0.8$

Excellent discrimination: $0.8 \leq \text{aROC} < 0.9$

Outstanding discrimination: $\text{aROC} \geq 0.9$

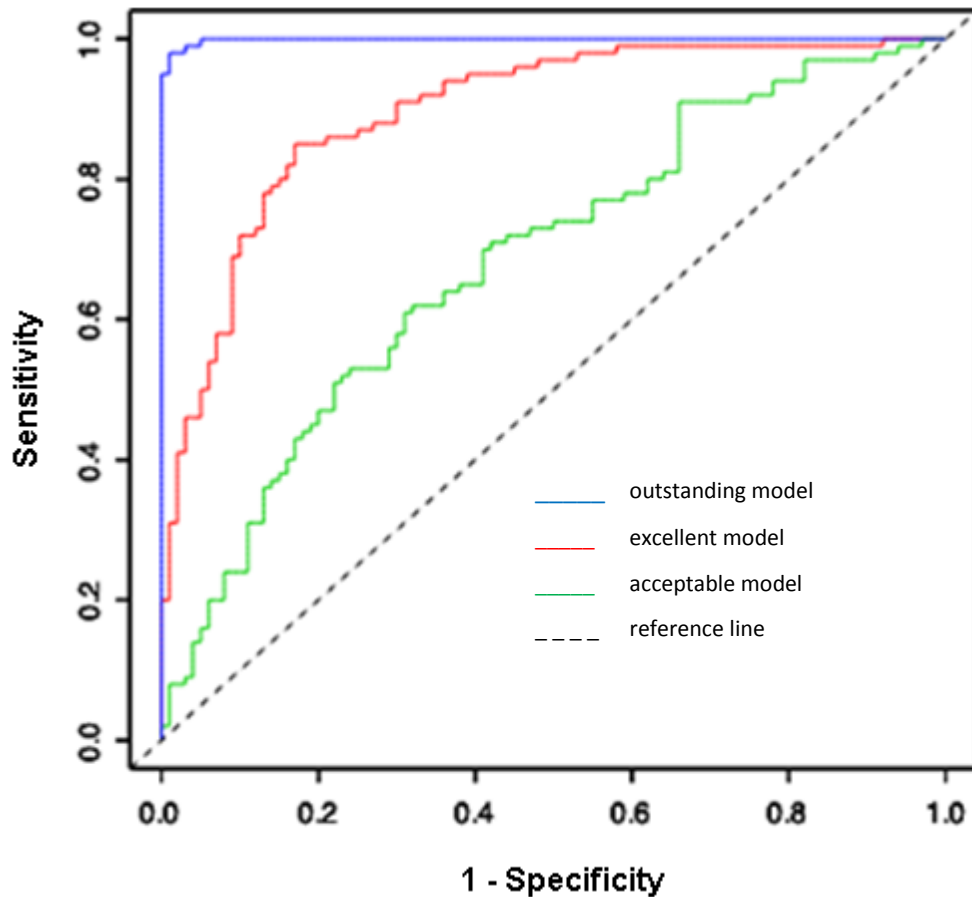


Figure 1. Examples of receiver operating characteristic curves ranging from acceptable to outstanding predictive values (190)

Sensitivity and specificity characterize a model's ability of to distinguish between individuals, which will develop the event of interest and individuals who will not, but they provide no information about the individual's probability to actually develop the outcome of interest. This aspect is covered by the concept of predictive values (192). The positive predictive value (PPV) represents the probability that an individual will develop an outcome of interest, given a positive test result; the negative predictive value (NPV) represents the probability that an individual will not develop the outcome of interest, given a negative test result. Predictive values are influenced by the measures sensitivity and specificity as well as by the prevalence of the outcome of interest in the study population. Given all other factors remain constant, PPV increases with growing prevalence, whereas NPV decreases with growing prevalence (193).

With regard to the implementation of a risk score into practice, an appropriate threshold (cut-off value) of this continuous measure needs to be chosen, which enables one to discriminate between test positive/high-risk individuals and test negative/low-risk individuals. A simple measure to determine such a threshold is Youden's index (J). J maximizes sensitivity and specificity across the range of possible cut-off values (194, 195). It is defined as $J = \text{sensitivity} + \text{specificity} - 1$, and ranges from 0 to 1, with 1 implying perfect separation of events and non-events by the continuous measure (195).

Calibration

Calibration of a prediction model describes the agreement between the predicted risks of developing the outcome of interest by a model and the actually observed incidence (196). To assess this agreement, individuals are frequently divided into categories of predicted risk, which are then compared with the observed incidence in each category, as shown by the Hosmer and Lemeshow goodness of fit test (HLT) (191, 197, 198). The HLT compares the observed risk with the predicted risk across deciles of predicted risk, which is often illustrated by means of calibration plots (191) (**Figure 2**). Points lying above the 45° line represent underestimation of true risk, while points lying below the line represent overestimation of true risk.

It is important to mention that risk prediction cannot be both perfectly calibrated and perfectly discriminatory (200). It has been shown mathematically that a prediction model maximizing discriminatory ability does so at the expense of calibration, while a model maximizing calibration does so at the expense of discriminatory ability (200). As a consequence, the purpose of a risk prediction model is crucial to determine whether discriminatory ability or calibration is of greater relevance. For instance, if a prognostic model is primarily used to identify high-risk individuals to undergo prevention measures, its discriminatory ability would attach greater importance than its accuracy to estimate absolute risks. In contrast, in the context of public health decision-making and the frame of carrying out a cost-benefit analysis a model's accuracy to estimate the actual risk of future events may be of greater interest (9, 189).

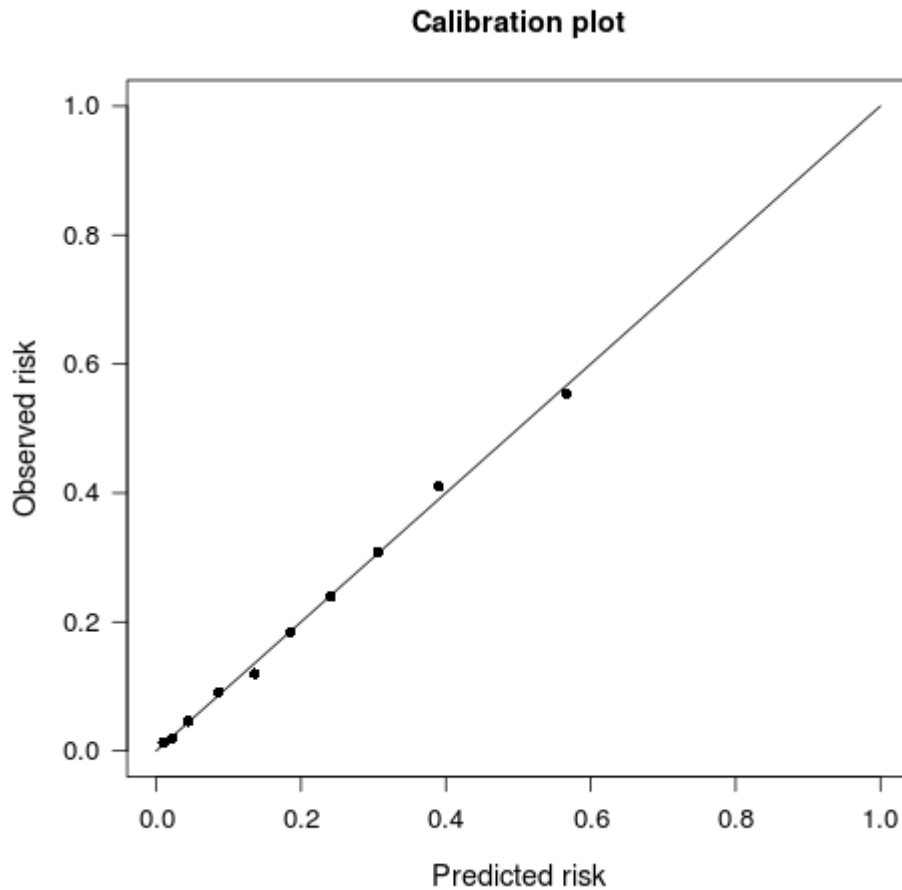


Figure 2. Example of a calibration plot of observed risk against predicted risk across deciles of predicted risk (199)

Validation

An elementary aspect of the evaluation of a model's performance is its validation (178). Thereby, two types of validation are distinguished with reference to the internal validity and external validity of a model. Internal validity describes the assessment of the performance measures discrimination and calibration in the model's derivation sample, which are prone to evaluate the model's performance too optimistically. External validity comprises the assessment of these measures in a derivation-independent, but comparable, study population, which is assumed to provide information on a model's generalizability (180).

1.3.2 Risk scores for gain in weight and waist circumference in adults: state of research

In 2013 the first risk score predicting adult weight gain was published (17). The risk score was developed based on the data of six study centers of the large-scale, multicenter EPIC study ($n_{\text{development}} = 53,748$) and was externally validated based on the data of eight further EPIC centers ($n_{\text{validation}} = 130,446$). The risk score was based on a set of 13 easily obtainable predictors (age, sex, baseline weight, educational level, sports activity, smoking habits, alcohol consumption, and intake of red and processed meat, poultry, fish, bread, and cake and cookies as well as soft drinks) and tailored to predict substantial weight gain (SWG) in the course of the following five years. SWG was defined as gaining $\geq 10\%$ of baseline-based body weight. The score's ability to discriminate between cases and non-cases assessed by the aROC (95 %CI) was 0.64 (0.63, 0.65) in the development sample and 0.57 (0.56, 0.58) in the external validation sample, with variations across the cohorts in each sample. In the development sample discriminatory accuracy ranged from 0.64 (0.62, 0.65) in the Danish cohort to 0.71 (0.68, 0.75) in the Dutch cohort, while in the validation sample the capacity of the score to discriminate between cases and non-cases varied between 0.56 (0.55, 0.57) in the French cohort and 0.67 (0.64, 0.71) in the Italian cohort. As determined by Youden's index, at score values of ≥ 200 , PPV and NPV were 9 % and 96 %, respectively. Calibration of the score was generally well. In the development sample only a minor overestimation of risk in the highest and lowest risk deciles was present. Across cohorts of the development sample calibration was quite homogeneous. In the validation sample minor overestimations of risk existed in the lower and upper deciles, and minor underestimations in the middle range of deciles. Meaningful variations in calibration across cohorts were solely observed in two cohorts of the validation sample.

Recently, the European risk score predicting SWG in the course of five years was adapted using the data of the German part of the EPIC study. In view of varying prevalence of overweight and obesity across European countries as well as diverging environmental and sociocultural circumstances going along with different dietary and lifestyle behaviors (7), it was proposed that the predictive performance of country-specific risk scores could be superior to universal scores for large populations (9). In the course of the adaptation, one of the German EPIC cohorts was used as the adaptation sample (EPIC-Potsdam: $n_{\text{adaptation}} = 12,332$) and the other as the external validation sample (EPIC-Heidelberg: $n_{\text{validation}} =$

10,221). However, the predictive performance of the risk score adapted for Germany-specific characteristics and behaviors remained nearly unchanged. The discriminatory abilities were 0.66 (0.65, 0.68) in the adaptation sample and 0.68 (0.65, 0.70) in the external validation sample, compared with the discriminatory abilities of 0.67 (0.65, 0.69) and 0.68 (0.65, 0.70) based on the reproduced, transnational European risk score. In a further step, the adapted risk score was extended by seven supplemental, easily obtainable weight gain-related predictors. Even though the extension by supplemental predictors could improve the score's ability to discriminate between cases and non-cases in the adaptation sample to aROC (95 %CI) of 0.72 (0.70, 0.73), confirmation of the improvement in the external validation sample was not possible. In the validation sample the discriminatory ability remained nearly on the level of the adaptation: 0.68 (0.66, 0.70) - a finding possibly ascribable to over-adaptation by using one single adaptation cohort. As a consequence, the hypothesis emerged that derivation of a Germany-wide valid risk score showing acceptable levels of predictive performance might be realized by applying a meta-analytical approach comprising diverse German cohort studies, which assessed required predictor information.

With regard to accumulation of abdominal body fat, no risk score predicting absolute risks among adults currently exists.

1.4 Public health relevance of the thesis

With 67.1 % of men and 53.0 % of women, the current prevalence of overweight and obesity in Germany is high (1). According to representative data from the German National Health Interview and Examination Survey 1998 (BGS98) and the German Health Interview and Examination Survey for Adults (DEGS) conducted from 2008 to 2011, the overweight prevalence remained quite constant, whereas the obesity prevalence further increased from 18.9 % to 23.3 % for men and 22.5 % to 23.9 % for women. The increase in obesity prevalence occurred particularly among young adults (1). Excessive body fat accumulation is a well-established risk factor for non-communicable chronic health disorders such as cardiovascular diseases (e.g., heart disease and stroke), type-II-diabetes, degenerative changes of the musculoskeletal system (e.g., osteoarthritis) and certain types of cancer (e.g., endometrial, breast, and colon) (3, 4). A large body of evidence has accumulated over time,

which suggests that the extent of metabolic risk related to excess body fat may be associated with body fat localization (5, 201, 202). In particular, visceral fat accumulation may exert unique pathogenic effects (203-205). The visceral fat compartment acts as an endocrine organ, segregating cytokines and other vasoactive substances that may influence the risk of developing metabolic disorders (202, 205, 206). Owing to its relation to the development of a wide range of diseases, also frequently going along with reduced life expectancy, excess body fat causes enormous health care expenditures. In 2008 3.25 % of total health care expenditures were attributable to the consequences of overweight and obesity in Germany (8). Thereby, the largest proportion (82 %) of direct costs was driven by metabolic health disorders including endocrinological and cardiovascular diseases (8).

Faced with the above-mentioned issues, primary prevention efforts on the subject of overweight and obesity are of great importance relating to public health in Germany. In particular, a preventive tool predicting the risks of substantial gain in weight and waist, which aims at timely identification of high-risk individuals to ensure an early assignment to targeted prevention measures, may counteract the continuing increase in obesity prevalence, and in further consequence, contribute to reduced numbers of future body fat-associated diseases and economic burdens for German society.

1.5 Challenges and research questions of the thesis

Risks scores allow timely, simple, and inexpensive identification of high-risk individuals by taking the multifactorial nature of disease development into account. Concerning abdominal fat accumulation no such tool has been derived so far. Based on the recently published European risk score predicting five-year risk of substantial gain in weight, and the subsequent procedures of risk score adaptation and extension using data of the two German EPIC centers, the aim of the present thesis is to derive and validate German risk scores predicting substantial gain in weight and waist on the basis of diverse German cohort studies. The major challenges are threefold; they comprise the specification of complementary anthropometric measures and meaningful thresholds for substantial gain in overall body fat mass and specific abdominal body fat accumulation, the overcoming of over-adaptation to a pre-defined derivation sample, and the selection of an appropriate,

preferably minimized predictor set leading to the best possible predictive performances in the general German adult population.

In particular, the following research questions are addressed in the frame of the thesis:

1. How well do meta-analytically derived risk scores predicting substantial gain in weight and waist perform across various German cohort studies by using complete available sets of easily obtainable predictors (referred to as maximum models)?
2. How well do meta-analytically derived risk scores predicting substantial gain in weight and waist perform across various German cohort studies by using homogeneous predictor sets (referred to as minimum models)?
3. Is it possible to reduce the homogeneous predictor set by the application of a variable selection procedure without considerable loss of predictive performance of the risk scores predicting substantial gain in weight and waist (referred to as selection models)?

2. Material and methods

2.1 EPI Germany Consortium: the subproject 2

Within the frame of the German Competence Network Obesity (CNO) the EPI Germany Consortium investigates causes and consequences of excessive weight gain, overweight, and obesity throughout the entire life span. The EPI Germany Consortium comprises two subprojects: while subproject 1 focuses its research on young participants from infancy to entry into adulthood, subproject 2 concentrates on participants from young adulthood to older age. The basis for the investigations in subproject 2 is data from seven German cohort studies (**Table 1**). Besides the two German centers of the European Prospective Investigation into Cancer and Nutrition (EPIC) study based in Potsdam and Heidelberg, the Study of Health In Pomerania (SHIP), the research platform KORA (*Kooperative Gesundheitsforschung in der Region Augsburg*), the CARLA (Cardiovascular disease, living and ageing in Halle) study, the PopGen controls as well as the nationwide German National Health Interview and Examination Survey 1998 (BGS98) / National Health Interview and Examination Survey for Adults (DEGS) are part of the subproject. All regional cohort studies were approved by local ethics committees, and all study participants gave their written informed consent. In total, 71,876 men and women were recruited between 1994 and 2007, of which 58,067 were followed-up between 2004 and 2012.

Table 1. Cohorts involved in subproject 2 of the EPI Germany Consortium of the CNO

Study	Region	Recruitment period	Follow-up period	n baseline	n follow-up
EPIC-Potsdam	Potsdam, Brandenburg	1994-1998	2004-2008	27.548	24.569
EPIC-Heidelberg	Heidelberg, Baden-Wuerttemberg	1994-1998	2004-2007	25.540	21.695
SHIP	Greifswald/Stralsund/Anklam, Mecklenburg-Western Pomerania	1997-2001	2008-2012	4.308	3.300
KORA	Augsburg, Bavaria	1999-2001	2006-2008	4.261	3.080
CARLA	Halle, Saxony-Anhalt	2002-2005	2007-2010	1.779	1.436
PopGen	Kiel, Schleswig-Holstein	2005-2007	2010-2011	1.316	942
BGS98/DEGS	national	1997-1999	2008-2011	7.124	3.045

EPIC Potsdam and EPIC Heidelberg

The EPIC study was designed for research on the relations between diet, nutritional status, lifestyle and environmental factors, and the incidence of cancer and other non-communicable diseases (207, 208). The study was established in 1990 within the “Europe Against Cancer” program of the European Commission and is coordinated by the International Agency for Research on Cancer of the World Health Organization (WHO) in Lyon, France. The EPIC study is an ongoing multi-center cohort study involving more than half-a-million study participants allocated to 23 study centers in 10 European countries - Spain, France, the United Kingdom, Norway, Sweden, Denmark, the Netherlands, Italy, Greece, and Germany. The German centers of the EPIC study are instituted at the German Cancer Research Center in Heidelberg and at the German Institute of Human Nutrition in Potsdam. German EPIC study participants were recruited (1994-1998) from the local general population of Heidelberg and Potsdam as well as from the surrounding areas; they were mainly aged between 35 and 65 years (209). They completed questionnaires on socio-demography and lifestyle, a validated food frequency questionnaire (FFQ) and attended an interview on medical history (210, 211). Additionally, the study participants were examined by trained staff regarding blood pressure and anthropometry, and blood samples were taken. In 1997 follow-up data collection of the German EPIC centers had begun and was regularly conducted every two to three years. In the present thesis the decennial follow-up data collection (2004-2008) of the German EPIC participants is included. Active follow-up data collection was conducted by mailed questionnaires and comprised updated information on health conditions, diagnosed diseases and related treatment as well as updated anthropometric and lifestyle data.

SHIP

The SHIP study was initiated in 1997 for a comprehensive investigation of a broad range of health disorders. The study documents common risk factors, pre-clinical changes, and manifest diseases with innovative, non-invasive methods (212). At enrolment (1997-2001), the study participants, belonging to the age group of 20 to 79 years, were drawn by an age- and sex-stratified random sample from residence registries in the cities of Greifswald,

Stralsund, Anklam, and 29 communities in the surrounding areas. Baseline data collection included measurements of blood pressure and anthropometry, neurological screenings, and various ultrasound examinations. Standard operating procedures were established, and examiners were trained and certified according to strict criteria. Moreover, study participants filled out questionnaires and completed a computer-based interview requesting information on socio-demography, working life, health behavior, dietary habits and medical history. Follow-up data collections were conducted in five to six-year intervals by re-invitation of the study participants and included, in addition to the baseline assessments, dermatological examinations, cardiopulmonary exercise tests, examinations of endothelial function, bone stiffness, and whole body magnetic resonance imaging (MRI) scans. The present thesis included updated information on anthropometric measurements of the second follow-up examination (2008-2012).

KORA

KORA is a research platform for population-based surveys, which was established in 1996 to continue the previous WHO-MONICA (Monitoring of Trends and Determinants of Cardiovascular Disease) studies in the region of Augsburg in the southern part of Germany (213). The KORA platform aimed to combine emerging and existing studies in the fields of epidemiology, health economics, and health care research, in particular, with regard to long-term follow-up of the MONICA surveys (S1-S4). Survey participants were drawn from the German population aged 25 to 74 in a two-stage procedure (214): first, Augsburg city and 16 communities from the neighboring counties were selected by cluster sampling; thereafter, stratified random sampling was performed within each community. S1-S4 surveys were performed at five-year intervals. In all surveys baseline information, including socio-demography, lifestyle habits, and medical and family history, was collected by trained medical staff during standardized face-to-face interviews. Moreover, all participants underwent standardized medical examinations including measurements of blood pressure and anthropometry. All surveys were followed up for mortality and onset of myocardial infarction. S3 and S4 surveys were additionally followed for another period of five years by repeated examinations including anthropometric measurements and updated information

on medical history (F3 and F4). The present thesis uses the data of survey S4 (1999-2001) and follow-up survey F4 (2006-2008).

CARLA

The CARLA study was primarily designed to investigate established cardiovascular risk factors and reduced heart rate variability as an indicator of autonomous dysfunction in a representative sample of an elderly East German population (215). The study was launched as population-based cross-sectional study with the aim of a prospective follow-up of examined study participants. CARLA study participants were aged 45 to 80 years; they were randomly drawn from the population registry in Halle city using a multi-step recruitment strategy (216). Baseline data collection (2002-2005) included a standardized, computer-assisted, personal face-to-face interview on socio-demography, behavioral and lifestyle factors, medical and family history, and self-administered questionnaires including a validated FFQ. Detailed medical examinations, including measurements of anthropometry, blood pressure, electrocardiograms, and echocardiography were performed by trained staff. In addition, blood samples were collected. The five-year follow-up examination took place between 2007 and 2010. Follow-up data collection comprised updated information on socio-demography, lifestyle, and medical history as well as repeated measurements of anthropometry.

PopGen controls

The PopGen biobank was initiated in 2002 to investigate genetic risk factors for complex diseases; it comprises a number of population-based studies (217, 218). Besides specific patient groups, a control sample was randomly recruited from the general population. Controls aged 19 to 77 were primarily sampled via population registries and extended by volunteer blood donors. At baseline assessment (2005-2007), participants completed a general questionnaire, including socio-demographic factors and anthropometry, and underwent medical examinations. Additionally, blood samples were taken and ad hoc measurements of a selected range of biological markers and genetic factors were

performed. After five years, all participants were re-invited for a follow-up (2010-2011). Follow-up data collection comprised a general questionnaire, medical examinations, sampling and ad hoc analyses of biological material, and was further extended by the assessment of dietary and lifestyle variables.

BGS98/DEGS

The aim of BGS98 (1997-1999) and DEGS (2008-2011) was to repeatedly obtain nationwide health data for adults aged 18 to 79 years. In BGS98 German residents from 120 cities and communities throughout all federal states were drawn by a two-stage, multi-stratified sampling procedure (219). Re-contactable BGS98 study participants were invited to take part in DEGS (220). The procedures of data collection in BGS98 and DEGS were standardized and harmonized to ensure the comparability of the study results over time. Data collection included interview-assessed information on socio-demography, lifestyle, health behaviors, and medical history. In addition, blood samples were collected and study participants were examined including measurements of anthropometry, heart rate, and blood pressure.

Five cohort studies of the EPI-Germany subproject 2 were used for the derivation and validation of the German risk scores predicting substantial gain in weight and waist: the two German EPIC cohorts, the SHIP study, the research platform KORA and the nationwide BGS98/DEGS cohort. Owing to the lack of a baseline assessment of relevant variables and insufficient numbers of study participants meeting the inclusion criteria, the CARLA study, and the PopGen controls could not be included.

2.1.1 Assessment of relevant variables

Anthropometry

Baseline anthropometry was measured by trained staff according to standardized procedures in all cohort studies. Measurements were obtained without shoes in light clothing with a precision of 0.1 kg for body weight, and a precision of 0.1 cm to 1 cm for body height, and 0.1 cm to 0.5 cm for body circumferences. At follow-up assessment, body weight and body circumferences were measured according to the same cohort-specific standardized procedures, except for participants in the German EPIC cohorts. In the German EPIC cohorts the participants self-reported information on body weight and body circumferences. However, self-reported anthropometry was corrected for potential misreporting using the EPIC-Oxford prediction equations (221). These equations predict sex- and age-specific body measures by linear regression models using data from the study participants with both measured and self-reported anthropometry. The EPIC-Oxford prediction equations are given as follows:

EPIC-Oxford prediction equations for men:

$$\text{Body weight}_{\text{corrected}} = 0.561 + (1.012 * \text{body weight}_{\text{self-reported}}) + (0.006 * \text{age})$$

$$\text{Waist circumference}_{\text{corrected}} = 7.791 + (0.972 * \text{waist circumference}_{\text{self-reported}}) - (0.035 * \text{age})$$

EPIC-Oxford prediction equations for women:

$$\text{Body weight}_{\text{corrected}} = 0.444 + (1.010 * \text{body weight}_{\text{self-reported}}) + (0.006 * \text{age})$$

$$\text{Waist circumference}_{\text{corrected}} = 9.022 + (0.847 * \text{waist circumference}_{\text{self-reported}}) - (0.091 * \text{age})$$

BMI (kg/m²) was calculated as body weight in kilograms (kg) divided by the square of body height in meters (m²).

Dietary factors

In the German EPIC centers the study participants were asked at baseline for their usual dietary intake over the past 12 months using a validated FFQ (208, 222, 223). This FFQ included nine pre-defined categories ranging from never or less than once per month to five or more times per day. If portion-size (e.g., teaspoon, slice) of food items were not uniquely specified, food item-specific pictures were provided. If relevant, information on food preparation was requested as well. Intake of food items in grams per day was additionally calculated by multiplying food frequency and portion-size. The study participants of KORA, SHIP, and BGS98 were only asked for their usual food frequency over the past 12 months without assessment of portion-size by means of a short questionnaire covering the main food groups. In KORA and SHIP, pre-defined categories ranged from (almost) never to (almost) daily. In BGS98, categories ranged from (almost) never to more than once per day.

Socio-demographic, lifestyle, psychosocial, and other health-related factors

All cohort studies assessed detailed information on socio-demographic characteristics (e.g., age and educational level), various lifestyle factors (e.g., alcohol consumption, smoking status, and time spent in physical activity), and history of diagnosed diseases by means of extensive questionnaires, face-to-face interviews and/or computer-assisted interviews. Some of the studies additionally requested information on psychosocial factors (e.g., life satisfaction, living in partnership) and further health-related factors (e.g., suffering from sleep disorders, history of major weight loss). While the request of information on socio-demography and lifestyle was quite consistent across cohorts, some psychosocial and further health-related factors were assessed by more different ways. Life satisfaction was requested using four pre-defined categories ranging from satisfied to unsatisfied in the German EPIC centers, while seven categories ranging from very satisfied to very unsatisfied were used in BGS98 and a short questionnaire covering 12 items related to mental health was used in the SHIP study. Suffering from sleep disorders was medically diagnosed in the German EPIC centers, but such disorders were self-reported in the studies SHIP, KORA, and BGS98 using categories ranging from never to often. In the German EPIC centers, a history of weight loss was requested at baseline as any weight loss of more than 5 kg within the last

two years. In KORA and BGS98, a history of weight loss was requested as previous major weight loss related to the last year (KORA) and any major weight loss which was not further specified regarding mass and time span (BGS98), respectively. **Table 2** shows assessed variables by cohort.

Table 2. Assessment of study-relevant data across the cohorts

Relevant variables	Study population				
	EPIC-Potsdam	EPIC-Heidelberg	SHIP	KORA	BGS98
	27,548	25,540	4,308	4,261	7,124
Socio-demography					
Age	✓	✓	✓	✓	✓
Sex	✓	✓	✓	✓	✓
Educational level	✓	✓	✓	✓	✓
Anthropometry					
Body weight	✓	✓	✓	✓	✓
Body height	✓	✓	✓	✓	✓
Waist circumference	✓	✓	✓	✓	✓
History of weight loss			✗	✓	✓
Lifestyle factors					
Sports	✓	✓	✓	✓	✓
Watching TV	✓	✓	✗	✗	✗
Sleep duration	✓	✗	✓	✗	✓
Smoking habits	✓	✓	✓	✓	✓
Alcohol consumption	✓	✓	✓	✓	✓
Dietary factors					
Meat (red, white and processed)	✓	✓	✓	✓	✓
Fish	✓	✓	✓	✓	✓
Bread	✓	✓	✓	✓	✓
Cake and cookies	✓	✓	✓	✓	✓
Soft drinks	✓	✓	✓	✓	✓
Fruits and vegetables	✓	✓	✓	✓	✓
Chocolate					
Psychosocial Factors					
Life satisfaction	✓	✓	✓	✗	✓
Living in partnership	✓	✓	✓	✓	✓*
Further health-related factors					
Sleep disorders	✓	✓	✓	✓	✓

Assessed variables are marked by check marks (✓). Not assessed variables are marked by crosses (✗). *Living in partnership was assessed in BGS98, but due to high number of missing values ($n_{\text{missings}} = 1,024$) the variable was not considered as predictor candidate in this cohort.

2.1.2 Analytical study population

The study population of the present thesis comprises five population-based German cohort studies involved in subproject 2 of the EPI Germany Consortium; while four cohort studies were conducted in different German regions (Brandenburg, Baden-Wuerttemberg, Mecklenburg-Western Pomerania, Bavaria), one cohort study was conducted nationwide. The data of these five cohorts has been used to derive German risk scores predicting substantial gain in weight and WC as well as for the evaluation of their predictive performance.

The exclusion criteria applied are illustrated in **Figure 3**. The exclusion criteria referred to unavailability of follow-up information, missing anthropometric data, as well as pregnancy at either baseline or follow-up. To avoid the impacts of chronic health disorders on changes in body weight or body fat distribution, the participants with prevalent diabetes, cardiovascular diseases, or cancer at time of the baseline examination have been additionally excluded. Owing to natural changes in body composition at an advanced age, the age of participants has been restricted to <70 years at follow-up. For the purpose of timely identification of high-risk individuals for substantial gain in weight and WC, and consequently to prevent general and/or abdominal obesity, the analytical study population comprises solely non-obese individuals ($BMI < 30 \text{ kg/m}^2$ and $WC_{\text{men}} < 102 \text{ cm}$, $WC_{\text{women}} < 88 \text{ cm}$). Furthermore, study participants with missing values in any exclusion criteria, insufficient information, or missing values in any covariate have been excluded. Since the latter differed between minimum models (inclusively selection models) and maximum models, the final analytical study population includes 32,204 men and women for the minimum and selection models and 31,005 men and women for the maximum models. Divided according to the cohorts, 15,465 study participants have been included from EPIC-Potsdam, 12,229 from EPIC-Heidelberg, 1,318 from SHIP, 1,671 from KORA, and 1,521 from BGS98/DEGS for the minimum and selection models; while 14,633 study participants have been included from EPIC-Potsdam, 11,942 from EPIC-Heidelberg, 1,264 from SHIP, 1,656 from KORA, and 1,510 from BGS98/DEGS for the maximum models.

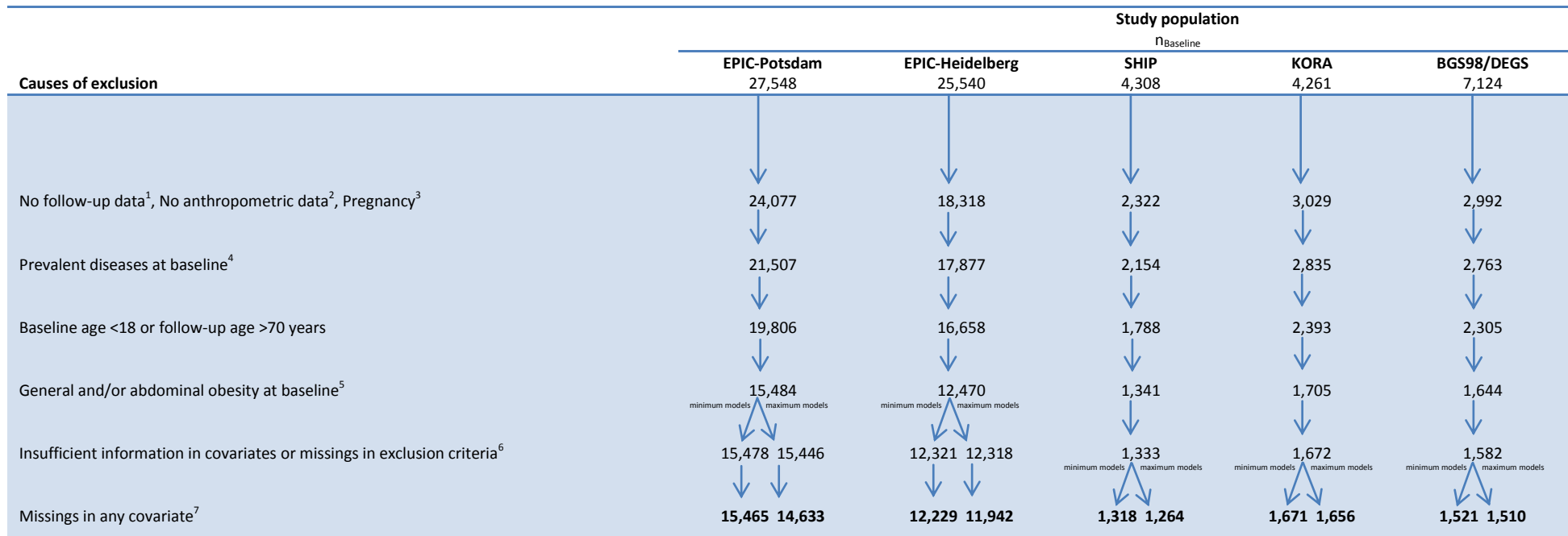


Figure 3. Flow diagram of participants excluded from the present study

¹No follow-up questionnaire (e.g., due to death before follow-up, emigration or non-response to invitation); ²Missing data on baseline or follow-up body weight, waist circumference or body height; ³Pregnancy at baseline or follow-up; ⁴Baseline diabetes, cardiovascular disease or cancer; ⁵Definition of general obesity: BMI ≥ 30 kg/m²; definition of abdominal obesity: WC_{men} ≥ 102 cm, WC_{women} ≥ 88 cm; ⁶Insufficient information (e.g., “don’t know”) in assessment of pregnancy and suffering from sleep disorders in the minimum and selection models (left branch of each cohort) and in addition insufficient information in assessment of life satisfaction, and history of weight loss in the maximum model (right branch of each cohort); ⁷Covariates of the minimum and selection models (left branch of each cohort) and covariates of the maximum models (right branch of each cohort).

2.2 Statistical approach

Statistical analyses have been performed using SAS Enterprise Guide release 4.3 (SAS Institute, Cary, NC). The variable selection is carried out by RSF using R software (version 3.0.1, <http://r-project.org>, package “randomForestSRC”). Meta-analyses are conducted using Review Manager (version 5.3.5, RevMan, Copenhagen: the Nordic Cochrane Center, the Cochrane Collaboration 2014).

2.2.1 Definition of study variables

2.2.1.1 Case status

To derive the German risk scores predicting absolute risks of substantial weight gain (SWG) and substantial waist circumference gain (SWCG) in the course of the following five years and to allow comparisons with the existing European risk score for SWG, complementary anthropometric measures and meaningful thresholds needed to be chosen. SWG has been specified as gaining $\geq 10\%$ of baseline weight within the following five years in accordance with the European risk score. Owing to the high correlation of anthropometric measures of overall body fatness and abdominal body fat accumulation (25), SWCG is defined by the residuals of WC regressed on BMI (WC_{BMI}). WC_{BMI} represents those changes of WC, which are independent of the changes in BMI and thus specifically reflect abdominal fat accumulation. SWCG is specified as gaining ≥ 2.5 cm of baseline WC_{BMI} within the following five years. The relative scale for the assessment of SWG is applied to ensure a better comparability between individuals with different initial body weights, while the absolute scale for the assessment of SWCG is applied to capture disproportional gains in WC across the whole range of overall fat mass. The thresholds of 10 % gain in body weight and 2.5 cm gain in residual waist circumference have been chosen for two reasons: on the one hand the thresholds were considered to reflect major gain in weight and specific gain in waist, possibly accompanied by adverse health effects; on the other hand the thresholds appeared sufficiently high to exclude random fluctuation in body weight and WC, but still allowing for some gain going along with the natural process of aging.

To take account of individually varying follow-up time and celerity of gain in anthropometric measures, a survival analysis is applied for statistical analysis. Thereby, the study participants are observed for incidence of SWG and SWCG from their baseline assessment of anthropometry to their anthropometric assessment at follow-up. The study participants gaining $\geq 10\%$ of their baseline weight and/or ≥ 2.5 cm of their baseline WC_{BMI} during five years of follow-up have been assigned to the corresponding set of cases, while the study participants not experiencing SWG and/or SWCG have been assigned to the corresponding set of non-cases. To estimate the theoretical point in time when the thresholds of 10 % (for SWG) and 2.5 cm (for SWCG) are passed (also referred to as survival time or time-to-event), the observed follow-up time is modified using average annual proportions of baseline-based changes in body weight and WC_{BMI} under the assumption of linearity. The calculation equations for the theoretical point in time of SWG (1) and SWCG (2) are given as follows:

(1)

Percentage body weight change:

$$\text{Body weight change [\%]} = (\text{body weight}_{\text{follow-up}} - \text{body weight}_{\text{baseline}}) / \text{body weight}_{\text{baseline}}$$

↓

Percentage average annual body weight change:

$$\text{Annual body weight change [\%]} = (\text{body weight}_{\text{follow-up}} - \text{body weight}_{\text{baseline}}) / (\text{body weight}_{\text{baseline}} * \text{follow-up time})$$

↓

Calculation of the theoretical point in time of SWG ($\geq 10\%$):

$$\text{Theoretical point in time} = 0.1 / \text{annual body weight change [\%]}$$

(2)

Absolute residual waist circumference (WC_{BMI}) change:

$$WC_{BMI} \text{ change [cm]} = (WC_{BMI \text{ follow-up}} - WC_{BMI \text{ baseline}})$$

↓

Absolute annual WC_{BMI} change:

$$\text{Annual } WC_{BMI} \text{ change [cm]} = (WC_{BMI \text{ follow-up}} - WC_{BMI \text{ baseline}}) / \text{follow-up time}$$

↓

Calculation of the theoretical point in time of SWCG (≥ 2.5 cm):

$$\text{Theoretical time point} = 2.5 / \text{annual } WC_{BMI} \text{ change [cm]}$$

2.2.1.1 Predictor variables

Predictor variables for SWG and SWCG are primarily based on reported and hypothesized associations with changes in body weight and/or WC in the literature (as addressed in Chapter 1.2.2). Predictor variables exclusively cover baseline data. Additionally, a feasible assessment of relevant predictors has been taken into account, in particular, with regard to the practical applicability of the derived risk scores. Therefore, dietary factors are represented by main food groups instead of individual nutrients. If necessary, scales are harmonized between cohorts. All relevant dietary factors (fruits and vegetables, meat, whole grain bread, cake and cookies, chocolate, soft drinks, fish) have been classified into three food group-specific consumption categories (see **Table 3**). In EPIC-Potsdam, SHIP, KORA, and BGS98 information on food frequencies were directly available, whereas in EPIC-Heidelberg data on dietary factors was solely available in grams per day. Based on the specified consumption categories and the food group-specific portion sizes (i.e., 125 g for fruits and vegetables, 80 g for whole grain bread, 125 g for fish, 125 g for meat, 135 g for cake and cookies, 15 g for chocolate, 250 g for soft drinks), consumption categories were approximated for EPIC-Heidelberg. Since intake of whole grain bread was not requested in EPIC-Heidelberg and SHIP it was replaced by the assessed intake of non-white bread. Educational level, sporting activities and life satisfaction have also been classified into three categories; alcohol consumption included four categories (see **Table 3**). The remaining predictor variables are dichotomous (e.g., smoking status, suffering from sleep disorders, living in partnership) or are treated continuously (e.g., sleep duration, time spent watching TV). The maximum predictor set includes 22 variables, comprising socio-demographic, anthropometric, lifestyle, dietary, psychosocial and further health-related factors. Depending on the cohort 19 (KORA) to 22 (EPIC-Potsdam) predictor variables are available for the maximum models. 17 of these predictors are adequately assessed in all cohorts, representing the homogenous predictor set of the minimum models. To identify most predictive variables for SWG and SWCG, a variable selection procedure for right-censored survival data (RSF) has been performed on the basis of the homogenous predictor set. The selection procedure (generating 1000 trees and using 1 split point) is applied separated by outcome and cohort. Most predictive variables across the cohorts (ranked by “minimal depth”) have been chosen for the selection models.

Table 3. Predictor variables and their corresponding scales

Predictor variables	Scales
Socio-demographic and anthropometric factors	
Age	continuous (per year)
Sex	dichotomous (female vs. male ^{REF})
Education	3 categories: no/primary school ^{REF} , secondary/professional school, university
Body weight	continuous (per kg)
Waist circumference	continuous (per cm)
Body height	continuous (per cm)
History of weight loss*	dichotomous (yes vs. no ^{REF})
Lifestyle factors	
Sports	3 categories: 0 h/week ^{REF} , >0 to <2 h/week, ≥2 h/week
Alcohol consumption	4 categories: 0 g/day, >0 to <6 g/day ^{REF} , ≥6 to <18 g/day, ≥18 g/day
Smoking habits	dichotomous (current non-smoking ^{REF} vs. current smoking)
Watching TV*	continuous (per h/day)
Sleep duration*	continuous (per h/day)
Dietary factors¹	
Fruits and vegetables	3 categories: <1.5 portions/day ^{REF} , ≥1.5 portions to <2 portions/day, ≥2 portions/day
Meat	3 categories: <4-6 portions/week ^{REF} , ≥4-6 portions/week to <1 portion/day, ≥1 portion/day
Whole grain bread ²	3 categories: <2-3 portions/week ^{REF} , ≥2-3 portions/week to <4-6 portions/week, ≥4-6 portions/week
Cake and cookies	3 categories: <1 portion/week ^{REF} , ≥1 portion/week to <2-3 portions/week, ≥2-3 portions/week
Soft drinks	3 categories: <2-3 portions/month ^{REF} , ≥2-3 portions/month to 4-6 portions/week, ≥4-6 portions/week
Fish	3 categories: <2-3 portions/month ^{REF} , ≥2-3 portions/month to <1 portion/week, ≥1 portion/week
Chocolate	3 categories: <1 portion/week ^{REF} , ≥1 portion/week to <2-3 portions/week, ≥2-3 portions/week
Psychosocial factors	
Living in partnership*	dichotomous (yes vs. no ^{REF})
Life satisfaction*	3 categories: satisfied ^{REF} , rather (un-)satisfied, unsatisfied
Further health-related factors	
Sleep disorders	dichotomous (yes vs. no ^{REF})

¹Dietary factors were used as consumption frequency/day in EPIC-Potsdam, SHIP, KORA, and BGS98 and as usual portion size/day in EPIC-Heidelberg (portion sizes by food group: 125 g fruits and vegetables; 80 g whole grain bread, 125 g fish, 125 g meat, 135 g cake and cookies, 15 g chocolate, 250 g soft drinks). ²In EPIC-Heidelberg and SHIP non-white bread was used instead of whole grain bread. REF = reference. Predictor variables only included in the maximum models are marked by asterisk (*).

2.2.2 Descriptive statistics

General characteristics are calculated across the cohorts. Continuous variables are presented as arithmetic means and standard deviations, except from WC_{BMI} which is presented by percentiles. Categorical variables are presented as proportions. Additionally, incidences of SWG and SWCG, including incidence rates (per 10,000 person-years, PY), have been computed for each cohort.

2.2.3 Construction and evaluation of the German risk scores

Risk functions for detecting incident SWG and SWCG are generated using Cox PH regression models (224). In Cox PH regression models the time period until the event occurs is accounted for, which is a clear advantage over logistic regression models that solely consider whether an event occurs or not. Thus, individual varying follow-up times are taken into consideration, whereas extrapolations of observed changes in body weight and WC_{BMI} to a specified non-observed time point were not required.

To assign score values (weights) for each predictor, regression coefficients (β) are used. To avoid over-adaptation for one specific cohort, the score values are meta-analytically derived on the basis of all five cohorts. In consideration of the inconsistencies between the cohorts, for example due to differences in assessment methods and procedures, questionnaire designs, and other cohort effects, cohort-specific regression coefficients were initially obtained separately for all predictors and are subsequently meta-analytically combined using random-effects meta-analyses. Random-effects meta-analyses weight the cohort-specific effects by the inverse of their variance (225) and allow for variation of the effect sizes between cohorts, which is a clear advantage over fixed-effects meta-analyses (225, 226). In the minimum and selection models all cohorts contribute to the meta-analytically estimated final score values, while in the maximum models only cohorts that assessed the corresponding predictors could be used. To quantify the amount of inconsistency between cohorts χ^2 -statistic, the corresponding p-values and resulting I^2 are reported. I^2 describes the percentage of variation in effect estimates across cohorts, which is due to heterogeneity rather than sampling error (chance) (227, 228). To illustrate the results of random-effects meta-analyses, forest plots are generated using RevMan. Based on the meta-analytically

combined regression coefficients - rounded to two decimal places and multiplied by 100 - the final score values are specified. Based on these score values for each individual, a risk score (RS_i) is calculated as a linear combination of individually weighted predictors. Re-scaling by addition of a constant is performed to avoid negative score values.

Calculation of individual risk scores:

$$RS_i = \text{constant} + \sum_{i=1}^m (100 * \beta_i) * x_i$$

x = predictor

β = regression coefficient

m = number of predictors

constant = 800 (SWG) and 2,500 (SWCG)

The estimation of the absolute risks of gaining ≥ 10 % of body weight and ≥ 2.5 cm of WC_{BMI} in the course of the following five years is based on the survival function of the Cox PH regression model (229).

Survival function of the Cox PH regression model:

$$S(t, X) = S_0(t) \exp \left(\sum_{i=1}^m (\beta_i * x_i) \right)$$

t = time point

X = set of predictors: $x_1 \dots x_m$

Corresponding to the survival function of the Cox PH regression model, the estimated individual survival probability (S) at a specified point in time (t) for a defined set of predictors (X) is given by a background survival function ($S_0(t)$) raised to a power equal to the exponential of the sum of each individually weighted predictor ($\sum_{i=1}^m (\beta_i * x_i)$). According to the definition, the background survival probability represents the survival probability at a specified point in time when all predictors take values equal to zero. However, to allow meta-analytically combinability, the estimation of background survival probability for SWG was already modified in the course of the development of the European risk score (17).

Thereby, mean values of each predictor (x_{mean}) are used instead of zero values. Following this modification, the estimation of background survival probability is performed and cohort-specific background survival probabilities during the following five years ($S_M(5y)$) are meta-analytically combined.

Background survival function at five years:
$$S_M(5y) = S(5y, X_{\text{mean}}) = S_0(t) \exp\left(\sum_{i=1}^m (\beta_i * x_{\text{mean}})\right)$$

The individual probability of gaining $\geq 10\%$ of body weight and ≥ 2.5 cm of WC_{BMI} in the course of the following five years ($P(5y)_i$) are ultimately computed by the insertion of the individual risk score (RS_i) into the survival function and subtraction from 1. In order to account for the modified background survival probability, the linear combination in the survival function formula is corrected for the averages of the participants' predictors (RS_M).

$$P(5y)_i = 1 - S_M(5y)^{\exp((RS_i - RS_M)/100)}$$

with

$$RS_M = 500 + \sum_{i=1}^m (100 * \beta_i) * x_{\text{mean } i}$$

Predictive performances of the derived risk scores are evaluated by means of discrimination and calibration using SAS macros. The discrimination is assessed by the c index and the corresponding 95% confidence intervals (CI), specifically adapted for time-to-event analyses, based on Harrel (184) and extended by Pencina and D'Agostino (187). The discrimination is illustrated by ROC curves. Calibration is tested by a modified version of the Hosmer-Lemeshow test (HLT) for time-to-event analyses provided by D'Agostino and Nam (197), and visualized by plotting the observed versus predicted risk across the deciles of predicted risk. To specify appropriate cut-off values of the risk scores, Youden's index is applied.

3. Results

3.1 Description of the study population

3.1.1 General characteristics

The general characteristics of the study population, separated by cohorts, are presented in **Table 4**. To get an overview of the whole range of the considered predictor variables, the table shows such general characteristics for the maximum models.

A total of 31,005 individuals were included for derivation and validation of the maximum models for SWG and SWCG. By far the largest contribution in terms of participants came from the German EPIC centers (85.7 %), whereas the studies SHIP (4.1 %), KORA (5.3 %) and BGS98/DEGS (4.9 %) contributed in approximately equal parts. The mean follow-up duration clearly exceeds the case definition-relevant limit of five years, ranging from 7.2 years (KORA) to 11.9 years (BGS98/DEGS). Participants from SHIP and BGS98/DEGS were the youngest, with a mean baseline age of slightly less than 40 years. The German EPIC centers included the oldest participants, with 47.3 years in EPIC-Potsdam and 48.6 years in EPIC-Heidelberg. The proportion of men ranged from 36.7 % (EPIC-Potsdam) to 49.5 % (KORA). Across the cohorts most of the individuals completed a secondary or professional school education (41.6 % in EPIC-Heidelberg to 75.0 % in KORA) or attained a university degree (13.8 % in SHIP to 41.6 % in EPIC-Potsdam). During follow-up the participants on average gained 3.1 % (KORA) to 6.3 % (SHIP) of their baseline body weight as well as 3.7 % (KORA) to 11.8 % (EPIC-Potsdam) of their baseline WC. On average, the participants annually gained 299 g (KORA) to 429 g (EPIC-Potsdam) in body weight and 4 mm (KORA and BGS98/DEGS) to 11 mm (EPIC-Potsdam) in WC. Regarding gain in WC_{BMI} , percentiles were used as descriptive measures. Annual changes in WC_{BMI} ranged from 2 mm (SHIP and BGS98/DEGS) to 4 mm (EPIC-Heidelberg) for the 75th percentile, and from 4 mm (SHIP and BGS98/DEGS) to 8 mm (EPIC-Heidelberg) for the 90th percentile. The overall proportion of generally obese individuals at follow-up ranged from 5.8 % (EPIC-Heidelberg) to 13.8 % (SHIP), while the proportion of abdominally obese individuals varied between 18.4 % (SHIP) and 32.6 % (EPIC-Potsdam). For cohorts assessing corresponding information, the proportion of individuals who reported major weight loss ranged from 1.2 % (BGS98/DEGS) to 12.5 % (EPIC-Potsdam).

With regard to lifestyle factors, the proportion of the most active individuals (≥ 2 h sports/week) was highest in EPIC-Heidelberg (41.7 %) and quite similar among other cohorts (23.9 % to 25.0 %). The proportion of individuals practicing no sporting activities was highest in EPIC-Potsdam (56.0 %), but it did not fall below 25 % for other cohorts. Across the cohorts more than two-thirds were non-smokers (66.2 % in SHIP and BGS98/DEGS to 79.1 % EPIC-Potsdam). The proportion of alcohol abstainers was highest in SHIP (28.3 %), followed by KORA (23.3 %) and BGS98/DEGS (12.5%); the proportion of low-alcohol consumers ranged between 22.5 % (SHIP) and 47.7 % (BGS98/DEGS), while the proportion of individuals with highest alcohol consumption was above one-third in EPIC-Heidelberg and KORA. Average sleep duration and daily hours spent watching TV were quite similar across the cohorts assessing that information. Concerning dietary factors, the proportions of individuals with most frequent consumption of fruits and vegetables (68.7 %) and fish (57.9 %) but also with most frequent consumption of meat (75.6 %), cake and cookies (51.2 %), and chocolate (53.6 %) were highest in EPIC-Potsdam; the proportion of individuals with most frequent consumption of fruits and vegetables was markedly lower (17.6 % to 39.1 %) in other cohorts, while the proportion of individuals with most frequent intake of fish, cake and cookies as well as chocolate differed to a lesser extent. The proportion of individuals with most frequent intake of soft drinks (21.3 %) was highest in BGS98/DEGS but lowest in the German EPIC-centers (6.2 % in EPIC-Potsdam and 9.8 % in EPIC-Heidelberg). Regarding psychosocial and further health-related factors, more than three-fourth of individuals were living in partnership (77.2 % in KORA to 86.4 % in German EPIC centers). In three out of four cohorts most individuals stated that they were rather satisfied with their lives, while the proportion of individuals that were satisfied with their lives exceeded proportion of (rather) unsatisfied individuals (16.5 % in BGS98/DEGS to 64.6 % in SHIP vs. 3.3 % in SHIP to 14.0 % in BGS98/DEGS) across the four cohorts assessing corresponding information. The proportions of individuals suffering from sleep disorders were quite similar across the cohorts, ranging from 11.9 % in EPIC-Heidelberg to 17.2 % in KORA.

Owing to slight differences between the study population for the maximum models and the study population for the minimum as well as selection models, the general characteristics of the study population for the minimum and selection models using homogeneous predictor sets across cohorts can be found in the Appendix (**App. Table 1**).

Table 4. General characteristics of the study population

	Study population				
	EPIC-Potsdam	EPIC-Heidelberg	SHIP	KORA	BGS98/DEGS
N	14,633	11,942	1,264	1,656	1,510
Duration of follow-up (y)	8.5 (0.9)	8.5 (0.7)	11.2 (0.8)	7.2 (0.3)	11.9 (0.9)
Socio-demography					
Age at baseline (y)	47.3 (8.1)	48.6 (7.4)	39.5 (10.8)	43.0 (10.5)	39.2 (10.6)
Men (%)	36.7	44.6	45.0	49.5	47.7
Education (%)					
No/primary school	12.8	21.8	12.8	5.0	6.8
Secondary/professional school	45.6	41.6	73.4	75.0	73.7
University	41.6	36.5	13.8	20.1	19.5
Anthropometry					
Weight					
At baseline (kg)	68.6 (10.8)	69.7 (11.1)	71.5 (12.0)	71.1 (11.5)	70.7 (11.6)
At follow-up (kg)	72.3 (11.8)	72.7 (12.5)	75.8 (13.1)	73.2 (12.5)	74.4 (13.3)
Absolute change (kg)	3.6 (5.0)	3.0 (6.0)	4.3 (6.3)	2.1 (4.7)	3.7 (6.4)
Annual change (g/year)	429 (590)	357 (719)	382 (561)	299 (661)	311 (537)
Change (% of baseline weight)	5.5 (7.5)	4.5 (9.1)	6.3 (9.1)	3.1 (6.7)	5.4 (9.2)
Height					
At baseline (cm)	167.8 (8.5)	169.7 (8.6)	170.7 (8.9)	169.7 (9.2)	170.8 (9.1)
BMI					
At baseline (kg/m ²)	24.3 (2.6)	24.1 (2.7)	24.4 (2.8)	24.6 (2.7)	24.1 (2.7)
At follow-up (kg/m ²)	25.6 (3.1)	25.2 (3.3)	26.1 (3.4)	25.1 (3.0)	25.5 (3.4)
Obese at follow-up (%)	7.5	5.8	13.8	5.9	9.2
Waist					
At baseline (cm)	80.5 (9.8)	82.2 (10.1)	80.6 (10.2)	84.0 (9.7)	82.8 (10.0)
At follow-up (cm)	89.8 (10.4)	90.8 (12.2)	86.2 (11.0)	87.0 (10.7)	87.5 (11.2)
Absolute change (cm)	9.2 (5.6)	8.6 (8.7)	5.6 (6.5)	3.0 (5.6)	4.7 (7.0)
Annual change (mm/year)	11 (7)	10 (10)	5 (6)	4 (8)	4 (6)
Change (% of baseline waist)	11.8 (7.4)	10.8 (10.9)	7.3 (8.4)	3.7 (6.8)	6.0 (8.9)
Abdominal obese at follow-up (%)	32.6	32.1	18.4	18.8	22.1
Residual waist					
Absolute change (cm)					
75 th percentile	2.6	3.4	2.4	2.2	2.4
90 th percentile	5.2	7.2	4.9	4.4	5.2
Annual change (mm/year)					
75 th percentile	3	4	2	3	2
90 th percentile	6	8	4	6	4
History of major weight loss (%)	12.5	5.2	not assessed	2.7	1.2
Lifestyle factors					
Sports					
0 h/week	56.0	37.4	41.1	25.4	33.6
>0 to <2 h/week	19.0	21.0	35.1	49.9	41.6
≥2 h/week	25.0	41.7	23.9	24.7	24.8
Smoking habits (%)					
Current non-smoking	79.1	77.8	66.2	71.0	66.2
Current smoking	20.9	22.2	33.8	29.1	33.8
Alcohol consumption					
0 g/day	2.4	4.2	28.3	23.3	12.5
>0 to <6 g/day	37.6	31.7	22.5	23.1	47.7
≥6 to <18 g/day	33.0	29.4	22.2	19.8	22.2
≥18 g/day	27.0	34.7	27.0	33.9	17.6
Sleep duration (h/day)	7.2 (1.0)	not assessed	7.3 (1.1)	not assessed	7.9 (1.0)
Watching TV (h/day)	1.8 (1.0)	1.6 (2.9)	not assessed	not assessed	not assessed
Dietary factors					
Fruits and vegetables					
<1.5 portion/day	13.7	34.4	56.7	52.3	51.3
≥1.5 portion to <2 portions/day	17.5	26.5	25.7	23.6	19.1
≥2 portions/day	68.8	39.1	17.6	24.2	29.7
Meat					

Results

<4-6 portions/week	12.3	58.3	21.1	39.1	27.9
≥4-6 portions/week to <1 portion/day	12.1	20.7	19.2	27.3	31.5
≥1 portion/day	75.6	21.0	59.7	33.6	40.7
Whole grain bread¹					
<2-3 portions/week	48.9	9.3	6.3	30.0	33.4
≥2-3 portions/week to <4-6 portions/week	17.1	15.3	7.5	33.1	25.9
≥4-6 portions/week	34.1	75.4	86.2	36.9	40.7
Cake and cookies					
<1 portion/week	15.6	19.2	32.0	31.9	21.2
≥1 portion/week to <2-3 portions/week	33.2	31.0	29.6	30.0	33.6
≥2-3 portions/week	51.2	49.8	38.5	38.0	45.2
Soft drinks					
<2-3 portions/month	70.8	69.9	56.5	56.3	40.6
≥2-3 portions/month to 4-6 portions/week	23.0	20.3	28.6	30.9	38.2
≥4-6 portions/week	6.2	9.8	15.0	12.8	21.3
Fish					
<2-3 portions/month	18.0	38.6	35.1	33.0	22.6
≥2-3 portions /month to <1 portion/week	24.1	23.3	29.8	28.0	25.1
≥1 portion/week	57.9	38.1	35.2	39.0	52.3
Chocolate					
<1 portion/week	22.4	29.5	42.1	37.6	27.3
≥1 portion/week to <2-3 portions/week	24.0	19.6	17.5	22.3	24.0
≥2-3 portions/week	53.6	50.9	40.4	40.2	48.7
Psychosocial factors					
Living in partnership (%)	86.4	86.4	79.9	77.2	39.9 ²
Life satisfaction (%)					
Satisfied	24.4	40.8	64.6	not assessed	16.5
Rather satisfied	64.5	53.1	32.0	not assessed	69.5
(Rather) unsatisfied	11.1	6.2	3.3	not assessed	14.0
Further health-related factors					
Sleep disorders (%)	13.2	11.9	16.5	17.2	16.1

¹ In EPIC-Heidelberg and SHIP intake of non-white bread was used instead of whole grain bread. ² In BGS98/DEGS living in partnership had 1,024 missing values.

3.1.2 Incidences of SWG and SWCG

In the course of the overall follow-up period, 6,383 individuals gained ≥10 % of their baseline weight (**Table 5**). Divided according to the cohorts, the number of individuals experiencing ≥10 % weight gain ranged from 226 (KORA) to 3,420 (EPIC-Potsdam). Regarding abdominal fat accumulation, 8,746 individuals increased their WC_{BMI} to ≥2.5 cm. Separated by cohorts, the number varied between 309 (SHIP) and 3,965 (EPIC-Potsdam). In the representative BGS98/DEGS cohort, approximately a quarter of individuals experienced a weight gain of ≥10 %, which corresponded quite exactly to the proportion of individuals experiencing the complementary outcome of gaining ≥2.5 cm of WC_{BMI}. In total 1,661 individuals experienced SWG during the first five years of the follow-up. Separated by cohorts, the number ranged from 66 (SHIP) to 935 (EPIC-Potsdam). Concerning abdominal body fat, a total of 5,073 individuals experienced SWCG within the first five years. In accordance with the cohorts, the number of individuals experiencing SWCG varied between 91 (SHIP) and 2,496 (EPIC-Potsdam). Across all cohorts the proportion of individuals that experienced SWCG within the

first five years (6.9 % to 20.4 %) was generally higher than proportion of individuals experiencing SWG (4.1 % to 6.1 %).

Table 5. Follow-up times, cases of SWG and SWCG during total follow-up and within the first five years

	Study population				
	EPIC-Potsdam	EPIC-Heidelberg	SHIP	KORA	BGS98/DEGS
Follow-up time	132,442	104,626	14,740	11,956	18,108
SWG					
Cases during total follow-up (%)	3,420 (22.1)	1,973 (16.1)	392 (29.7)	226 (13.5)	372 (24.5)
Cases during first five years of follow-up (%)	935 (6.1)	504 (4.1)	66 (5.0)	89 (5.3)	67 (4.4)
SWCG					
Cases during total follow-up (%)	3,965 (25.6)	3,729 (30.5)	309 (23.4)	370 (22.1)	373 (24.5)
Cases during first five years of follow-up (%)	2,143 (13.9)	2,496 (20.4)	91 (6.9)	229 (13.7)	114 (7.5)

Follow-up times are presented in person-years (PY). Cases are presented in absolute numbers including proportions in parentheses.

To facilitate comparisons of SWG and SWCG incidences by cohorts, incidence rates (per 10,000 person-years, PY) are additionally illustrated in **Figure 4**. Incidence rates for SWG ranged across the cohorts from 189 (EPIC-Heidelberg and KORA) to 266 (SHIP), whereas incidence rates for SWCG varied from 206 (BGS98/DEGS) to 356 (EPIC-Heidelberg). In the nationwide BGS98/DEGS sample both incidence rates for SWG and SWCG were at approximately equal levels of slightly above 200. Owing to the slightly reduced study population follow-up times, the number of cases, and incidence rates of the study population for the maximum models can be found in the Appendix (**App. Table 2** and **App. Figure 1**).

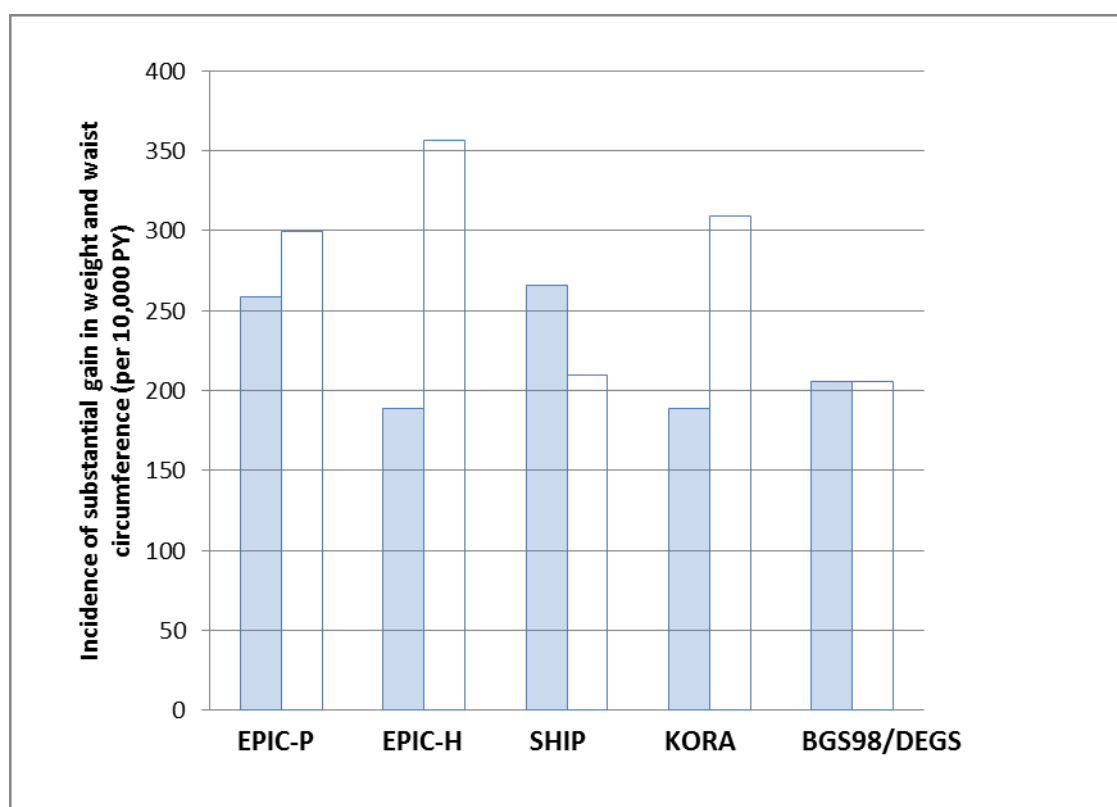


Figure 4. Incidence rates of SWG and SWCG (per 10,000 person-years)

Incidence rate = (Cases during total follow-up/person-years) * 10,000; Incidence rates of SWG are illustrated by blue bars and incidence rates of SWCG are illustrated by white bars.

3.2 Derivation of the German risk scores

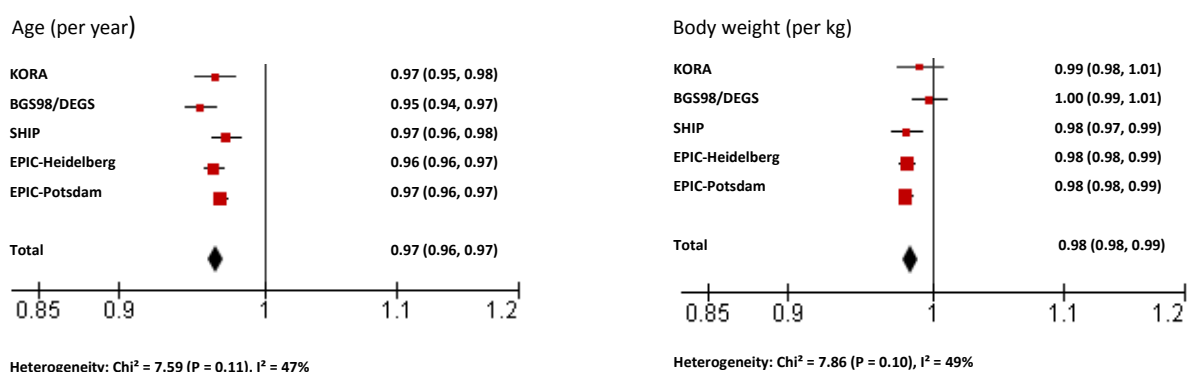
3.2.1 Description of the German risk scores

The maximum models for SWG have been defined on the basis of the following 21 baseline predictor variables: age, sex, educational level, body weight, body height, sports, alcohol consumption, smoking, suffering from sleep disorders, intake of fruits and vegetables, meat, whole grain bread, cake and cookies, chocolate, soft drinks and fish, weight loss history, watching TV, sleep duration, living in partnership, and life satisfaction. For the maximum models of SWCG, baseline WC was further added to the maximum predictor set, which amounted to a total of 22 predictor variables. The predictor variables of the minimum models for SWG had been reduced to the following 16 factors, which were available in all cohort studies: age, sex, educational level, body weight, body height, sports, alcohol consumption, smoking, suffering from sleep disorders, intake of fruits and vegetables, meat, whole grain bread, cake and cookies, chocolate, soft drinks, and fish. For the minimum

models of SWCG, baseline WC was further added to the minimum predictor set. The predictor variables of the minimum model for SWG correspond quite well to the predictor set of the European risk score (17). After applying the variable selection procedure of RSF to the homogenous predictor set of the minimum models, separated by cohorts and outcomes, those variables which were most predictive according to “minimal depth” across all cohorts were included in the selection models. Considering the numbers of cases for reliable variable selection (≥ 10 cases per variable) across the cohorts, the selection models for SWG included five predictor variables (age, sex, body weight, smoking, and consumption of soft drinks), and the selection model for SWCG included seven predictor variables (age, sex, body weight, body height, WC, educational level, and meat intake).

Maximum, minimum, and selection models were run in each of the five cohorts, and the combined effects across the cohorts were estimated using random-effects meta-analyses. As observed by the range of cohort-specific hazard ratios (HR) in the minimum, maximum and selection models, heterogeneity has been found between the predictors and the cohorts. However, confidence intervals did overlap for most predictors, but not all. The forest plots for the relationship between age and body weight with the risk for SWG and SWCG in the selection models are demonstrated by way of example in **Figure 5**. The remaining plots of the selection models can be found in the Appendix (**App. Figure 2**).

a, Associations with risk for SWG



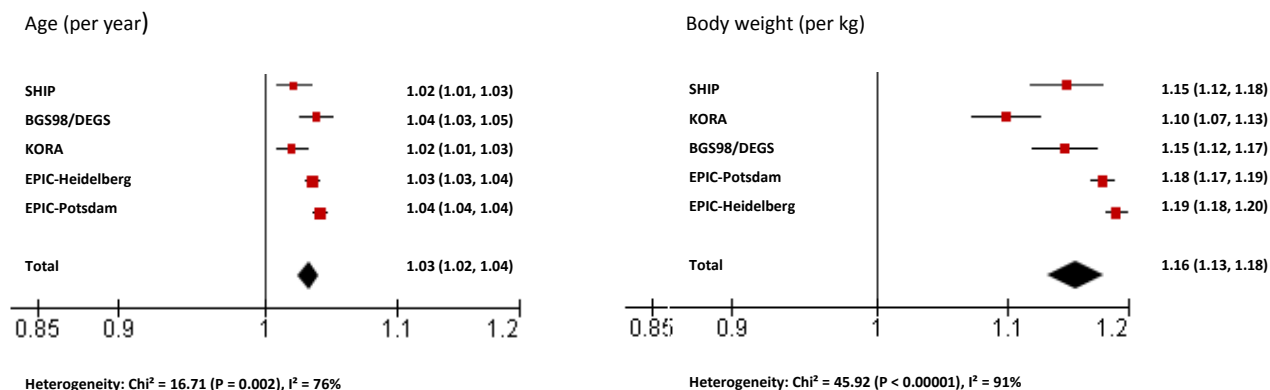
b, Associations with risk for SWCG

Figure 5. Associations of age and body weight with risk for SWG (a) and SWCG (b) in the selection models

The meta-analytically combined estimates of relative risks for the associations of predictors with SWG and SWCG are presented in **Table 6** for the maximum models, in **Table 7** for the minimum models, and in **Table 8** for the selection models, respectively. As expected, for a wide range of predictors, risk associations with SWG and SWCG were quite consistent. On the one hand higher educational level, body height, sporting activity, higher intake of fruits and vegetables, whole grain bread, cake and cookies, and fish, as well as living in partnership, were inversely related with the risk of SWG and SWCG; on the other hand alcohol abstinence, self-reported weight loss, smoking, higher intake of meat, and chocolate, high consumption of soft drinks, time spent watching TV, dissatisfaction with life, and suffering from sleep disorders were found to be positively related with such risks. Interestingly, some factors also revealed the opposing risk associations between SWG and SWCG. These predictors mainly comprised age, sex, body weight, and alcohol consumption. For instance, the relative risk of gaining $\geq 10\%$ of body weight decreased with advanced age, but the relative risk of gaining ≥ 2.5 cm of residual WC increased. Furthermore, women were at a higher risk of experiencing SWG, while men were exposed to higher risks of experiencing SWCG. In addition, moderate alcohol consumption was inversely associated with the risk of SWG, but tends to be positively associated with risk of SWCG.

Table 6. Meta-analytically combined estimates of relative risk for the associations of predictors with SWG and SWCG in the maximum models

Predictor variables	Maximum models					
	SWG			SWCG		
	β	Hazard ratio (95% CI)	Points allocated	β	Hazard ratio (95% CI)	Points allocated
Age (per year)	-0.0361	0.96 (0.96, 0.97)	-3.61	0.0346	1.04 (1.03, 1.04)	3.46
Sex (female vs. male)	0.1288	1.14 (0.98, 1.32)	12.88	-0.7005	0.50 (0.31, 0.79)	-70.05
Secondary/Professional school (vs. None/Primary school)	-0.1296	0.88 (0.78, 0.99)	-12.96	-0.1210	0.89 (0.76, 1.03)	-12.10
University (vs. None/Primary school)	-0.2676	0.77 (0.70, 0.83)	-26.76	-0.2757	0.76 (0.64, 0.90)	-27.57
Body weight (per kg)	-0.0156	0.98 (0.98, 0.99)	-1.56	0.1493	1.16 (1.14, 1.19)	14.93
Waist circumference (per cm)	-	-	-	-0.1512	0.86 (0.84, 0.88)	-15.12
Body height (per cm)	-0.0021	1.00 (0.99, 1.01)	-0.21	-0.0788	0.92 (0.91, 0.94)	-7.88
Self-reported weight loss (yes vs. no)	0.6877	1.99 (0.96, 4.13)	68.77	0.0437	1.04 (0.86, 1.27)	4.37
>0 to 2 h sports/week (vs. 0 h/week)	-0.0620	0.94 (0.88, 1.00)	-6.20	-0.0586	0.94 (0.86, 1.03)	-5.86
≥2 h sports/week (vs. 0 h/week)	-0.1267	0.88 (0.78, 0.99)	-12.67	-0.1830	0.83 (0.73, 0.95)	-18.30
0 g alcohol/day (vs. >0 to <6 g/day)	0.0527	1.05 (0.88, 1.26)	5.27	0.0780	1.08 (0.97, 1.20)	7.80
≥6 to <18 g alcohol/day (vs. >0 to <6 g/day)	-0.1133	0.89 (0.84, 0.95)	-11.33	0.0425	1.04 (0.94, 1.16)	4.25
≥18 g alcohol/day (vs. >0 to <6 g/day)	-0.1013	0.90 (0.79, 1.03)	-10.13	0.1016	1.11 (0.99, 1.23)	10.16
Smoking (vs. non-smoking)	0.3436	1.41 (1.31, 1.51)	34.36	0.0561	1.06 (1.00, 1.12)	5.61
Watching TV (h per day)	0.0276	1.03 (0.97, 1.09)	2.76	0.0021	1.00 (0.99, 1.01)	0.21
Sleep duration (h per day)	-0.0457	0.96 (0.93, 0.99)	-4.57	0.0048	1.00 (0.98, 1.03)	0.48
≥1.5 to <2 portions fruits and vegetables/day (vs. <1.5 portions/day)	-0.0450	0.96 (0.89, 1.03)	-4.50	-0.0718	0.93 (0.87, 0.99)	-7.18
≥2 portions fruits and vegetables/day (vs. <1.5 portions/day)	-0.0307	0.97 (0.90, 1.05)	-3.07	-0.1052	0.90 (0.84, 0.97)	-10.52
≥4-6 portions meat/week to <1 portion/day (vs. <4-6 portions/week)	0.0669	1.07 (0.95, 1.20)	6.69	0.0393	1.04 (0.94, 1.15)	3.93
≥1 portion meat/day (vs. <4 - 6 portions/week)	0.1137	1.12 (1.03, 1.22)	11.37	0.0667	1.07 (0.98, 1.17)	6.67
≥2-3 portions whole grain bread/week to <4-6 portions/week (vs. <2-3 portions/week)	-0.0589	0.94 (0.87, 1.02)	-5.89	-0.0598	0.94 (0.87, 1.02)	-5.98
≥4-6 portions whole grain bread/week (vs. <2-3 portions/week)	-0.1534	0.86 (0.80, 0.92)	-15.34	-0.0730	0.93 (0.88, 0.99)	-7.30
≥1 portion cake and cookies/week to <2-3 portions/week (vs. <1 portion/week)	-0.1298	0.88 (0.75, 1.03)	-12.98	-0.0005	1.00 (0.89, 1.13)	-0.05
≥2-3 portions cake and cookies/week (vs. <1portion /week)	-0.2488	0.78 (0.62, 0.98)	-24.88	-0.0028	1.00 (0.94, 1.06)	-0.28
≥2-3 portions soft drinks/month to <4-6 portions/week (vs. <2-3 portion/month)	0.1020	1.11 (1.04, 1.18)	10.20	-0.0097	0.99 (0.94, 1.05)	-0.97
≥4-6 portions soft drinks/week (vs. <2-3 portion/month)	0.1933	1.21 (1.06, 1.39)	19.33	0.0571	1.06 (0.98, 1.15)	5.71

Results

≥2-3 portions fish/month to <1 portions/week (vs. <2-3 portion/month)	-0.0232	0.98 (0.91, 1.05)	-2.32	-0.0290	0.97 (0.85, 1.11)	-2.90
≥1 portion fish/week (vs. <2-3 portion/month)	-0.0026	1.00 (0.91, 1.10)	-0.26	-0.0197	0.98 (0.91, 1.06)	-1.97
≥1 portion chocolate/week to <2-3 portions/week (vs. <1 portion/week)	0.0527	1.05 (0.98, 1.14)	5.27	0.0237	1.02 (0.96, 1.09)	2.37
≥2-3 portions chocolate/week (vs. <1portion/week)	0.0459	1.05 (0.98, 1.12)	4.59	0.0485	1.05 (0.99, 1.11)	4.85
Living in partnership (yes vs. no)	-0.2258	0.80 (0.74, 0.85)	-22.58	-0.0369	0.96 (0.90, 1.03)	-3.69
Rather satisfied with life (vs. satisfied)	0.1397	1.15 (1.04, 1.27)	13.97	-0.0191	0.98 (0.92, 1.04)	-1.91
(Rather) unsatisfied with life (vs. satisfied)	0.3562	1.43 (1.16, 1.75)	35.62	0.0299	1.03 (0.94, 1.13)	2.99
Sleep disorders (yes vs. no)	0.1672	1.18 (1.10, 1.27)	16.72	0.0579	1.06 (0.99, 1.13)	5.79

SWG = substantial weight gain; SWCG = substantial waist circumference gain.

By comparing the predictor associations in the maximum models with the associations of the corresponding predictors in the minimum models minor differences regarding the strength of risk associations have been observed (**Table 7**). Most obvious differences were seen for SWG in terms of stronger positive associations with smoking and, in particular, with suffering from sleep disorders in the minimum models.

Table 7. Meta-analytically combined estimates of relative risk for the associations of predictors with SWG and SWCG in the minimum models

Predictor variables	Minimum models					
	SWG			SWCG		
	β	Hazard ratio (95% CI)	Points allocated	β	Hazard ratio (95% CI)	Points allocated
Age (per year)	-0.0377	0.96 (0.96, 0.97)	-3.77	0.0341	1.03 (1.03, 1.04)	3.41
Sex (female vs. male)	0.1115	1.12 (0.97, 1.29)	11.15	-0.7191	0.49 (0.32, 0.75)	-71.91
Secondary/Professional school (vs. None/Primary school)	-0.1294	0.88 (0.78, 0.99)	-12.94	-0.1197	0.89 (0.77, 1.03)	-11.97
University (vs. None/Primary school)	-0.2822	0.75 (0.68, 0.83)	-28.22	-0.2704	0.76 (0.64, 0.91)	-27.04
Body weight (per kg)	-0.0171	0.98 (0.98, 0.99)	-1.71	0.1484	1.16 (1.14, 1.18)	14.84
Waist circumference (per cm)	-	-	-	-0.1509	0.86 (0.84, 0.88)	-15.09
Body height (per cm)	-0.0000	1.00 (0.99, 1.01)	-0.00	-0.0786	0.92 (0.91, 0.94)	-7.86
>0 to 2 h sports/week (vs. 0 h/week)	-0.0732	0.93 (0.87, 0.99)	-7.32	-0.0540	0.95 (0.87, 1.03)	-5.40
≥2 h sports/week (vs. 0 h/week)	-0.1187	0.89 (0.77, 1.02)	-11.87	-0.1847	0.83 (0.73, 0.95)	-18.47
0 g alcohol/day (vs. >0 to <6 g/day)	0.0701	1.07 (0.90, 1.28)	7.01	0.0573	1.06 (0.95, 1.18)	5.73

Results

≥6 to <18 g alcohol/day (vs. >0 to <6 g/day)	-0.1252	0.88 (0.83, 0.94)	-12.52	0.0331	1.03 (0.93, 1.15)	3.31
≥18 g alcohol/day (vs. >0 to <6 g/day)	-0.1175	0.89 (0.80, 0.99)	-11.75	0.0775	1.08 (0.98, 1.19)	7.75
Smoking (vs. non-smoking)	0.3889	1.48 (1.40, 1.56)	38.89	0.0590	1.06 (1.01, 1.12)	5.90
≥1.5 to <2 portions fruits and vegetables/day (vs. <1.5 portions/day)	-0.0597	0.94 (0.88, 1.01)	-5.97	-0.0661	0.94 (0.88, 0.99)	-6.61
≥2 portions fruits and vegetables/day (vs. <1.5 portions/day)	-0.0459	0.96 (0.86, 1.06)	-4.59	-0.1087	0.90 (0.84, 0.95)	-10.87
≥4-6 portions meat/week to <1 portion/day (vs. <4-6 portions/week)	0.0697	1.07 (0.98, 1.18)	6.97	0.0222	1.02 (0.94, 1.11)	2.22
≥1 portion meat/day (vs. <4 - 6 portions/week)	0.1061	1.11 (1.01, 1.22)	10.61	0.0707	1.07 (0.99, 1.16)	7.07
≥2-3 portions whole grain bread/week to <4-6 portions/week (vs. <2-3 portions/week)	-0.0609	0.94 (0.87, 1.02)	-6.09	-0.0497	0.95 (0.85, 1.06)	-4.97
≥4-6 portions whole grain bread/week (vs. <2-3 portions/week)	-0.1454	0.86 (0.81, 0.92)	-14.54	-0.0710	0.93 (0.88, 0.99)	-7.10
≥1 portion cake and cookies/week to <2-3 portions/week (vs. <1 portion/week)	-0.1300	0.88 (0.75, 1.02)	-13.00	0.0009	1.00 (0.90, 1.11)	0.09
≥2-3 portions cake and cookies/week (vs. <1portion/week)	-0.2643	0.77 (0.61, 0.96)	-26.43	-0.0003	1.00 (0.94, 1.06)	-0.03
≥2-3 portions soft drinks/month to <4-6 portions/week (vs. <2-3 portion/month)	0.1011	1.11 (1.04, 1.18)	10.11	0.0014	1.00 (0.95, 1.06)	0.14
≥4-6 portions soft drinks/week (vs. <2-3 portion/month)	0.1614	1.18 (1.03, 1.34)	16.14	0.0558	1.06 (0.98, 1.14)	5.58
≥2-3 portions fish/month to <1 portions/week (vs. <2-3 portion/month)	-0.0229	0.98 (0.90, 1.06)	-2.29	-0.0255	0.97 (0.87, 1.10)	-2.55
≥1 portion fish/week (vs. <2-3 portion/month)	-0.0098	0.99 (0.90, 1.09)	-0.98	-0.0194	0.98 (0.92, 1.05)	-1.94
≥1 portion chocolate week to <2-3 portions/week (vs. <1 portion/week)	0.0485	1.05 (0.97, 1.13)	4.85	0.0190	1.02 (0.96, 1.09)	1.90
≥2-3 portions chocolate/week (vs. <1portion/week)	0.0325	1.03 (0.95, 1.12)	3.25	0.0365	1.04 (0.98, 1.10)	3.65
Sleep disorders (yes vs. no)	0.2506	1.28 (1.20, 1.38)	25.06	0.0512	1.05 (0.99, 1.12)	5.12

SWG = substantial weight gain; SWCG = substantial waist circumference gain.

In the selection models, except from body weight, all estimates of relative risk for the associations with SWG and SWCG increased in strength, in comparison with the corresponding minimum models.

Table 8. Meta-analytically combined estimates of relative risk for the associations of predictors with SWG and SWCG in the selection models

Predictor variables	Selection models					
	SWG			SWCG		
	β	Hazard ratio (95% CI)	Points allocated	β	Hazard ratio (95% CI)	Points allocated
Age (per year)	-0.0347	0.97 (0.96, 0.97)	-3.47	0.0316	1.03 (1.02, 1.04)	3.16
Sex (female vs. male)	0.1522	1.16 (1.04, 1.30)	15.22	-0.7533	0.47 (0.30, 0.73)	-75.33

Results

Body weight (per kg)	-0.0158	0.98 (0.98, 0.99)	-1.58	0.1443	1.16 (1.13, 1.18)	14.43
Waist circumference (per cm)	-	not selected	-	-0.1451	0.86 (0.85, 0.88)	-14.51
Body height (per cm)	-	not selected	-	-0.0776	0.93 (0.91, 0.94)	-7.76
Secondary/Professional school (vs. None/Primary school)	-	not selected	-	-0.1399	0.87 (0.75, 1.01)	-13.99
University (vs. None/Primary school)	-	not selected	-	-0.3007	0.74 (0.61, 0.89)	-30.07
Smoking (vs. non-smoking)	0.4385	1.55 (1.47, 1.64)	43.85	-	not selected	-
≥4-6 portions meat/week to <1 portion/day (vs. <4-6 portions/week)	-	not selected	-	0.0358	1.04 (0.95, 1.13)	3.58
≥1 portion meat/day (vs. <4-6 portions/week)	-	not selected	-	0.0901	1.09 (1.00, 1.20)	9.01
≥2-3 portions soft drinks/month to <4-6 portions/week (vs. <2-3 portion/month)	0.1367	1.15 (1.08, 1.22)	13.67	-	not selected	-
≥4-6 portions soft drinks/week (vs. <2-3 portion/month)	0.2232	1.25 (1.15, 1.36)	22.32	-	not selected	-

SWG = substantial weight gain; SWCG = substantial waist circumference gain.

To assign score values for each predictor, regression coefficients (multiplied by 100) were used. For each individual, risk scores (RS_i) for SWG and SWCG were calculated by a linear combination of single predictors according to the following formulas (exemplarily shown for the selection models):

SWG:

$RS_i = 800$ $- 3.47 * \text{age (y)}$ $+ 15.22 * \text{female sex}$ $- 1.58 * \text{body weight (kg)}$ $+ 43.85 * \text{current smoking}$ $+ 13.67 * \text{soft drinks } (\geq 2-3 \text{ portions/month to } < 4-6 \text{ portions/week})$ $+ 22.32 \text{ soft drinks } (\geq 4-6 \text{ portions/week})$
--

SWCG:

$RS_i = 2,500$ $+ 3.16 * \text{age (y)}$ $- 75.33 * \text{female sex}$ $+ 14.43 * \text{body weight (kg)}$ $- 14.51 * \text{waist circumference (cm)}$ $- 7.76 * \text{body height (cm)}$ $- 13.99 * \text{secondary/professional school}$ $- 30.07 * \text{university}$ $+ 3.58 * \text{meat } (\geq 4-6 \text{ portions/week to } < 1 \text{ portion/day})$ $+ 9.01 * \text{meat } (\geq 1 \text{ portion/day})$

The meta-analytically combined estimate of the background survival probability for five years, estimated at average values of the predictors, were 0.9584 for SWG and 0.8981 for SWCG in the maximum models. It means that a hypothetical, average individual has a probability of 95.84 % and 89.81 % to survive the following five years without experiencing

$\geq 10\%$ gain in weight and ≥ 2.5 cm gain in WC_{BMI} , respectively. Consequently, the risks of gaining $\geq 10\%$ of baseline weight and ≥ 2.5 cm of baseline WC_{BMI} within five years were found to be 4.16 % and 10.19 % for this individual. The corresponding risk score values were 442.94 and 1,096.72. For the minimum models, the meta-analytically combined estimates of the background survival probability for five years were 0.9574 (SWG; risk score value: 482.51) and 0.8972 (SWCG; risk score value: 1,105.10), while they were 0.9554 (SWG; risk score value: 580.69) and 0.8955 (SWCG; risk score value: 1,133.32) for the selection models. Depending on the used predictor set, the individual probabilities of experiencing $\geq 10\%$ gain in body weight and ≥ 2.5 cm gain in WC_{BMI} within the following five years were calculated by inserting the individual's risk score into the corresponding survival functions as given below:

Survivals function of the maximum models:

$$P_{(SWG, 5y)} = 1 - 0.9584 + \exp^{[RSi - 442.94] / 100}$$

$$P_{(SWCG, 5y)} = 1 - 0.8981 + \exp^{[RSi - 1,096.72] / 100}$$

Survival functions of the minimum models:

$$P_{(SWG, 5y)} = 1 - 0.9574 + \exp^{[RSi - 482.51] / 100}$$

$$P_{(SWCG, 5y)} = 1 - 0.8972 + \exp^{[RSi - 1,105.10] / 100}$$

Survival functions of the selection models:

$$P_{(SWG, 5y)} = 1 - 0.9554 + \exp^{[RSi - 580.69] / 100}$$

$$P_{(SWCG, 5y)} = 1 - 0.8955 + \exp^{[RSi - 1,133.32] / 100}$$

3.2.2 Predictive performances of the German risk scores

The discriminatory abilities of the maximum models for SWG, as assessed by aROC (95% CI), reached acceptable values (≥ 0.70) (**Table 9 a**). It means that in at least 70 % of cases individuals who gained ≥ 10 % of their baseline-based body weight within five years had higher predicted risks than individuals who did not. Between the cohorts, discriminatory accuracy varied from 0.71 (0.68, 0.73) in EPIC-Heidelberg to 0.80 (0.76, 0.85) in the BGS98/DEGS cohort. Compared with the models for SWG, discriminatory abilities of the maximum models for SWCG were lower in most of the cohorts, ranging from 0.61 (0.58, 0.65) in KORA to 0.73 (0.72, 0.74) in EPIC-Heidelberg. The meta-analytically combined aROCs (95 % CI) of the maximum models were 0.73 (0.71, 0.76) for SWG, and 0.69 (0.66, 0.73) for SWCG. By applying the homogeneous predictor set of the minimum models discriminatory abilities slightly decreased across the cohorts (**Table 9 b**). The decrease was most pronounced in the EPIC-Potsdam cohort (for SWG) and in the BGS98/DEGS cohorts (for SWCG). Despite the homogeneous predictor set, the discriminatory abilities of the risk scores for SWG and SWCG considerably varied, ranging from 0.68 (0.67, 0.70) in EPIC-Potsdam to 0.79 (0.74, 0.84) in the BGS98/DEGS cohort for SWG and from 0.61 (0.58, 0.64) in KORA to 0.73 (0.72, 0.74) in EPIC-Heidelberg for SWCG. While the variation was attributable to one specific cohort (BGS98/DEGS) in case of SWG with comparatively higher discriminatory ability, aROCs (95% CI) for SWCG generally varied across the cohorts. The meta-analytically combined aROCs (95 % CI) were 0.71 (0.68, 0.75) for SWG, and 0.68 (0.65, 0.72) for SWCG. By further restriction of the predictor set to the most predictive variables, as applied in the selection models, discriminatory abilities only marginally changed (**Table 9 c**). The meta-analytically combined aROCs (95% CI) for the selection models were 0.70 (0.67 0.73) for SWG and 0.68 (0.64, 0.71) for SWCG, as illustrated in **Figure 6**.

By comparing the models using meta-analytically combined predictor weights with those using cohort-specific predictor weights (marked in gray in **Table 9 a-c**), only minor differences were observed for most models; clear differences in discriminatory accuracy were particularly seen for SWCG models in the KORA cohort.

Results

Table 9. Predictive performances of the risk scores across the cohorts

a, Maximum models: meta-analytically combined ROC (95% CI) for SWG: **0.73** (0.71, 0.76), and for SWCG: **0.69** (0.66, 0.73)

Outcome	EPIC-Potsdam		EPIC-Heidelberg		SHIP		KORA		BGS98/DEGS	
	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT
SWG	0.72 (0.71, 0.74)	<0.001	0.71 (0.68, 0.73)	<0.001	0.74 (0.68, 0.79)	0.821	0.72 (0.67, 0.77)	0.195	0.80 (0.76, 0.85)	<0.001
	0.73 (0.71, 0.74)	<0.001	0.72 (0.70, 0.74)	<0.001	0.74 (0.69, 0.80)	0.598	0.73 (0.68, 0.78)	0.798	0.80 (0.75, 0.85)	0.228
SWCG	0.70 (0.69, 0.71)	<0.001	0.73 (0.72, 0.74)	<0.001	0.72 (0.67, 0.77)	<0.001	0.61 (0.58, 0.65)	<0.001	0.70 (0.65, 0.75)	0.778
	0.73 (0.71, 0.74)	0.010	0.74 (0.73, 0.75)	0.055	0.73 (0.68, 0.78)	0.736	0.70 (0.67, 0.73)	0.186	0.71 (0.66, 0.76)	0.663

aROCs (95% CI) using meta-analytically combined predictor weights are marked in black; aROCs (95% CI) using cohort-specific predictor weights are marked in gray; HLT = Hosmer-Lemeshow test.

b, Minimum models: meta-analytically combined ROC (95% CI) for SWG: **0.71** (0.68, 0.75), and for SWCG: **0.68** (0.65, 0.72)

Outcome	EPIC-Potsdam		EPIC-Heidelberg		SHIP		KORA		BGS98/DEGS	
	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT
SWG	0.68 (0.67, 0.70)	<0.001	0.71 (0.69, 0.73)	<0.001	0.72 (0.66, 0.78)	0.138	0.69 (0.64, 0.74)	0.881	0.79 (0.74, 0.84)	<0.001
	0.69 (0.67, 0.70)	<0.001	0.71 (0.69, 0.73)	<0.001	0.71 (0.66, 0.77)	0.409	0.70 (0.65, 0.75)	0.703	0.79 (0.74, 0.83)	0.544
SWCG	0.70 (0.69, 0.71)	<0.001	0.73 (0.72, 0.74)	<0.001	0.72 (0.67, 0.77)	0.002	0.61 (0.58, 0.64)	<0.001	0.64 (0.59, 0.69)	<0.001
	0.72 (0.71, 0.73)	0.007	0.74 (0.73, 0.74)	0.028	0.73 (0.68, 0.78)	0.273	0.69 (0.66, 0.72)	0.192	0.71 (0.66, 0.75)	0.679

aROCs (95% CI) using meta-analytically combined predictor weights are marked in black; aROCs (95% CI) using cohort-specific predictor weights are marked in gray; HLT = Hosmer-Lemeshow test.

c, Selection models meta-analytically combined ROC (95% CI) for SWG: **0.70** (0.67, 0.73), and for SWCG: **0.68** (0.64, 0.71)

Outcome	EPIC-Potsdam		EPIC-Heidelberg		SHIP		KORA		BGS98/DEGS	
	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT
SWG	0.67 (0.66, 0.69)	<0.001	0.70 (0.68, 0.72)	<0.001	0.70 (0.64, 0.76)	0.713	0.66 (0.61, 0.71)	0.143	0.78 (0.73, 0.82)	0.066
	0.68 (0.66, 0.69)	<0.001	0.70 (0.68, 0.72)	<0.001	0.70 (0.65, 0.77)	0.358	0.66 (0.61, 0.72)	0.435	0.77 (0.73, 0.82)	0.101
SWCG	0.70 (0.69, 0.71)	<0.001	0.73 (0.72, 0.74)	<0.001	0.71 (0.66, 0.76)	0.010	0.60 (0.57, 0.63)	<0.001	0.64 (0.59, 0.69)	<0.001
	0.72 (0.71, 0.73)	0.038	0.73 (0.72, 0.74)	0.110	0.72 (0.66, 0.77)	0.704	0.69 (0.66, 0.72)	0.177	0.69 (0.64, 0.74)	0.839

aROCs (95% CI) using meta-analytically combined predictor weights are marked in black; aROCs (95% CI) using cohort-specific predictor weights are marked in gray; HLT = Hosmer-Lemeshow test.

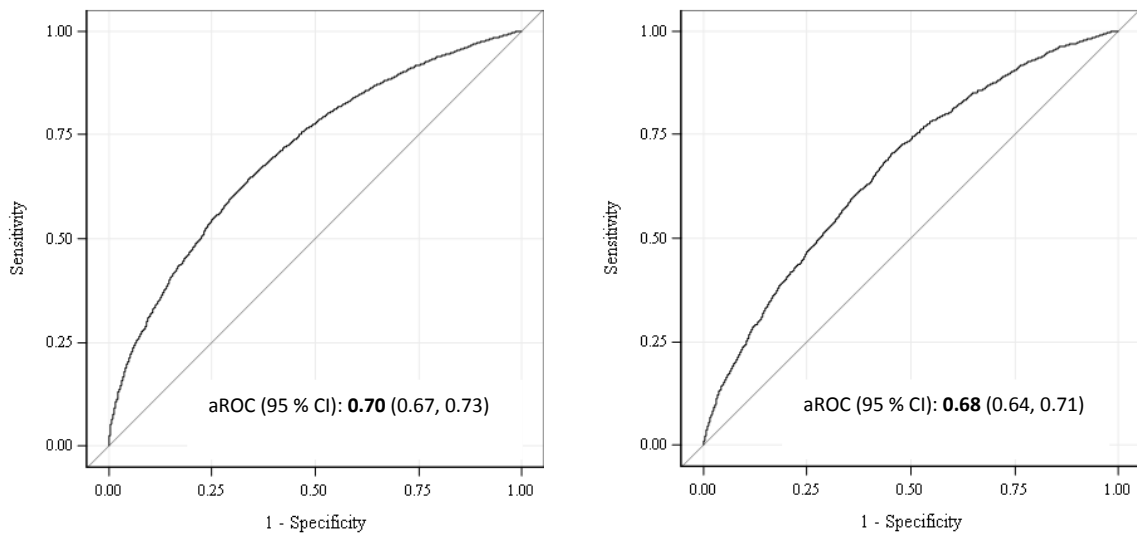


Figure 6. Meta-analytically combined aROCs of the selection models for the prediction of SWG (left) and SWCG (right) over five years

To determine an appropriate cut-off value for identification of high-risk individuals, measures of discriminatory accuracy, including sensitivity, specificity, PPV, and NPV, can be used. Owing to the uniform predictor set across the cohorts and nearly unchanged discriminatory abilities compared to the minimum models, such measures of the selection models are shown in the following. **Table 10 a-e** and **Table 11 a-e** provide these measures of the selection models for SWG and SWCG, separated by the cohorts. For instance, a cut-off at the score value of ≥ 550 in the model for SWG corresponded in the EPIC-Potsdam cohort to a sensitivity and specificity of 73.5 and 52.8, respectively. It means that individuals who gained $\geq 10\%$ of their body weight in the course of five years had the probability of 73.5 % to have a score value of 550 or above. Furthermore, individuals who did not experience SWG within five years had the probability of 52.8 % to have a score value below this. The corresponding Youden's index, which maximizes sensitivity and specificity in the EPIC-Potsdam cohort, was 0.263. The probability to gain $\geq 10\%$ of baseline weight within five years for individuals with score values ≥ 550 was 9.1 % (PPV), while the probability to gain $< 10\%$ of baseline weight for individuals below this score value was 96.9 % (NPV). In EPIC-Heidelberg and KORA sensitivity and specificity were also maximized at the score value of ≥ 550 , while the corresponding Youden's indices were 0.315 and 0.253, respectively. In SHIP and BGS98/DEGS, however, the

appropriate cut-off to identify high-risk individuals was determined at a score value of ≥ 600 . Nonetheless, across all the cohorts NPVs were clearly above 94 % at score values of ≥ 550 as well as ≥ 600 - it implies that in all cohorts the score had a high ability to exclude SWG in the course of five years.

Table 10. Sensitivity, specificity, PPV and, NPV for various cut-off points of the selection models for SWG

a, EPIC-Potsdam (J = 0.263)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥ 525	48.7	88.9	32.8	0.217	7.8	97.9
≥ 550	31.1	73.5	52.8	0.263	9.1	96.9
≥ 575	14.3	53.3	70.5	0.238	10.4	95.9
≥ 600	5.4	27.4	86.6	0.140	11.6	94.9
≥ 625	1.4	12.5	95.1	0.076	13.8	94.4

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

b, EPIC-Heidelberg (J = 0.315)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥ 525	43.4	88.3	37.7	0.260	5.7	98.7
≥ 550	26.9	73.6	57.9	0.315	7.0	98.1
≥ 575	11.4	50.4	74.1	0.245	7.7	97.2
≥ 600	4.3	25.6	89.2	0.148	9.2	96.5
≥ 625	1.1	11.5	96.0	0.075	11.0	96.2

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

c, SHIP (J = 0.311)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥ 525	66.7	97.0	17.8	0.148	5.9	99.2
≥ 550	53.3	87.9	34.3	0.222	6.5	98.2
≥ 575	37.1	78.8	48.0	0.268	7.4	97.8
≥ 600	21.5	66.7	64.4	0.311	9.0	97.4
≥ 625	10.3	47.0	79.8	0.268	10.9	96.6

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

d, KORA (J = 0.253)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥ 525	61.5	88.8	23.9	0.127	6.2	98.1
≥ 550	44.8	85.4	39.9	0.253	7.2	98.0
≥ 575	27.1	67.4	56.4	0.238	7.8	97.0
≥ 600	12.9	46.2	74.1	0.203	9.2	96.2
≥ 625	4.6	25.8	87.9	0.137	10.7	95.5

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

e, BGS98/DEGS (J = 0.418)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥525	72.1	100.0	15.1	0.151	5.1	100.0
≥550	57.4	97.0	29.1	0.261	5.9	99.6
≥575	39.2	91.0	44.2	0.352	7.0	99.1
≥600	23.3	79.1	62.7	0.418	8.9	98.5
≥625	12.2	59.7	78.4	0.381	11.1	97.7

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

Regarding the selection model for SWCG, more heterogeneity has been found with respect to the determination of an appropriate cut-off to identify high-risk individuals (**Table 11**). In EPIC-Heidelberg and SHIP sensitivity and specificity were maximized at the score value of ≥1,100, while the corresponding Youden's indices were 0.349 and 0.289, respectively. In KORA the appropriate cut-off was determined at ≥1,050 (J = 0.146), in EPIC-Potsdam at ≥1,150 (J = 0.280), and in BGS98/DEGS at a score value of ≥1,200 (J = 0.207). Despite this heterogeneity, NPVs were still above 82 % at these various cut-offs across all cohorts, implying that at score values of ≥1,050 to ≥1,200 the score was still adequately able to exclude SWCG in the course of five years in all cohorts.

Table 11. Sensitivity, specificity, PPV and, NPV for various cut-off points of the selection model for SWCG

a, EPIC-Potsdam (J = 0.280)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥1,000	83.8	99.7	4.1	0.038	14.3	99.0
≥1,050	56.3	95.1	17.8	0.129	15.7	95.9
≥1,100	25.5	80.1	47.4	0.275	19.7	93.7
≥1,150	7.5	49.4	78.6	0.280	27.1	90.6
≥1,200	1.6	20.8	94.8	0.156	38.8	88.1

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

b, EPIC-Heidelberg (J = 0.349)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥1,000	73.8	98.6	10.6	0.092	22.0	96.6
≥1,050	46.3	92.2	30.7	0.229	25.4	93.9
≥1,100	20.5	74.2	60.7	0.349	32.6	90.2
≥1,150	6.4	45.0	85.9	0.309	45.1	85.9
≥1,200	1.5	18.7	96.9	0.156	60.8	82.2

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

c, SHIP (J = 0.289)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥1,000	84.5	100.0	4.0	0.040	7.2	100.0
≥1,050	57.7	96.7	16.4	0.131	7.9	98.7
≥1,100	27.7	84.6	44.3	0.289	9.9	97.5
≥1,150	8.5	53.8	74.2	0.280	13.3	95.6
≥1,200	2.1	29.7	93.1	0.228	23.3	94.7

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

d, KORA (J = 0.146)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥1,000	67.3	96.5	13.5	0.100	15.0	96.1
≥1,050	35.1	79.9	34.7	0.146	16.3	92.0
≥1,100	11.6	44.5	66.4	0.109	17.0	88.3
≥1,150	2.4	16.6	89.2	0.058	13.0	87.1
≥1,200	0.2	3.1	97.7	0.008	15.4	86.5

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

e, BGS98/DEGS (J = 0.207)

Score value	Percentage of the population	Sensitivity (%)	Specificity (%)	Youden's Index (J)	PPV (%)	NPV (%)
≥1,000	92.8	100.0	1.8	0.018	7.6	100.0
≥1,050	77.5	97.4	7.6	0.050	7.8	97.4
≥1,100	55.2	89.5	23.5	0.130	8.6	96.8
≥1,150	32.7	72.8	46.3	0.191	9.9	95.8
≥1,200	15.3	51.8	68.9	0.207	10.9	94.6

PPV = positive predictive value; NPV = negative predictive value. Youden's Index J = (sensitivity [%] + specificity [%] - 100) / 100

In addition to the scores' ability to discriminate between cases and non-cases, their ability to quantify absolute risks of experiencing SWG and SWCG within five years have been evaluated. Thereby, comparisons between observed and predicted probabilities across the deciles of predicted risk have been performed. Complementary to the measures of discriminatory accuracy, **Figure 7** and **8** illustrate the calibration plots for SWG and SWCG of the selection models by the cohorts. In the lower deciles the estimated probability of experiencing SWG within five years tends to overestimate the observed proportion of incident cases across all cohorts. In the middle deciles the predicted and observed risks of SWG agreed quite well, while in the higher deciles some underestimation of the observed incidence was present. Underestimation in the higher deciles was most pronounced in EPIC-Potsdam. In total, calibration plots for experiencing weight gain of ≥10 % in the course of five years showed same patterns across all cohorts, which were even tendentially observed if

cohort-specific predictor weights and background survival probabilities were used (lower row in **Figure 7**).

Regarding SWCG, calibration plots across the cohorts were more heterogeneous. While in the German EPIC cohorts and in the KORA cohort the score generally underestimated the observed incidences of individuals gaining ≥ 2.5 cm of residual waist in the course of five years, the score overestimated the proportion of the cases across the deciles in SHIP and BGS98/DEGS. If cohort-specific predictor weights and background survival probabilities were used, predicted absolute risks agreed well with observed cases in all cohorts (lower row in **Figure 8**).

Results

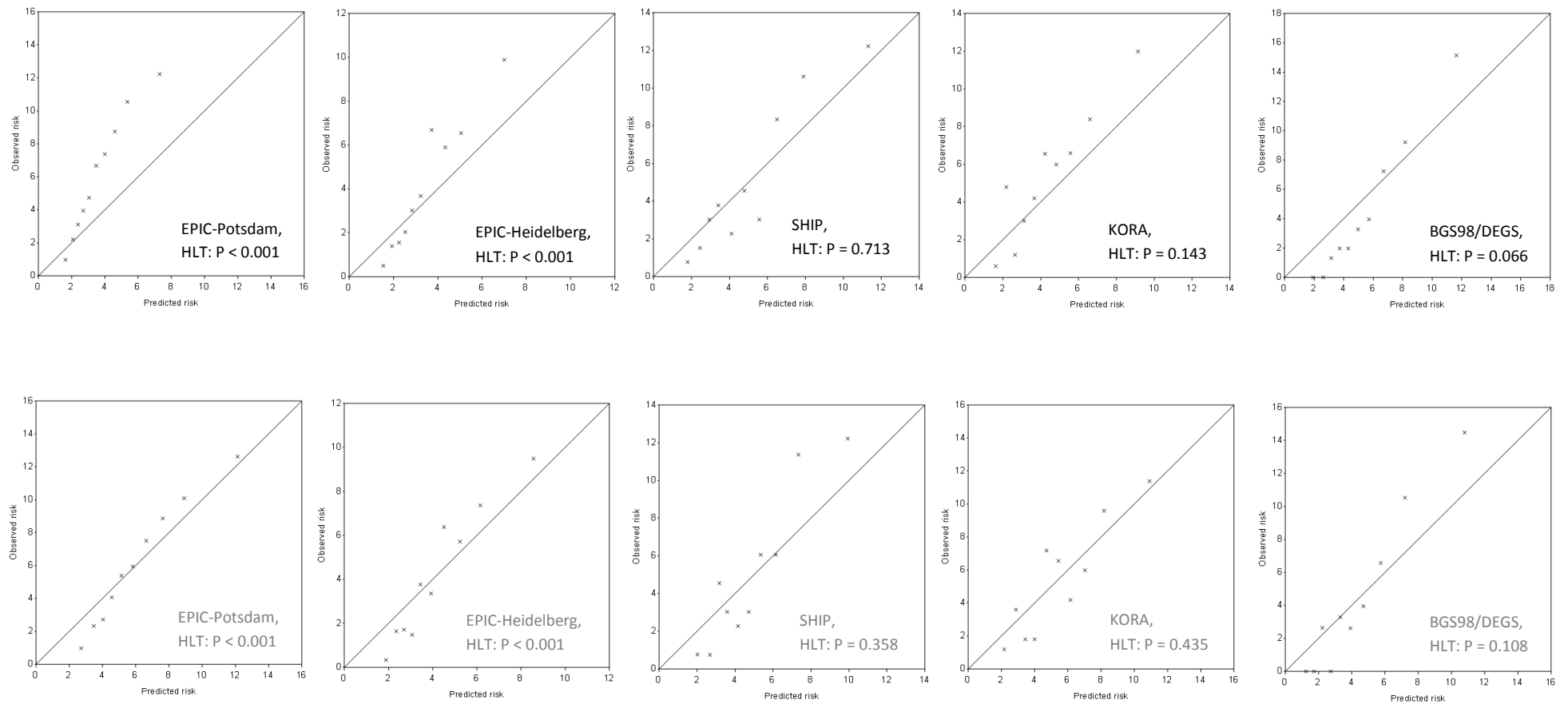


Figure 7. Calibration plots for SWG using the selection model

Upper row shows plots based on meta-analytical combined predictor weights and background survival probabilities, lower row shows plots based on cohort-specific predictor weights and background survival probabilities

Results

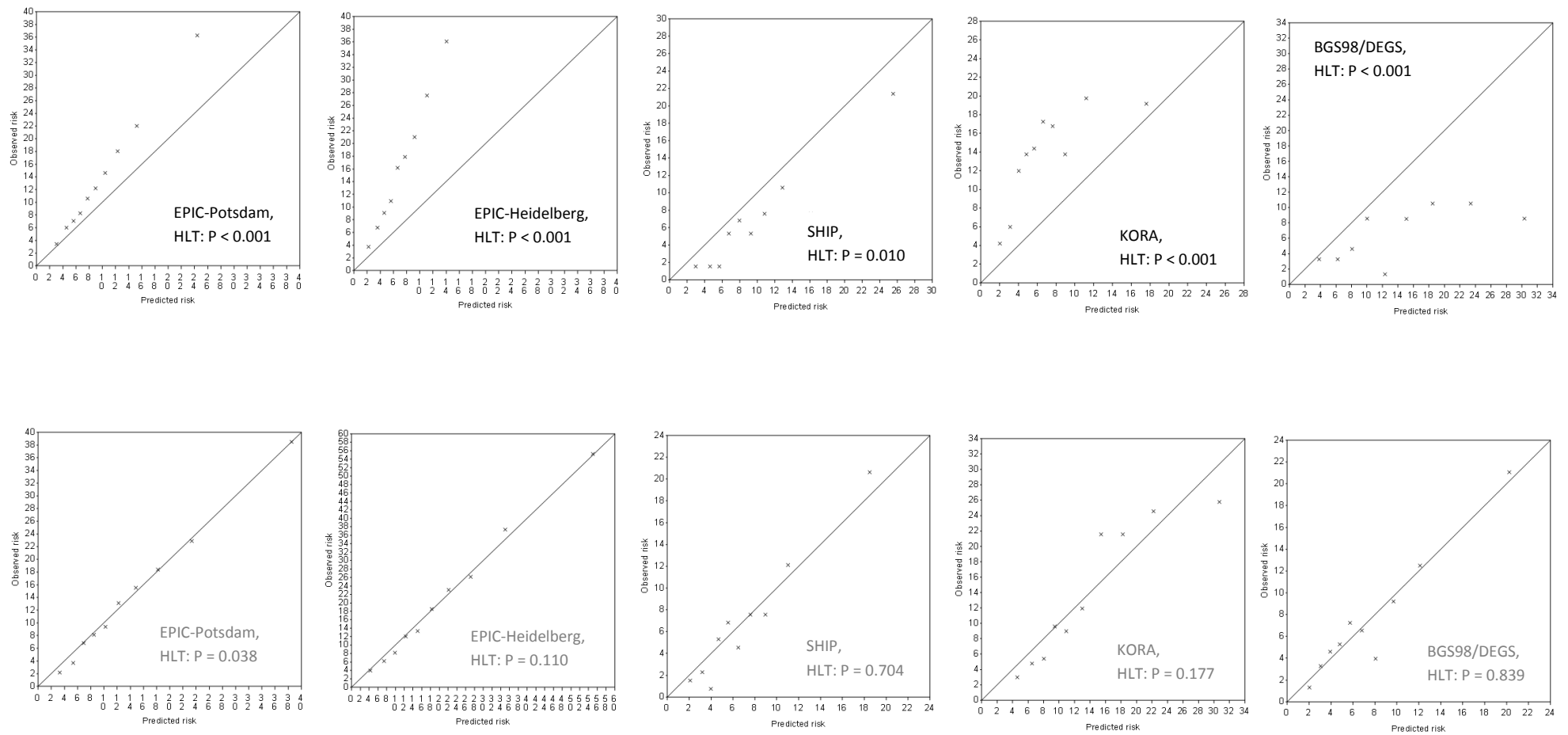


Figure 8. Calibration plots for SWCG using the selection model

Upper row shows plots based on meta-analytical combined predictor weights and background survival probabilities, lower row shows plots based on cohort-specific predictor weights and background survival probabilities

4. Discussion

Risk scores predicting SWG and SWCG within five years were derived and validated in the present study by using a database of over 31,000 adults from five German cohort studies. The predictor candidates comprised easily obtainable information on socio-demographic and anthropometric characteristics, dietary and lifestyle factors, and psychosocial and further health-related conditions. To precisely predict the five-year risks based on limited information and to avoid over-adaption to a pre-specified derivation sample, a three-step meta-analytical approach was applied. Thereby, risk scores were first derived using maximal available sets of predictor candidates. Next, the predictors were restricted to homogeneous sets, including solely predictors, which were assessed across all cohorts. Finally, homogenous predictor sets were reduced to the most predictive factors by carrying out a selection procedure. Across the steps meta-analytically combined aROCs (95 % CI) for the scores predicting SWG slightly diminished from 0.73 (0.71, 0.76) to 0.71 (0.68, 0.75) and 0.70 (0.67, 0.73). Discriminatory abilities remained unchanged when cohort-specific predictor weights were used, and the quantification of absolute risks of experiencing SWG showed same patterns across the cohorts. Discriminatory abilities for the scores predicting SWCG were relatively constant across the steps, with aROCs (95 % CI) of 0.69 (0.66, 0.73), 0.68 (0.65, 0.72) and 0.68 (0.64, 0.71), but more variation was observed between the cohorts. Additionally, using cohort-specific predictor weights could improve discriminatory abilities in some cohorts, and quantification of absolute risks of experiencing SWCG was inconsistent as well. Overall, the findings indicate that a Germany-wide risk score for SWG, taking account of the complete range of available, easily obtainable predictors is able to reach acceptable levels of discriminatory ability in diverse German cohorts. Furthermore, independent of the applied predictor set, Germany-wide risk scores for SWG are not inferior to the cohort-specific ones. The findings also suggest that the performances of the risk scores for SWCG are inferior to those of SWG and additionally appear more cohort-specific.

4.1 Material and methods

4.1.1 Study design and study population

The present study includes data regarding male and female study participants taking part in one of the five numerically most important, well-conducted German cohort studies. Four studies randomly recruited their participants from the local general population in the cities of their study centers and surrounding areas, including the federal states of Brandenburg (EPIC-Potsdam), Baden-Wuerttemberg (EPIC-Heidelberg), Mecklenburg-Western Pomerania (SHIP) and Bavaria (KORA). The data of these studies was complemented by a nationwide study that included participants from all federal states. Across all cohorts baseline assessments were conducted in comparable time frames - starting in the mid to late 90s. The largest proportion of participants came from the two German EPIC cohorts (about 85 %), while the cohorts of the SHIP study, the research platform KORA, and the nationwide BGS98/DEGS cohort contributed in more or less equal parts (about 5 %). The major strengths of the present study are certainly its prospective study design, large sample size, and inclusion of men and women taking part in various German cohorts. Consideration of all numerically important German cohort studies was assumed to allow generalizability of the derived risk scores to the German population.

4.1.2 Data quality of predictors and outcomes

Derivation of highly predictive risk scores requires valid estimates of both predictor-outcome associations and incidences of the predicted outcomes. The predictor range included in the present study comprised easily obtainable, non-invasive risk factors related to gain in body weight and/or WC. To assess baseline predictor information, different methods were applied across the German cohorts, covering self-administered questionnaires as well as computer-assisted and face-to-face interviews. Each of these methods has advantages and disadvantages. Questionnaires are most appropriate for data acquisition in case of large numbers of participants, because their application is inexpensive and time-saving. They collect data in a highly standardized and anonymous way, which facilitates data processing and data evaluation. However, in case of questionnaires, questions and response options are mostly pre-specified and misunderstanding cannot be cleared-up instantaneously, as it is possible by face-to-face interviews. Face-to-face interviews additionally provide the

opportunity to collect in-depth data regarding the participants, which potentially increases the accuracy of assessed information, and they are not necessarily bound to pre-defined questions. Since responses must be given directly to an interviewer, response behavior may be more affected in comparison with questionnaires, in particular, regarding issues relating to social desirability. Application of different data assessment methods might have contributed to the observed heterogeneity of the predictors' risk estimates between the cohorts.

Dietary intake was assessed by means of FFQs requesting usual food consumption over the past 12 months. FFQs are most commonly used in epidemiological studies and acquire dietary information in a simple cost-effective and time-efficient way (230). The validity of FFQs is measured by correlating FFQ-assessed dietary information with more precise reference methods of dietary assessment. The correlation coefficients of individual food items typically range from 0.4 to 0.6 (231). Consequently, the FFQ-based assessment of dietary factors might have resulted in misestimation of their predictive potential in the derived scores. Various aspects have to be taken into account regarding the accuracy of FFQ-obtained information. For instance, dietary data collected by FFQ largely depends on the long-term memory of the participants and on the spectrum of the requested food items (232). All included cohort studies assessed habitual dietary intake of the preceding 12 months, but the scope of the covered food items differed markedly. While the German EPIC studies assessed very detailed information relating to consumption frequency and portion size (covering 148 items), the available dietary information of the other studies was limited to consumption frequencies and covered less than 40 items. The heterogeneity of diet-related risk estimates in the present study might partly be explained by such differences in the assessment of dietary intakes between the cohorts. In addition, the validity of the assessed information may vary between food items. For instance, some food groups are more socially desirable (e.g., fruits and vegetables) and therefore more prone to be over-reported, while others are negatively connoted (e.g., sugary and fatty foods) and more likely to be under-reported. Similar influences of misreporting on the validity of collected information are reported for physical activity-related factors (e.g., sports activity, sedentary behavior) (233). Such bias might have limited the ability to obtain accurate risk estimates and also might have impaired the predictive performance of the derived risk scores. In

contrast, a greater validity is shown for self-reports of alcohol consumption (234, 235) and smoking status (236), and can also be assumed for social characteristics (e.g., educational level, living in partnership). Regarding the assessment of life satisfaction, history of weight loss, and presence of sleep disorders, in particular, differing requests of corresponding information across the included cohorts have to be kept in mind. Such differences might have introduced heterogeneity in the estimation of relative risks and potentially limited their prediction strengths.

Additional heterogeneity might be introduced in the present study with respect to the outcome assessment. While standardized measurements of anthropometry at baseline were taken by trained staff in all cohort studies, at follow-up anthropometry was measured according to the same procedures mentioned solely in the studies SHIP, KORA, and DEGS. In the German EPIC cohorts anthropometric follow-up information was self-reported by the participants. However, to account for potential misreporting, anthropometric self-reports were corrected using specifically developed correction equations, as described previously (221). With regard to the predicted outcomes, a further limitation may exist concerning the determination of incident cases. Since anthropometric assessments of only two time points were used to assess anthropometric changes over time, incidences of SWG and SWCG were not identified until follow-up data collection. To estimate the theoretical points in time when the thresholds of SWG and SWCG were passed, linear anthropometric changes over time were taken as the basis. However, on individual level, body weight and WC rather tend to fluctuate over time (237, 238). As a consequence, the assumption of linear gain might have resulted in misclassification of cases and non-cases and misestimation of the cases' survival times - all of which might have reduced the predictive performances of the derived risk scores.

4.1.3 Methodological approach

Definition of the predictor sets

The predictor candidates of the present study comprised factors that are reported or hypothesized as being associated with changes in body weight and/or WC. With the objective to derive simple and inexpensive public health tools, only those predictor

candidates that are easily obtainable (e.g., by means of a questionnaire) were taken into account. In total, 22 predictor candidates were identified in the provided database; they determined the predictor set of the maximum models for SWG and SWCG. The approach of taking all identified predictor candidates into a model has been suggested to avoid bias associated with the procedure of predictor selection; such bias is frequently paralleled with over-adaptation to the model's derivation sample (178, 183). As not all included cohort studies assessed the full range of the identified predictor candidates, the maximum models of the cohorts were slightly different and varying numbers of cohorts contributed to the meta-analytically combined risk associations. These differences might have introduced some heterogeneity in the predictive performances of the maximum models across the cohorts. To reduce heterogeneity, the minimum models were defined. These models comprised those 17 predictors which were commonly assessed across the cohorts and each cohort contributed to the meta-analytically combined risk associations. In addition, the predictor range of the homogeneous minimum models corresponded closely to the predictor range of the European risk score (17). To examine whether the predictor range can be reduced further without considerable loss of predictive performances, and owing to the criticism that inclusion of too many predictors is to avoid when the number of cases is limited (178, 183), a selection procedure was applied on the basis of the homogeneous predictor set. The predictors were selected using RSF generating 1000 trees (=1000 bootstrap samples) separated by outcomes and cohorts. To derive models with sufficiently high numbers of outcome events per predictor, the relatively smaller number of participants in the cohorts of SHIP, KORA, and BGS98/DEGS in comparison with that in the German EPIC cohorts, was taken into account. Ranged by "minimal depth", the five most important predictors were included in the selection model for SWG and the seven most important predictors were included in the selection model for SWCG.

Validation of the risk scores

To evaluate the predictive performances of the derived risk scores, the measures of discrimination and calibration were used. Discrimination quantifies a model's ability to correctly assign individuals to one of two groups (e.g., cases and non-cases) (187); its most

common quantitative index is the aROC or c index (187, 188), which simply and directly addresses a model's discriminatory ability by its probabilistic interpretation (239). Nevertheless, several shortcomings have to be taken into account. For example, in spite of rough guidelines (191), it is not clearly definable how high should be a model's c index to justify its implementation into practice (240). Moreover, the c statistic is rank-based and the actual risk distribution is left out of consideration (189). It means that given the differences in rank are the same, marginal risk differences between two low-risk individuals (e.g., 1.0 % vs. 1.1 %) have the same impact on a model's c index, as marked risk differences between two individuals at moderate versus high risk (e.g., 5 % vs. 20 %) (189). This weakness is of prior relevance in prospective population-based cohort studies, since these studies are generally characterized by high proportions of individuals at low and very low risk as well as low proportions of individuals at high risk (189). Furthermore, measures of discrimination do not indicate whether a model's predicted risk agrees to the actually observed risk (200). To evaluate this model characteristic, the calibration of the derived risk scores was assessed by comparing their predicted risks with the actually observed incidences. Similar to the measure of discrimination, it is not clear how much miscalibration would entail that a model cannot be used (240). A commonly performed test to evaluate a model's calibration is the Hosmer-Lemeshow test (HLT); this statistical test compares a model's predicted risk with the observed incidence across the deciles of predicted risk, which can be visualized by calibration plots (191). However, the value of the HLT is called into question because its null hypothesis states that the model is well calibrated. Given a high p value, the test can consequently only tell that there is insufficient evidence of miscalibration (240). Furthermore, statistical tests highly depend on the sample size and consequently the HLT will always give a low p value when the sample size is large enough. Given the differing sample sizes across the cohorts, this can also be observed in the present study. Therefore, the visual evaluation by means of calibration plots is additionally attached for the assessment of a model's calibration.

As the present study aimed at meta-analytical derivation of German risk scores by including the data of the numerically most important German cohort studies, evaluation of the models' predictive performances, separated by cohorts, appeared most appropriate. To examine further the general applicability of the derived models, supplementary validation

techniques such as cross-validation could be applied (178). Such techniques give an insight into the model's generalizability with regard to a derivation-independent but comparable data set.

4.2 Results

4.2.1 Associations of predictors with substantial gain in weight and waist circumference

Risk scores aim to predict future events most accurately on the basis of a few factors that are preferably most strongly linked to the outcome of interest. Hence, causal links between single factors and the outcome are not necessarily required (180). It is well-established that the development of excessive body fat accumulation is a highly complex process. Multiple factors, which are potentially strongly interrelated, may ultimately affect the balance between energy intake and energy expenditure (7). While some factors included in the present risk scores may have causal relations with substantial gain in weight and/or WC, others may just represent the markers of an underlying construct of causal and non-causal associations. Against this background, the associations of predictors with SWG and SWCG have to be interpreted in the present study.

The maximum predictor set comprised 22 risk factors, including socio-demographic characteristics (age, sex, education), anthropometric factors (body weight, WC, body height, history of weight loss), lifestyle factors (sporting activity, alcohol consumption, smoking habits, watching TV, sleep duration), dietary factors (intake of fruits and vegetables, meat, whole grain bread, cake and cookies, soft drinks, fish, chocolate) as well as psychosocial and further health-related factors (living in partnership, life satisfaction, suffering from sleep disorders). Depending on individual risk factors, two to five cohorts contributed to the corresponding meta-analytically combined risk association in the maximum models for SWG and SWCG. To define a homogeneous predictor set across the cohorts, those risk factors which were not adequately assessed by all cohort studies were excluded: the excluded factors comprised history of weight loss, watching TV, sleep duration, living in partnership, and life satisfaction. Thus, the homogenous minimum models comprised 17 risk factors, and all five cohorts contributed to the risk associations. Likewise, all cohorts contributed to the risk associations of the risk factors in the selection models, which were selected by applying

a variable selection procedure and considering the number of cases across the cohorts. As a result, five risk factors remained in the model for SWG: age, sex, body weight, smoking habits, and consumption of soft drinks; and seven factors remained in the model for SWCG: age, sex, body weight, WC, body height, educational level, and consumption of meat.

In general, risk associations for most risk factors were quite consistent for SWG and SWCG; they were largely in line with the reported and hypothesized associations in the literature (see also Chapter 1.2.2). As reported in previous observational and/or experimental studies, most of the dietary factors included in the present study may reflect causal associations - for instance, higher intake of fiber-rich food such as whole grain bread and fruits and vegetables, was associated with lower risk of SWG and SWCG. Previous studies on these food groups mainly supported protective effects concerning the development of excessive body fat (84-88, 89, 93), whereby effects on satiety, satiation, and modulation of hormonal responses were proposed as common underlying physiological mechanisms (80-82, 90). Furthermore, higher consumption of sugar-rich food, such as soft drinks and chocolate, showed positive relations with SWG and SWCG. In accordance with this, previous prospective studies observed positive associations for higher chocolate consumption with long-term weight gain (108) and for increased consumption of sugar-sweetened beverages with greater risk of weight gain as well as obesity (116, 117). In contrast, inverse associations (in particular with SWG) were observed for intake of cake and cookies, which, being usually rich in sugar and fat potentially increases the risk of gaining body fat. Existing evidence investigating the underlying causal relation is limited and inconsistent (107, 109). Possibly, this observation resulted from selective under-reporting of socially undesirable foods like cake and cookies, by individuals that are prone to gain body fat. With regard to the consumption of protein sources that are frequently also rich in fat, such as meat (red, white, and processed meat) and fish, differing associations with the analyzed outcomes were revealed. While higher intake of meat was positively related with SWG and SWCG, higher fish consumption showed inverse relations with the same. Previous observational studies investigating the impact of various protein sources on excess body fat mainly found positive associations, which were particularly assigned to the intake of red and processed meat as well as poultry, instead of fish and other protein sources (97-101). These differing associations may in part be explained by food group-specific lipid profile: while meat and

meat products mainly contain saturated fatty acids and low density lipoprotein cholesterol, fish is a major source of long-chain n-3 poly-unsaturated fatty acids (PUFA). In clinical trials long-chain n-3 PUFAs show weight loss-specific effects (241, 242). Moreover, these associations may also reflect a general lower health consciousness as well as unhealthier dietary and lifestyle behaviors of individuals with higher meat intake, in comparison with individuals with higher fish intake. In addition to most dietary factors, lower risk for SWG and SWCG with higher sporting activity is based on a clear causal background (169, 170). The remaining factors included in the present study, however, rather represent the markers of the interplay between causal and non-causal associations, reflecting underlying lifestyle factors or behavioral patterns that affect the regulation of body fat accumulation. For instance, higher educational level was associated with lower risks relating to SWG and SWCG, while current smoking and time spent watching TV were associated with higher risks. Previous studies observed that highly educated individuals are more likely to integrate health-promoting dietary habits and to be more physical active (51). In contrast, particularly heavy smoking and prolonged time spent on screen-based activities are related to unhealthy dietary pattern (147, 148, 176, 177) and less exercise (145), potentially resulting in excess body fat accumulation. With regard to the psychosocial and further health-related factors, dissatisfaction with life and presence of sleep disorders were positively associated with the analyzed outcomes. Stress-related hormonal dysregulations going along with body fat accumulation promoting dietary and lifestyle behaviors are proposed as common underlying mechanisms (63, 69-74, 158-161). Living in partnership, as a marker of social support, was inversely related to SWG and SWCG in the present study, which is also supported by previous studies (58, 59). Furthermore, the observed positive associations with history of weight loss is also based on the underlying regulatory mechanism of body fat content, which favor rapid rebuilding of lost energy storages in response to weight reduction (243, 244, 245).

Interestingly, some factors showed opposing associations with SWG and SWCG. Increasing age, male sex, and higher body weight were associated with lower risks for SWG, but higher risks for SWCG. The observed associations suggest that substantial increases in body fat during aging specifically affects abdominal fat storage, which is also related to gradual hormonal changes (41-43). Additionally, they may reflect sex-specific patterns of body fat

gain. Surplus energy is predominantly stored in the gluteal and femoral area (referred to as gynoid or peripheral fat distribution) of women, while abdominal fat storage predominates in men (referred to as android or central fat distribution). Moreover, higher alcohol intake was associated with lower risk for SWG, but tended to increase risk for SWCG. The reduced risk of major gain in weight is possibly related to the level of education; in particular, higher alcohol consumption among women is related to higher educational level (134, 135), which is observed to go along with more health-promoting behavior (51), and presumably also with a more weight-conscious lifestyle. In contrast, various biological and psychosocial mechanisms support the relation between alcohol intake and gain in body fat (121-124): higher alcohol consumption is frequently hypothesized to be particularly related to higher risk of abdominal fat accumulation (246, 247), even though findings on this issue are inconsistent (131, 132, 134).

4.2.2 Predictive performances of the German risk scores

Discriminatory abilities of the meta-analytical maximum models for SWG and SWCG, which considered the complete range of available information, reached acceptable levels ($aROC \geq 0.70$) in most of the cohorts, - except from KORA for SWCG: 0.61 (0.58, 0.65). In the cohorts of EPIC-Heidelberg, SHIP, and BGS98/DEGS acceptable levels were achieved, although their maximum predictor sets were reduced by one (EPIC-Heidelberg: sleep duration) and two predictors (SHIP: history of weight loss, watching TV; BGS98/DEGS: living in partnership, watching TV), respectively. Use of cohort-specific predictor weights could improve the discriminatory accuracy for SWCG in KORA to 0.70 (0.67, 0.73), but the remaining discriminatory abilities remained nearly at the levels of the meta-analytical models. By applying the homogeneous predictor set of the minimum models, discriminatory abilities in most cohorts decreased slightly. Most pronounced decreases were observed in EPIC-Potsdam for SWG: 0.68 (0.67, 0.70) vs. 0.72 (0.71, 0.74) as well as in BGS98/DEGS for SWCG: 0.64 (0.59, 0.69) vs. 0.70 (0.65, 0.75). Use of cohort-specific predictor weights could improve the discriminatory abilities for SWCG in BGS98/DEGS to 0.71 (0.66, 0.75), but for SWG discriminatory abilities maintained the levels of the meta-analytical models - without exception. Further restriction to the predictor sets of the selection models remained discriminatory abilities almost unchanged.

These findings confirmed the initial hypothesis that the discriminatory ability of a Germany-wide risk score predicting SWG, which is based on easily obtainable information of several German cohorts, is not inferior to cohort-specific models. This was valid independent of the chosen predictor set and even if the meta-analytical risk estimates were based on less than five cohorts (when corresponding predictor information was not assessed). It implies that the meta-analytically derived risk estimates for SWG reflect the effects of single predictors across various German cohorts reasonably well. Moreover, taking account of the complete predictor range was able to slightly improve the discriminatory ability in comparison with the European risk score for SWG adapted to the German population. The discriminatory ability of the European risk score was comparable to that of the selection model for SWG. However, the selection model was based on less than half of the predictors included in the European risk score. Moreover, the findings of the present study showed that the discriminatory ability of a Germany-wide risk score predicting SWCG were more or less independent of the used predictor set but varied more across the cohorts, and additionally, the use of cohort-specific risk estimates could clearly improve discriminatory abilities in two cohorts. It implies that impaired performances depend rather on the meta-analytically derived risk estimates for SWCG than on the applied predictor range. The possible reasons for differences between meta-analytical and cohort-specific risk estimates are differences in data collection or coding of predictors that are specifically related to SWCG. Overall, for most of the cohorts, the discriminatory abilities of the risk scores for SWG and SWCG were lower than reported for most diabetes risk scores mainly reaching excellent levels of discrimination ($aROC \geq 0.80$) (248-250). Discriminatory abilities achieved in the present study are more comparable to those observed for risk scores predicting breast, lung or colorectal cancer, predominantly ranging from 0.60 to 0.75 (251-256). The most obvious reason for lower levels of discriminatory abilities is the omission of strong predictors. However, regarding the wide range of factors related to gain in body weight and/or WC which were considered in the present study, the discriminatory abilities of risk scores predicting SWG and SWCG on the basis of easily obtainable information appear to be generally limited. To explain this limitation, several aspects may be taken into account - for instance, the predictors were mainly based on self-reported information by the participants, which is prone to random error and self-reporting bias. In particular, socially undesirable behavioral factors (e.g., foods rich in sugar and fat, sedentary behavior) may suffer from under-

reporting, while others are positively regarded (e.g., sports activity) and may be more likely to be over-reported. In addition, only baseline information on the predictors was considered, which was not examined regarding changes over time.

With respect to quantification of absolute risks, similar patterns across the cohorts were observed for the risk scores for SWG (slight overestimations in the lower range and slight underestimation in the upper range of the score), which could be simply addressed by measures of re-calibration. Quantification of absolute risks by the risk scores for SWCG, however, showed more variation, which presumably owes to varying incidences of SWCG across the cohorts. Cohorts with higher incidences (EPIC-Potsdam, EPIC-Heidelberg, and KORA) generally underestimated absolute risks, while cohorts with lower incidences (SHIP, BGS98/DEGS) generally overestimated them.

Since only the internal validity of the derived risk scores was determined in the present study, which is prone to evaluate the models' performances too optimistically, a supplementary sensitivity was performed. Thereby, the cohorts SHIP, KORA, and BGS98/DEGS ($n = 4,430$) were used for meta-analytical derivation of the risk scores, which were subsequently externally validated in the EPIC-Potsdam cohort ($n = 14,633$) and in the EPIC-Heidelberg cohort ($n = 11,942$). As expected, the discriminatory accuracy across the derivation samples remained quite constant, while the discriminatory accuracy of the external validation samples was slightly lower in comparison with the models derived on the basis of all cohorts. One possible explanation for lower discriminatory ability is over-adaptation of the models to the derivation samples. Furthermore, the differences may be caused by varying data assessment methods and varying predictor coding between the derivation samples and external validation samples (178). The predictive performances of this sensitivity analysis are presented in detail in **App. Table 3-5** and **App. Figure 3**. As seen for the overall models, restriction to the homogenous predictor set of the minimum model only lessened the discriminatory ability for SWG in the EPIC-Potsdam cohort, while the remaining aROCs remained remarkably stable - even with further predictor restrictions. Thus, regardless of the predictor range, external validations revealed aROCs ≥ 0.67 for SWG and aROCs ≥ 0.66 for SWCG. Concerning SWG, the discriminatory ability of the selection models for SWG, derived on the basis of SHIP, KORA, and BGS98/DEGS achieved comparable levels to that of the European risk score adapted to the German population, as it was seen

for the overall models. Additionally, the quantification of absolute risks again showed same patterns for SWG and varied more for SWCG in both external validation samples. Hence, general conclusions of the present study remained virtually unchanged and support the application of the chosen approach. Inclusion of all available German cohorts for derivation of Germany-wide valid risk scores avoided the risk of over-adaptation to certain cohorts and reduced bias related to different assessment methods as well as predictor-coding across the cohorts, which are favorable conditions for allowing generalizability of the derived risk scores for the German population. However, the validity of the derived risk scores could be further evaluated by statistical standard techniques such as cross-validation (178).

4.3 Strengths and limitations

Among the major strengths of the present study are its prospective design, large sample size, and the inclusion of men and women taking part in the numerically most important population-based cohort studies in Germany. Involvement of these studies was assumed to allow generalizability of the derived risk scores for the German population. A further strength of this study is the availability of information on a wide range of easily obtainable risk factors for gain in body weight and/or WC across the included cohorts. Beyond that, the assessment of this information was conducted in comparable time-frames, with baseline assessments starting in the mid to late 90s. Additionally, harmonization efforts were performed to achieve same or at least comparable predictor scales across the cohorts. Furthermore, it was possible to apply uniform exclusion criteria to the cohorts, which allowed derivation of obesity-preventive tools by putting restriction on a meaningful age range and by avoiding the bias related to chronic health disorders. As a further strength, baseline anthropometry was measured by trained staff according to standardized procedures in all cohort studies. In KORA, SHIP, and DEGS, anthropometry was also measured at follow-up, while in the German EPIC cohorts follow-up anthropometry was based on self-reports by the participants. To account for potential misreporting (257-259), self-reports were corrected by specifically developed correction equations (221).

Among the limitations of the present study are the classification of the cases and non-cases for SWG and SWCG, assuming linear changes in body weight and WC over time. It is well-

known that these anthropometric measures rather tend to fluctuate (237, 238). As a consequence, individual non-linear changes in body weight and WC may have resulted in misclassifications of cases and non-cases as well as misestimations of the cases' survival times, and finally, may have impaired predictive performances of the risk scores. Moreover, the predictive performances may have been affected by the procedure of predictor assessment. The predictors were assessed by means of self-administered questionnaires as well as computer-assisted or face-to-face interviews, which may have induced self-reporting bias and random error (232, 260). In addition to the variation in methods of predictor assessment between the cohorts, varying requests of predictors across the cohorts may have induced the heterogeneity of relative risk estimates. A further potential limitation refers to the evaluation of the validity of the derived risk scores. The usual approach to derive generalizability of a country-specific risk score by means of external validation in a derivation-independent German study sample was not chosen; instead, the internal validity of the derived risk scores was assessed, separated by cohorts, and additional sensitivity analyses were performed.

4.4 Implications for public health

Risk scores are targeted at the reliable stratification of individuals into risk categories as well as at the adequate prediction of individuals' risks of the occurrence of future events based on limited information (9, 11, 13). By supporting health-care professionals and individuals in decision-making e.g., regarding initiating and withholding health-care interventions, risk scores are valuable public health instruments aiming at improved cost-effectiveness of care and better individual outcomes (180, 181). Till date, several risk scores are implemented in prevention programs or are publicly available as web-based instruments in the field of several non-communicable chronic diseases. Similarly, the present derivation of German risk scores predicting SWG and SWCG within five years intended to generate such tools for early, simple, and economic identification of high-risk individuals as well as accurate prediction of individuals' absolute risks. Provided that the derived risk scores are evaluated to be valid for the German population, they might find entrance into German obesity prevention programs or lay the ground for easy-to-use risk assessment charts of general and abdominal obesity.

The model characteristics of the derived risk scores, however, were just moderate. Taking account of the measures of discriminatory accuracy (sensitivity, specificity, PPV, NPV), in conjunction with the outcome-related impact of misclassifications, and in terms of false-positives and false-negatives, generally determines an optimal threshold to identify high-risk individuals (192). Owing to the higher impact of false-positives, potentially going along with substantially increased health-care expenditures through needless assignments to cost-intensive obesity prevention programs, a threshold attaining higher specificity might be more appropriate for the risk score of SWG. At the threshold of a score value of 600, for instance, specificity ranged from 62.7 % to 89.2 % across the cohorts. Furthermore, 4.3 % to 23.3 % of individuals would be captured, but only 8.9 % to 11.6 % of them will actually gain ≥ 10 % of baseline-body weight within five years. Owing to the stronger relationship between accumulated abdominal fat and metabolic risks, which potentially leads to serious chronic health disorders, and to avoid misguided reassurance, the choice of an appropriate threshold for the risk score of SWCG might be based on more balanced proportions of sensitivity and specificity (192). However, the measures of discriminatory accuracy for the risk score of SWCG considerably varied across the cohorts. For instance, at the threshold of a score value of 1,150, specificity ranged from 46.3 % to 89.2 % and sensitivity from 16.6 % to 72.8 %. Nevertheless, both derived risk scores were adequately able to exclude SWG (for score value < 600) and SWCG (for score value $< 1,150$) in the course of five years with probabilities above 94 % and 85 %, respectively.

With regard to the quantification of individuals' absolute risks, the risk score for SWG displayed superior performance to the German risk score for SWCG. While comparisons between the observed and predicted probabilities for SWG generally indicated similar patterns across the German cohorts, much more heterogeneity was found for the model for SWCG. Germany-wide valid risk scores with accurate quantification of individuals' absolute risks might have different applications. For instance, accurate risk quantification might be important in the frame of obesity prevention programs to improve compliance to dietary and/or lifestyle interventions by informing about the expected beneficial effect. In addition, they might be used for the conception of country-specific intervention trials because absolute risks reflect the number of incident cases, which in further consequence determines the power of the conceived trial (261).

4.5 Conclusion and outlook

Based on the pre-existing state of research, the present thesis supports the hypothesis that meta-analytically derived German risk scores predicting SWG, on the basis of easily obtainable information of the numerically most important German cohort studies, are generalizable in terms of the German population. This conclusion was drawn for all investigated predictor sets. In addition, the findings demonstrate that restrictions on the complete range of predictor candidates to a few but most important predictors slightly diminished predictive performances across the cohorts; nevertheless, even the meta-analytically derived German risk score with the most reduced predictor set revealed a comparable level of discriminatory abilities across the cohorts, as provided by the European risk score for SWG applied to the German EPIC cohorts - although it was based on less than half of the predictors. Regarding the prediction of SWCG, the present thesis is the first study that reported on such a tool among adults. In comparison with the derived risk scores for SWG, the discriminatory abilities of the scores for SWCG were generally lower but still reached acceptable levels in most of the cohorts. Additionally, the discriminatory abilities were largely independent of the range of predictors.

With regard to the wide range of easily obtainable predictors and the maximal achieved levels of discriminatory performance, however, the predictability of weight gain and specific waist gain based on such information may be limited in general. Indubitably, quantification of absolute risks would require additional efforts of re-calibration and internal validation of the derived risk scores could be evaluated further through statistical standard procedures, including bootstrapping and cross-validation. Future research may elucidate whether the assessment of dietary and lifestyle changes over time could improve the predictability. This issue is of particular interest in order to encourage individuals to adopt health-conscious, obesity-preventive behavioral pattern. Furthermore, future studies may focus on strategies that facilitate adoption of such behavioral pattern and support maintenance of a health-promoting dietary and lifestyle, by taking measures against the “obesogenic environment”.

Summary

Currently, more than 60 % of German adults are overweight or obese. During the last decade prevalence of overweight stagnated at a high level, while obesity prevalence grew further - especially among young adults. Excessive accumulation of body fat is a well-established risk factor for many chronic health disorders and premature death. Regarding the extent of metabolic risks, accumulation of abdominal (visceral) fat gained paramount scientific interest by exerting unique pathogenic effects. According to recent estimations for Germany, nearly 16.8 billion euros of annual health-care expenditures are attributed to consequences of overweight and obesity; whereby 82 % of direct costs are driven by metabolic disorders. In order to counteract further increase in obesity prevalence and to reduce the numbers of future body fat-associated health disorders, it is important to derive practicable and informative preventive instruments for public health in Germany. Therefore, the present thesis aimed to derive Germany-wide valid risk scores predicting substantial gain in weight and waist circumference (WC) in the course of five years on the basis of easily obtainable information.

Derivation of the German risk scores was based on the data of over 31,000 participants from five German cohort studies; these comprised the two German cohorts of the European Prospective Investigation into Cancer and Nutrition (EPIC) study, the Study of Health In Pomerania (SHIP)-cohort, a cohort of the research platform KORA (*Kooperative Gesundheitsforschung in der Region Augsburg*), and the nationwide cohort of the German National Health Interview and Examination Survey 1998 / National Health Interview and Examination Survey for Adults (BGS98/DEGS). Substantial weight gain (SWG) and substantial WC gain (SWCG) were specified as gaining ≥ 10 % of baseline weight and ≥ 2.5 cm of baseline residual waist (WC_{BMI}), respectively, during the follow-up. The predictor candidates comprised information on socio-demographic and anthropometric characteristics, dietary and lifestyle factors, as well as on psychosocial and further health-related conditions. To most precisely predict the five-year risks based on limited information and to avoid over-adaption to a pre-specified derivation sample, a three-step meta-analytical approach was applied: first, the risk scores were derived using maximal available sets of predictor candidates; next, the predictors were restricted to a homogeneous set of predictors across the cohorts; and finally, homogenous predictor sets were further reduced to the most

predictive factors by applying the selection procedure of random survival forest (RSF). To assign weights for each predictor, cohort-specific multivariable regression coefficients were pooled using random-effect meta-analyses. Based on the pooled coefficients, the risk scores were calculated as a linear combination of the included predictors. Across the steps, the risk scores were validated by the assessment of discrimination (area under the receiver operating characteristic curve, aROC) and calibration.

In the course of the follow-up period, 6,383 individuals gained ≥ 10 % of their baseline weight and 8,746 individuals increased their baseline $WC_{BMI} \geq 2.5$ cm. Incidence rates (per 10,000 person-years) across the cohorts ranged from 189 to 266 for SWG and from 206 to 356 for SWCG. Depending on cohorts 19 to 22 predictors were available for the maximum models. Of them, 17 were adequately assessed in all cohorts and determined the homogeneous predictor set of the minimum models. After predictor selection by RSF, and by considering the numbers of cases across the cohorts, five and seven predictors were included in the selection models for SWG and SWCG, respectively. Pursuant to the stepwise applied predictor sets for the risk scores predicting SWG, aROCs (95 % CI) slightly decreased from 0.73 (0.71, 0.76) to 0.71 (0.68, 0.75) and 0.70 (0.67, 0.73), while aROCs (95 % CI) for the risk scores predicting SWCG were relatively constant with values of 0.69 (0.66, 0.73), 0.68 (0.65, 0.72), and 0.68 (0.64, 0.71). For SWG and SWCG discriminatory abilities varied between the cohorts. Use of cohort-specific predictor weights left discrimination for SWG unchanged, while for SWCG some improvements of cohort-specific models were observed. Regarding the calibration of the risk scores, similar patterns were observed for SWG across all the cohorts, while more variability existed for SWCG.

The findings of this thesis support the generalizability of meta-analytically derived risk scores predicting SWG on the basis of easily obtainable information from the numerically most important German cohort studies. Moreover, discriminatory performance remains remarkable constant even after reduction to few but most important predictors. For SWCG, performances are inferior and appear more cohort-specific. With regard to the broad spectrum of considered factors, however, predictability of gain in body weight and WC based on easily obtainable information seems generally limited.

Zusammenfassung

Mehr als 60 % der deutschen Erwachsenen gelten derzeit als übergewichtig oder adipös. Während die Übergewichtsprävalenz in den vergangenen zehn Jahren auf hohem Niveau stagnierte, stieg die Adipositasprävalenz weiter an - insbesondere bei jungen Erwachsenen. Überschüssiges Körperfett ist ein etablierter Risikofaktor für viele chronische Gesundheitsstörungen und vorzeitige Todesfälle. Im Hinblick auf das Ausmaß metabolischer Risiken, hat insbesondere das abdominale (viszerale) Fett durch seine spezifischen pathogenen Effekte wissenschaftliche Aufmerksamkeit erlangt. Aktuelle Zahlen für Deutschland zufolge, sind etwa 16,8 Milliarden Euro der jährlichen Gesundheitsausgaben den Folgen von Übergewicht und Adipositas zuzuschreiben, wobei 82 % der direkten Kosten auf metabolische Erkrankungen zurückzuführen sind. Um einem weiteren Anstieg der Adipositasprävalenz entgegen zu wirken und die Zahl zukünftiger körperfett-assoziiierter Gesundheitsstörungen zu senken, ist die Erstellung einfacher und informativer Präventionsinstrumente für die öffentliche Gesundheit in Deutschland von großer Bedeutung. Die vorliegende Dissertation hat sich daher zum Ziel gesetzt, für Deutschland gültige Risikoscores zu erstellen, die starke Zunahmen von Gewicht und Taillenumfang im Laufe von fünf Jahren, auf Basis einfach zu erhebender Informationen, vorhersagen.

Die Erstellung der deutschen Risikoscores basierte auf den Daten von über 31.000 Teilnehmern von fünf deutschen Kohortenstudien; diese umfassten die beide deutschen Kohorten der *European Prospective Investigation into Cancer and Nutrition* (EPIC) -Studie, die SHIP (*Study of Health In Pomerania*)-Kohorte, eine Kohorte der Forschungsplattform KORA (Kooperative Gesundheitsforschung in der Region Augsburg) und die überregionale Kohorte des Bundes-Gesundheitssurvey 1998 / Deutsche Erwachsenen-Gesundheits-Studie (BGS98/DEGS). Starke Gewichtszunahmen (SGZ) bzw. starke Taillenumfangszunahmen (STUZ) entsprachen Zunahmen von ≥ 10 % des Körpergewicht bzw. ≥ 2.5 cm des residualen Taillenumfangs im Laufe der Beobachtungszeit. Die Prädiktoren umfassten soziodemographische und anthropometrische Charakteristika, Ernährungs- und Lebensstilfaktoren sowie psychosoziale und weitere gesundheitsbezogenen Faktoren. Um die Fünf-Jahresrisiken auf Basis von eingeschränkten Informationen möglichst präzise vorherzusagen und um eine Überanpassung an eine vordefinierte Kohorte zu vermeiden, wurde ein dreistufiger meta-analytischer Ansatz verfolgt: zunächst wurden die Risikoscores

unter Verwendung der maximal verfügbaren Prädiktorenzahl erstellt; daraufhin wurde die Prädiktorenzahl auf einen unter den Kohorten einheitlichen Prädiktorensatz eingeschränkt; schließlich wurden dieser Prädiktorensatz unter Anwendung des Variablenselektionsverfahrens *random survival forest* (RSF) auf die prädikativsten Faktoren weiter reduziert. Die Prädiktorengewichtungen basierten jeweils auf der meta-analytischen Zusammenfassung (mittels *random-effects meta-analysis*) der kohorten-spezifischen Regressionskoeffizienten. Auf Basis der so zusammengefassten Regressionskoeffizienten wurden die Risikoscores durch lineare Kombination der jeweils eingeschlossenen Prädiktoren berechnet. Auf jeder Stufe (bzw. mit jedem Prädiktorensatz) wurden die Risikoscores mittels Diskrimination (area under the receiver operating characteristic curve, aROC) und Kalibrierung validiert.

Im Verlauf der Beobachtungszeit nahmen 6.383 Personen ≥ 10 % ihres Körpergewichts zu und 8.746 Personen steigerten ihren residualen Taillenumfangs ≥ 2.5 cm. Die Inzidenzraten (pro 10.000 Personenjahre) für SGZ schwankten zwischen den Kohorten von 189 bis 266 und für STUZ von 206 bis 356. Je nach Kohorte standen 19 bis 22 Prädiktoren für die Maximummodelle zur Verfügung; davon wurden 17 in allen Kohorten erhoben und stellten somit den einheitlichen Prädiktorensatz der Minimummodelle dar. Nach der Prädiktorselektion mittels RSF und unter Berücksichtigung der Fallzahlen der Kohorten, wurden fünf und sieben Prädiktoren für die Selektionsmodelle von SGZ und STUZ verwendet. Dem dreistufigen Ansatz folgend, sank die aROC (95 % Konfidenzintervall) der Risikoscores für SGZ von 0.73 (0.71, 0.76) über 0.71 (0.68, 0.75) auf 0.70 (0.67, 0.73), während sie für die Risikoscores für STUZ mit 0.69 (0.66, 0.73), 0.68 (0.65, 0.72), und 0.68 (0.64, 0.71) relativ konstant blieb. Sowohl für SGZ als auch für STUZ schwankten die aROCs zwischen den Kohorten. Die Verwendung kohorten-spezifischer Prädiktorgewichtungen hatte für SGZ keine Auswirkungen, für STUZ konnten jedoch zum Teil Zunahmen der aROC beobachtet werden. Die Kalibrierungsplots der Risikoscores für SGZ zeigten zwischen den Kohorten einen sehr ähnlichen Verlauf, wohingegen die Kalibrierungsplots der Risikoscores für STUZ mehr variierten.

Die Ergebnisse der Dissertation stützen die Verallgemeinerbarkeit von meta-analytisch erstellten Risikoscores zur Vorhersage einer SGZ, auf Basis von einfach zu erhebenden Daten der zahlenmäßig bedeutendsten deutschen Kohortenstudien, für die deutsche Bevölkerung.

Die Diskriminationsfähigkeit der Risikoscores für SGZ und STUZ bleibt auch nach Einschränkung der Prädiktorenzahl bemerkenswert konstant. Im Hinblick auf das weite Spektrum berücksichtigter Faktoren, scheint die Vorhersagbarkeit starker Zunahmen von Gewicht und Taillenumfang durch einfach zu erhebenden Prädiktoren jedoch generell begrenzt zu sein.

References

- (1) Mensink, G.B.M.; Schienkiewitz, A.; Haftenberger, M.; Lapert, T.; Ziese, T.; Scheidt-Nave, C. Overweight and obesity in Germany - Results of the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsbl*, 2013, 56: 786-794.
- (2) Bellanger, T.M.; Bray, G.A. Obesity related morbidity and mortality. *J La State Med Soc*, 2005, 157 Spec No 1: S42-9.
- (3) World Health Organization. Global status report on noncommunicable diseases 2010, Geneva, 2011.
- (4) World Health Organization, Fact sheet: Obesity and Overweight. Available online: <http://www.who.int/mediacentre/factsheets/fs311/en/> (accessed on 7th September 2015).
- (5) Després, J.P. Health consequences of visceral obesity. *Ann Med*, 2001, 33(8): 534-41.
- (6) World Health Organization, Noncommunicable diseases, Noncommunicable disease country profiles 2014, Germany. Available online: http://www.who.int/nmh/countries/deu_en.pdf?ua=1 (accessed on 7th September 2015).
- (7) World Health Organization, Obesity: Preventing and Managing The Global Epidemic. Report of a WHO Consultation. Report No.: 894, Geneva, 2000.
- (8) Lehnert, T.; Streltchenia, P.; Konnopka, A.; Riedel-Heller, S.G.; König, H.H. Health burden and costs of obesity and overweight in Germany: an update. *Eur J Health Econ* 2014, pp 1-11.
- (9) Buijsse, B.; Simmons, R. K.; Griffin, S. J.; Schulze, M. B. Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. *Epidemiol Rev*, 2011, 33: 46-62.
- (10) Hippisley-Cox, J.; Coupland, C.; Vinogradova, Y.; Robson, J.; May, M.; Brindle, P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ*, 2007, 335 (7611): 136.
- (11) Cui, J. Overview of risk prediction models in cardiovascular disease research. *Ann Epidemiol*, 2009, 19 (10): 711-717.
- (12) Brindle, P.; Beswick, A.; Fahey, T.; Ebrahim, S. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. *Heart*, 2006, 92: 1752-1759.
- (13) Usher-Smith, J.A.; Emery, J.; Kassianos, A. P.; Walter, F. M. Risk Prediction Models for Melanoma: A Systematic Review. *Cancer Epidemiol Biomarkers Prev*, 2014, 23 (8): 1450-1463.
- (14) Gail, M.H.; Brinton, L. A.; Byar, D. P.; Corle, D. K.; Green, S.B.; Schairer, C.; Mulvihill, J.J. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*, 1989, 81 (24): 1879-1886.
- (15) Freedman, A.N.; Slattery, M. L.; Ballard-Barbash, R.; Willis, G.; Cann, B. J.; Pee, D., Gail, M. H.; Pfeiffer, R. M. Colorectal cancer risk prediction tool for white men and women without known susceptibility. *J Clin Oncol*, 2009, 27 (5): 686-693.

- (16) Cassidy, A.; Myles, J.P.; van Tongeren, M.; Page, R.D.; Liloglou, T.; Duffy, S.W.; Field, J.K. The LLP risk model: an individual risk prediction model for lung cancer. *Br J Cancer*, 2008, 98 (2): 270-276.
- (17) Steffen A.; Sørensen, T.I.; Knüppel, S.; Travier, N.; Sánchez, M.J.; Huerta, J.M.; Quirós, J.R.; Ardanaz, E.; Dorronsoro, M.; Teucher, B.; Li, K.; Bueno-de-Mesquita, H.B.; van der A D.; Mattiello, A.; Palli, D.; Tumino, R.; Krogh, V.; Vineis, P.; Trichopoulou, A.; Orfanos, P.; Trichopoulos, D.; Hedblad, B.; Wallström, P.; Overvad, K.; Halkjær, J.; Tjønneland, A.; Fagherazzi, G.; Dartois, L.; Crowe, F.; Khaw, K.T.; Wareham, N.; Middleton, L.; May, A.M.; Peeters, P.H.; Boeing, H. Development and validation of a risk score predicting substantial weight gain over 5 years in middle-aged European men and women. *PLoS One*, 2013, 8(7): e67429. doi:10.1371/journal.pone.0067429.
- (18) Harvard School of Public Health: Obesity prevention source - obesity definition. Available online: <http://www.hsph.harvard.edu/obesity-prevention-source/obesity-definition/> (accessed on 7th September 2015).
- (19) Garrow, J.; Webster, J. Quetelet's index (W/H²) as a measure of fatness. *International Journal of Obesity and Related Metabolic Disorders*, 1985, 9: 147-153.
- (20) Gallagher, D.; Visser, M.; Sepulveda, D.; Pierson, R.N.; Harris, T.; Heymsfield, S.B. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? *Am J Epidemiol*, 1996; 143: 228-3.
- (21) Sun, Q.; van Dam, R. M.; Spiegelman, D; Heymsfield, S.B.; Willett, H. C.; Hu, F. B. Comparison of dual-energy x-ray absorptiometric and anthropometric measures of adiposity in relation to adiposity-related biologic factors. *Am J Epidemiol*, 2010, 172 (12): 1442-1454.
- (22) Pouliot, M.C.; Després, J.P.; Lemieux, S.; Moorjani, S.; Bouchard, C.; Tremblay, A.; Nadeau, A.; Lupien, P.J. Waist circumference and abdominal sagittal diameter: best simple anthropometric indexes of abdominal visceral adipose tissue accumulation and related cardiovascular risk in men and women. *Am J Cardiol*, 1994, 73(7): 460-8.
- (23) Han, T.S.; van Leer, E.M.; Seidell, J.C.; Lean, M.E. Waist circumference action levels in the identification of cardiovascular risk factors: prevalence study in a random sample. *BMJ*, 1995, 311(7017): 1401-1405.
- (24) World Health Organization: Waist Circumference and Waist–Hip Ratio: Report of a WHO Expert Consultation, Geneva, 2008.
- (25) Lear, S.A.; Humphries, K.H.; Kohli, S.; Birmingham, C.L. The Use of BMI and Waist Circumference as Surrogates of Body Fat Differs by Ethnicity. *Obesity*, 2007, 15: 11.
- (26) Zhang, C.; Rexrode, K.M.; van Dam, R.M.; Li, T.Y.; Hu, F.B. Abdominal obesity and the risk of all-cause, cardiovascular, and cancer mortality: sixteen years of follow-up in US women. *Circulation*, 2008, 117: 1658-67.
- (27) Food and Agriculture Organization, Human energy requirements. Report of a Joint FAO/WHO/UNU Expert Consultation, Rome, 2001.

- (28) Edholm, O.G.; Fletcher, J.G.; Widdowson, E.M.; McCance, R.A. The energy expenditure and food intake of individual men. *Br J Nutr*, 1955, 9(3): 286-300.
- (29) Kennedy, G. The role of depot fat in the hypothalamic control of food intake in the rat. *Proc R Soc Lond B*; 1953, 140: 578-592.
- (30) Allison, D. B.; Kaprio, J.; Korkeila, M.; Koskenvuo, M.; Neale, M. C.; Hayakawa, K. The heritability of BMI among an international sample of monozygotic twins reared apart. *Int J Obes Relat Metab Disord*, 1996, 20, 501-506.
- (31) Luke A.; Guo X.; Rotimi C. N.; Adeyemo A. A.; Wilks R.; Forrester T.; Lowe W.; Comuzzie A.; Martin L. J.; Zhu X.; Rotimi, C.N.; Cooper R.S. Heritability of obesity-related traits among Nigerians, Jamaicans and US black people. *Int J Obes*, 2001, 25, 1034-1041.
- (32) Zhu X.; Cooper R. S.; Luke A.; Chen G.; Wu X.; Chakravarti A.; Weder A. A genome-wide scan for obesity in African Americans. *Diabetes*, 2002, 51, 541-544.
- (33) Anderson J. W.; Konz E. C.; Frederich R.C.; Wood C.L. Long-term weight-loss maintenance: a meta-analysis of US studies. *Am J Clin Nutr*, 2001, 74, 579-584.
- (34) Speakman J.R.; Levitsky, D.A.; Allison, D.B.; Bray, M.S.; de Castro J.M.; Clegg, D.J.; Clapham, J.C.; Dulloo, A.G.; Gruer, L.; Haw, S.; Hebebrand, J.; Hetherington, M.M.; Higgs, S.; Jebb, S.A.; Loos, R.J.; Luckman, S.; Luke, A.; Mohammed-Ali, V.; O'Rahilly, S.; Pereira, M.; Perusse, L.; Robinson, T.N.; Rolls, B.; Symonds, M.E.; Westerterp-Plantenga, M.S. Set points, settling points and some alternative models: theoretical options to understand how genes and environments combine to regulate body adiposity. *Dis Model Mech*, 2011, 4(6): 733-45.
- (35) De Castro J. M.; Plunkett S. A general model of intake regulation. *Neurosci Biobehav Rev*, 2002, 26, 581-595.
- (36) Keith, S.; Redden, D.T.; Katzmarzyk, P.; Boggiano, M.M.; Hanlon, E.C.; Benca, R.M.; Ruden, D.; Pietrobelli, A.; Barger, J.L.; Fontaine, K.R.; Wang, C.; Aronne, L.J.; Wright, S.M.; Baskin, M.; Dhurandhar, N.V.; Lijoi, M.C.; Grilo, C.M.; DeLuca, M.; Westfall, A.O.; Allison, D.B. Putative contributors to the secular increase in obesity: exploring the roads less traveled. *Int J Obesity (Lond)*, 2006, 30: 1585-94.
- (37) St-Onge, M.-P.; Gallagher, D. Body composition changes with aging: The cause or the result of alterations in metabolic rate and macronutrient oxidation? *Nutrition*, 2010, 26(2): 152-155.
- (38) Bosy-Westphal, A.; Eichhorn, C.; Kutner, D.; Illner, K.; Heller, M.; Müller, M.J. The age-related decline in resting energy expenditure in humans is due to the loss of fat-free mass and to alterations in its metabolically active components. *J Nutr*, 2003, 133: 2356-2362.
- (39) Laaksonen, D.E.; Niskanen, L.; Punnonen, K.; Nyyssönen, K.; Tuomainen, T.P.; Valkonen, V.P.; Salonen, R.; Salonen, J.T. Testosterone and sex hormone-binding globulin predict the metabolic syndrome and diabetes in middle-aged men. *Diabetes Care*, 2004, 27(5): 1036-41.

- (40) Dhindsa, S.; Miller, M.G.; McWhirter, C.L.; Mager, D.E.; Ghanim, H.; Chaudhuri, A.; Dandona, P. Testosterone concentrations in diabetic and nondiabetic obese men. *Diabetes Care*, 2010, 33(6): 1186-92.
- (41) Tsai, E.C.; Boyko, E.J.; Leonetti, D.L.; Fujimoto, W.Y. Low serum testosterone level as a predictor of increased visceral fat in Japanese-American men. *Int J Obes Relat Metab Disord*, 2000, 24(4): 485-91.
- (42) Lovejoy, J.C.; Champagne, C.M.; De Jonge, L.; Xie, H.; Smith, S.R. Increased visceral fat and decreased energy expenditure during the menopausal transition. *Int J Obesity*, 2008, 32 (6): 949-958.
- (43) Lovejoy, J.C. Weight gain in women at midlife: the influence of menopause. *Obesity Management*, 2009, 5 (2): 52-56.
- (44) Grossman, M.; Kaestner R. Effects of Education on Health, J.R. Behrman and N. Stacey (eds.), *The Social Benefits of Education*, University of Michigan Press, 1997, Ann Arbor, MI, pp. 69-123.
- (45) Webbink, D.; Martin, N.G.; Visscher, P.M. Does education reduce the probability of being overweight? *J Health Econ*, 2010, 29(1): 29-38.
- (46) Yoon, Y.S.; Oh, S.W.; Park, H.S. Socio-Economic Status in Relation to Obesity and Abdominal Obesity in Korean Adults: A Focus on Sex Differences, *Obesity*, 2006, Vol. 14(5), pp. 909-19.
- (47) Kuntz, B.; Lampert, T. Socioeconomic Factors and Obesity. *Dtsch Arztebl Int*, 2010, 107(30): 517–522.
- (48) Arendt, J.N. Does education cause better health? A panel data analysis using school reforms for identification. *Economics of Education Review*, 2005, 24, 149-160.
- (49) Kenkel, D.S.; Lillard, D.R.; Mathios, A.D. The Roles of High School Completion and GED Receipt in Smoking and Obesity, NBER Working Papers 11990, 2006. Available online: www.nber.org/papers/w11990 (accessed on 7th September 2015).
- (50) Lindeboom, M.; Llena-Nozal, A.; van der Klaauw, B. Parental education and child health: Evidence from a schooling reform. Free University Amsterdam, 2007. Available online: <http://ftp.iza.org/dp2516.pdf> (accessed on 7th September 2015).
- (51) Cutler, D.; Lleras-Muney, A. Education and Health: Evaluating Theories and Evidence, NBER Working Paper 12352, 2006. Available online: www.nber.org/papers/w12352 (accessed on 7th September 2015).
- (52) Webbink, D.; Martin, N.G.; Visscher, P.M. Does Education Reduce the Probability of Being Overweight?, CPB Discussion Papers 102, CPB Netherlands Bureau for Economic Policy Analysis, 2008.
- (53) Aitsi-Selmi, A.; Batty, G.D.; Barbieri, M.A.; Silva, A.A.; Cardoso, V.C.; Goldani, M.Z.; Marmot, M.G.; Bettiol, H. Childhood socioeconomic position, adult socioeconomic position and social mobility in relation to markers of adiposity in early adulthood: evidence of differential effects by gender in the 1978/79 Ribeirao Preto cohort study. *Int J Obes (Lond)*, 2013, 37(3): 439-47.

- (54) Hajian-Tilaki, K.O.; Heidari, B. Association of educational level with risk of obesity and abdominal obesity in Iranian adults. *J Public Health (Oxf)*, 2010, 32(2): 202-9.
- (55) Cohen, S.; Syme, S. *Social Support and Health*. Orlando; FL: Academic Press, 1984.
- (56) Anderson, D.; Deshaies, G.; Jobin, J. Social support, social networks and coronary artery disease rehabilitation: a review. *Can J Cardiol*, 1996, 12: 739-744.
- (57) Berkman, L.F. The role of social relations in health promotion. *Psychosom Med*, 1995, 57: 245-254.
- (58) Elfhag, K.; Rössner S. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obes Rev*, 2005, 6(1): 67-85.
- (59) Verheijden, M.W.; Bakx, J.C.; van Weel, C.; Koelen, M.A.; van Staveren, W.A. Role of social support in lifestyle-focused weight management interventions. *Eur J Clin Nutr*, 2005, 59 Suppl 1: S179-86.
- (60) Lallukka, T.; Laaksonen, M.; Martikainen, P.; Sarlio-Lahteenkorva, S.; Lahelma, E. Psychosocial working conditions and weight gain among employees. *Int J Obes (London)*, 2005, 29: 909-915.
- (61) Ali, S.M.; Lindstrom, M. Socioeconomic, psychosocial, behavioural, and psychological determinants of BMI among young women: differing patterns for underweight and overweight/obesity. *Eur J Public Health*, 2006, 16: 325-331.
- (62) Rääkkönen, K.; Matthews, K.A.; Kuller, L.H. Anthropometric and psychosocial determinants of visceral obesity in healthy postmenopausal women. *Int J Obes Relat Metab Disord*, 1999, 23(8): 775-782.
- (63) Laitinen, J.; Näyhä, S.; Kujala, V. Body mass index and weight change from adolescence into adulthood, waist-hip ratio and perceived work ability among young adults. *Int J Obes*, 2005, 29: 697-702.
- (64) Diener, E. Subjective well-being. *Psychological bulletin*, 1984, 95: 542-575.
- (65) Mehmood, T.; Shaukat, M. Life satisfaction and psychological well-being among young adult female university students. *International Journal of Liberal Arts and Social Science*, 2014, 2(5): 143-153.
- (66) Kuppens, P.; Realo, A.; Diener, E. The role of positive and negative emotions in life satisfaction judgment across nations. *Journal of personality and Social Psychology*, 2008, 95(1): 66-75.
- (67) Beutell, M.E. Glaesmer, H.; Decker, O.; Fischbeck, S.; Brahler, E. Life satisfaction, distress and resiliency across the life span of women. *Menopause*, 2009, 16(6): 1132-1138.
- (68) Guney, S. Life satisfaction of university students in Turkey. 1st world positive psychology conference, Pennsylvania, USA, 2009.
- (69) Bjorntorp, P. Do stress reactions cause abdominal obesity and comorbidities? *Obes Rev*, 2001, 2: 73-86.

- (70) Dallman, M.F.; Pecoraro, N.C.; La Fleur, S.E.; Warne, J.P.; Ginsberg, A.B.; Akana, S.F.; Laugero, K.C.; Houshyar, H.; Strack, A.M.; Bhatnagar, S.; Bell, M.E. Glucocorticoids, chronic stress, and obesity. *Prog Brain Res*, 2006, 153: 75-105.
- (71) Kral, J.G. The pathogenesis of obesity: stress and the brain-gut axis. *Surg Obes Relat Dis*, 2005, 1: 25-34.
- (72) Steinman, L.; Conlon, P.; Maki, R.; Foster, A. The intricate interplay among body weight, stress, and the immune response to friend or foe, *J Clin Invest*, 2003, 111: 183-185.
- (73) Drapeau, V.; Therrien, F.; Richard, D.; Tremblay, A. Is visceral obesity a physiological adaptation to stress? *Panminerva Med*, 2003, 45: 189-195.
- (74) Gluck, M.E.; Geliebter, A.; Lorence, M. Cortisol stress response is positively correlated with central obesity in obese women with binge eating disorder (BED) before and after cognitive-behavioral treatment, *Ann NY Acad Sci*, 2004, 1032: 202-207.
- (75) Swinburn, B.; Egger G. The runaway weight gain train: too many accelerators, not enough brakes. *BMJ*, 2004, 329: 736-769.
- (76) Matthiessen, J.; Fagt, S.; Biloft-Jensen, A.; Beck A.M.; Ovesen, L. Size makes a difference. *Public Health Nutr*, 2003, 6: 65-72.
- (77) McAllister, E.J.; Dhurandhar, N.V.; Keith, S.W.; Aronne, L.J.; Barger, J.; Baskin, M.; Benca, R.M.; Biggio, J.; Boggiano, M.M.; Eisenmann, J.C.; Eloheid, M.; Fontaine, K.R.; Gluckman, P.; Hanlon, E.C.; Katzmarzyk, P.; Pietrobelli, A.; Redden, D.T.; Ruden, D.M.; Wang, C.; Waterland, R.A.; Wright, S.M.; Allison, D.B. Ten putative contributors to the obesity epidemic. *Crit Rev Food Sci Nutr*, 2009, 49: 868-913.
- (78) Lanningham-Foster L.; Nysse, L.J.; Levine, J.A. Labor saved, calories lost: the energetic impact of domestic labor-saving devices. *Obes Res*, 2003, 11(10): 1178-1181.
- (79) Bell, A.C.; Ge, K.; Popkin, B.M. The road to obesity or the path to prevention: motorized transportation and obesity in China. *Obes Res*, 2002, 10(4): 277-83.
- (80) Ludwig, D.S.; Pereira, M.A.; Kroenke, C.H.; Hilner, J.E.; Van Horn, L.; Slattery, M.L.; Jacobs, D.R. Jr. Dietary fiber, weight gain, and cardiovascular disease risk factors in young adults. *JAMA*, 1999, 282(16): 1539-46.
- (81) Howarth, N.C.; Saltzman, E.; Roberts, S.B. Dietary fiber and weight regulation. *Nutr Rev*, 2001, 59: 129-139.
- (82) Koh-Banerjee, P.; Rimm, E.B. Whole grain consumption and weight gain: a review of the epidemiological evidence, potential mechanisms and opportunities for future research. *Proc Nutr Soc*, 2003, 62(1): 25-9.
- (83) Rolls, B.J.; Ello-Martin, J.A.; Tohill, B.C. What can intervention studies tell us about the relationship between fruit and vegetable consumption and weight management? *Nutr Rev*, 2004, 62: 1-17.

- (84) Ledoux, T.A.; Hingle, M.D.; Baranowski, T. Relationship of fruit and vegetable intake with adiposity: a systematic review. *Obes Rev*, 2011, 12(5), e143-e150.
- (85) He, K.; Hu, F.B.; Colditz, G.A.; Manson, J.E.; Willett, W.C.; Liu, S. Changes in intake of fruits and vegetables in relation to risk of obesity and weight gain among middle-aged women. *Int J Obes Relat Metab Disord*, 2004, 28(12): 1569-74.
- (86) Kahn, H.S.; Tatham, L.M.; Rodriguez, C.; Calle, E.E.; Thun, M.J.; Heath, C.W. Jr. Stable behaviors associated with adults' 10-year change in body mass index and likelihood of gain at the waist. *Am J Public Health*, 1997, 87(5): 747-54.
- (87) Buijsse, B.; Feskens, E.J.; Schulze, M.B.; Forouhi, N.G.; Wareham, N.J.; Sharp, S.; Palli, D.; Tognon, G.; Halkjaer, J.; Tjønneland, A.; Jakobsen, M.U.; Overvad, K.; van der A, D.L.; Du, H.; Sørensen, T.I.; Boeing, H. Fruit and vegetable intakes and subsequent changes in body weight in European populations: results from the project on Diet, Obesity, and Genes (DiOGenes). *Am J Clin Nutr*, 2009, 90(1): 202-9.
- (88) Alinia, S.; Hels, O.; Tetens, I. The potential association between fruit intake and body weight--a review. *Obes Rev*, 2009; 10(6): 639-47.
- (89) van de Vijver, L.P.; van den Bosch, L.M.; van den Brandt, P.A.; Goldbohm, R.A. Whole-grain consumption, dietary fibre intake and body mass index in the Netherlands cohort study. *Eur J Clin Nutr*, 2009, 63(1): 31-8.
- (90) Fardet A. New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? *Nutr Res Rev*, 2010, 23(1): 65-134.
- (91) Fung, T.T.; Rimm, E.B.; Spiegelman, D.; Rifai, N.; Tofler, G.H.; Willett, W.C.; Hu, F.B. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr*, 2001, 73(1): 61-7.
- (92) Jang, Y.; Lee, J.H.; Kim, O.Y.; Park, H.Y.; Lee, S.Y. Consumption of whole grain and legume powder reduces insulin demand, lipid peroxidation, and plasma homocysteine concentrations in patients with coronary artery disease: randomized controlled clinical trial. *Arterioscler Thromb Vasc Biol*, 2001, 21(12): 2065-71.
- (93) World Cancer Research Fund and American Institute for Cancer Research, Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective, Washington DC: AICR, 2007.
- (94) Halton, T.L.; Hu, F.B. The effects of high protein diets on thermogenesis, satiety and weight loss: a critical review. *J Am Coll Nutr*, 2004, 23(5): 373-85.
- (95) Westerterp-Plantenga, M.S.; Nieuwenhuizen, A.; Tome, D.; Soenen, S.; Westerterp, K.R. Dietary protein, weight loss, and weight maintenance. *Annu Rev Nutr*, 2009, 29: 21-41.
- (96) Summerbell, C.D.; Douthwaite, W.; Whittaker, V.; Ells, L.J.; Hillier, F.; Smith, S.; Kelly, S.; Edmunds, L.D.; Macdonald, I. The association between diet and physical activity and subsequent excess weight gain and obesity assessed at 5 years of age or older: a systematic review of the epidemiological evidence. *Int J Obes (Lond)*, 2009, 33(Suppl 3): S1-S92.

- (97) Halkjær, J.; Olsen, A.; Overvad, K.; Jakobsen, M.U.; Boeing, H.; Buijsse, B.; Palli, D.; Tognon, G.; Du, H.; van der A, D.L.; Forouhi, N.G.; Wareham, N.J.; Feskens, E.J.M.; Sørensen, T.I.A.; Tjønneland, A. Intake of total, animal and plant protein and subsequent changes in weight or waist circumference in European men and women: the Diogenes project. *Int J Obes*, 2010, 1-10.
- (98) Vergnaud, A.C.; Norat, T.; Romaguera, D.; Mouw, T.; May, A.M.; Travier, N.; Luan, J.; Wareham, N.; Slimani, N.; Rinaldi, S.; Couto, E.; Clavel-Chapelon, F.; Boutron-Ruault, M.C.; Cottet, V.; Palli, D.; Agnoli, C.; Panico, S.; Tumino, R.; Vineis, P.; Agudo, A.; Rodriguez, L.; Sanchez, M.J.; Amiano, P.; Barricarte, A.; Huerta, J.M.; Key, T.J.; Spencer, E.A.; Bueno-de-Mesquita, B.; Büchner, F.L.; Orfanos, P.; Naska, A.; Trichopoulou, A.; Rohrmann, S.; Hermann, S.; Boeing, H.; Buijsse, B.; Johansson, I.; Hellstrom, V.; Manjer, J.; Wirfält, E.; Jakobsen, M.U.; Overvad, K.; Tjønneland, A.; Halkjaer, J.; Lund, E.; Braaten, T.; Engeset, D.; Odysseos, A.; Riboli, E.; Peeters, P.H. Meat consumption and prospective weight change in participants of the EPIC-PANACEA study. *Am J Clin Nutr*, 2010, 92(2): 398-407.
- (99) Rouhani, M.H.; Salehi-Abargouei, A.; Surkan, P.J.; Azadbakht, L. Is there a relationship between red and processed meat intake and obesity? A systematic review and meta-analysis of observational studies. *Obes Rev*, 2014, 15: 740-748.
- (100) Jakobsen, M.U.; Dethlefsen, C.; Due, K.M.; May, A.M.; Romaguera, D.; Vergnaud, A.C.; Norat, T.; Sørensen, T.I.; Halkjær, J.; Tjønneland, A.; Boutron-Ruault, M.C.; Clavel-Chapelon, F.; Fagherazzi, G.; Teucher, B.; Kühn, T.; Bergmann, M.M.; Boeing, H.; Naska, A.; Orfanos, P.; Trichopoulou, A.; Palli, D.; Santucci De Magistris, M.; Sieri, S.; Bueno-de-Mesquita, H.B.; van der A, D.L.; Engeset, D.; Hjartåker, A.; Rodríguez, L.; Agudo, A.; Molina-Montes, E.; Huerta, J.M.; Barricarte, A.; Amiano, P.; Manjer, J.; Wirfält, E.; Hallmans, G.; Johansson, I.; Khaw, K.T.; Wareham, N.J.; Key, T.J.; Chajès, V.; Slimani, N.; Riboli, E.; Peeters, P.H.; Overvad, K. Fish consumption and subsequent change in body weight in European women and men. *Br J Nutr*, 2013, 109(2): 353-62.
- (101) Rosell, M.; Håkansson, N.N.; Wolk, A. Association between dairy food consumption and weight change over 9 y in 19 352 perimenopausal women. *Am J Clin Nutr*, 2006, 84(6): 1481-1488.
- (102) Yamashita, T.; Sasahara, T.; Pomeroy, S.E.; Collier, G.; Nestel, P.J. Arterial compliance, blood pressure, plasma leptin, and plasma lipids in women are improved with weight reduction equally with a meat-based diet and a plant-based diet. *Metabolism*, 1998, 47(11): 1308-14.
- (103) Leslie, W.S.; Lean, M.E.; Baillie, H.M.; Hankey, C.R. Weight management: a comparison of existing dietary approaches in a work-site setting. *Int J Obes Relat Metab Disord*, 2002, 26(11): 1469-75.
- (104) Mahon, A.K.; Flynn, M.G.; Stewart, L.K.; McFarlin, B.K.; Iglay, H.B.; Mattes, R.D.; Lyle, R.M.; Considine, R.V.; Campbell, W.W. Protein intake during energy restriction: effects on body composition and markers of metabolic and cardiovascular health in postmenopausal women. *J Am Coll Nutr*, 2007, 26(2): 182-9.
- (105) Golomb, B.A.; Koperski, S.; White, H.L. Association between more frequent chocolate consumption and lower body mass index. *Arch Intern Med*, 2012, 172(6): 519-21.

- (106) O'Neil, C.E.; Fulgoni, V.L. 3rd; Nicklas, T.A. Candy consumption was not associated with body weight measures, risk factors for cardiovascular disease, or metabolic syndrome in US adults: NHANES 1999–2004. *Nutr Res*, 2011, 31(2): 122-130.
- (107) Parker, D.R.; Gonzalez, S.; Derby, C.A.; Gans, K.M.; Lasater, T.M.; Carleton, R.A. Dietary factors in relation to weight change among men and women from two southeastern New England communities *Int J Obes*, 1997, 21: 103-109.
- (108) Greenberg, J.A.; Buijsse, B. Habitual chocolate consumption may increase body weight in a dose-response manner. *PLoS ONE*, 2013, 8(8): e70271. doi:10.1371/journal.pone.0070271.
- (109) Sammel, M.D.; Grisso, J.A.; Freeman, E.W.; Hollander, L.; Liu, L.; Liu, S.; Nelson, D.B.; Battistini, M. Weight gain among women in the late reproductive years. *Fam Pract*, 2003, 20(4): 401-9.
- (110) Malik, V.S.; Schulze, M.B.; Hu, F.B. Intake of sugar-sweetened beverages and weight gain: a systematic review. *Am J Clin Nutr*, 2006, 84(2): 274-288.
- (111) Hu, F.B. Resolved: there is sufficient scientific evidence that decreasing sugar-sweetened beverage consumption will reduce the prevalence of obesity and obesity-related diseases. *Obes Rev*, 2013, 14(8): 606-19.
- (112) Liebman, M.; Pelican, S.; Moore, S.A.; Holmes, B.; Wardlaw, M.K.; Melcher, L.M.; Liddil, A.C.; Paul, L.C.; Dunnagan, T.; Haynes, G.W. Dietary intake, eating behavior, and physical activity-related determinants of high body mass index in rural communities in Wyoming, Montana, and Idaho. *Int J Obes Relat Metab Disord*, 2003, 27: 684-692.
- (113) French, S.A.; Jeffery, R.W.; Forster, J.L.; McGovern, P.G.; Kelder, S.H.; Baxter, J.E. Predictors of weight change over two years among a population of working adults: the Healthy Worker Project. *Int J Obes Relat Metab Disord*, 1994, 18(3): 145-54.
- (114) Giammattei, J.; Blix, G.; Marshak, H.H.; Wollitzer, A.O.; Pettitt, D.J. Television watching and soft drink consumption: associations with obesity in 11- to 13-year-old schoolchildren. *Arch Pediatr Adolesc Med*, 2003, 157(9): 882-6.
- (115) Gillis, L.J.; Bar-Or, O. Food away from home, sugar-sweetened drink consumption and juvenile obesity. *J Am Coll Nutr*, 2003, 22(6): 539-45.
- (116) Schulze, M.B.; Manson, J.E.; Ludwig, D.S.; Colditz, G.A.; Stampfer, M.J.; Willett, W.C.; Hu, F.B. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *JAMA*, 2004, 292: 927-934.
- (117) Ludwig, D.S.; Peterson, K.E.; Gortmaker, S.L. Relation between consumption of sugar-sweetened drinks and childhood obesity: a prospective, observational analysis. *Lancet*, 2001, 357(9255): 505-8.
- (118) DiMeglio, D.P.; Mattes, R.D. Liquid versus solid carbohydrate: effects on food intake and body weight. *Int J Obes Relat Metab Disord*, 2000, 24: 794-800.

- (119) Tordoff, M.G.; Alleva, A.M. Effect of drinking soda sweetened with aspartame or high-fructose corn syrup on food intake and body weight. *Am J Clin Nutr*, 1990, 51(6): 963-9.
- (120) Yeomans, M.R. Alcohol, appetite and energy balance: is alcohol intake a risk factor for obesity? *Physiol Behav*, 2010, 100(1): 82-9.
- (121) Yki-Järvinen, H.; Koivisto, V.A.; Ylikahri, R.; Taskinen, M.R. Acute effects of ethanol and acetate on glucose kinetics in normal subjects. *Am J Physiol*, 1988, 254: E175-E180.
- (122) Siler, S.Q.; Neese, R.A.; Chrstinansen M.P.; Hellerstein, M.K. The inhibition of gluconeogenesis following alcohol in humans. *Am J Physiol*, 1998, 275: E897-907.
- (123) DeCastro, J.M.; Orozco, S. Moderate alcohol intake and spontaneous eating pattern of humans: evidence of unregulated supplementation. *Am J Clin Nutr*, 1990, 52: 246-253.
- (124) Forsander, O.A. Hypothesis: factors involved in the mechanisms regulating food intake affect alcohol consumption. *Alcohol Alcohol*, 1994, 29(5): 503-12.
- (125) Yeomans, M.R.; Caton, S.; Hetherington, M.M. Alcohol and food intake. *Curr Opin Clin Nutr Metab Care*, 2003, 6(6): 639-44.
- (126) Flechtner-Mors, M.; Biesalski, H.K.; Jenkinson, C.P.; Adler, G.; Ditschuneit, H.H. Effects of moderate consumption of white wine on weight loss in overweight and obese subjects. *Int J Obes Relat Metab Disord*, 2004, 28(11): 1420-6.
- (127) Suter, P.M.; Jéquier, E.; Schutz, Y. Thermogenic effect of ethanol. *FASEB J*, 1993, 7: A725. (x) Wareham, N.J.; van Sluijs, E.M.; Ekelund, U. Physical activity and obesity prevention: a review of the current evidence. *Proc Nutr Soc*, 2005, 64: 229-247.
- (128) MacInnis, R.J.; Hodge, A.M.; Dixon, H.G.; Peeters, A.; Johnson, L.E.; English, D.R.; Giles, G.G. Predictors of increased body weight and waist circumference for middle-aged adults. *Public Health Nutr*, 2014, 17(5): 1087-97.
- (129) Rissanen, A.M.; Heliövaara, M.; Knekt, P.; Reunanen, A.; Aromaa, A. Determinants of weight gain and overweight in adult Finns. *Eur J Clin Nutr*, 1991, 45(9): 419-30.
- (130) Sayon-Orea, C.; Bes-Rastrollo, M.; Nuñez-Cordoba, J.M.; Basterra-Gortari, F.J.; Beunza, J.J.; Martinez-Gonzalez, M.A. Type of alcoholic beverage and incidence of overweight/obesity in a Mediterranean cohort: the SUN project. *Nutrition*, 2011, 27(7-8): 802-8.
- (131) Schütze, M.; Schulz, M.; Steffen, A.; Bergmann, M.M.; Kroke, A.; Lissner, L.; Boeing, H. Beer consumption and the 'beer belly': scientific basis or common belief? *Eur J Clin Nutr*, 2009, 63(9): 1143-9.
- (132) Bendsen, N.T.; Christensen, R.; Bartels, E.M.; Kok, F.J.; Sierksma, A.; Raben, A.; Astrup, A. Is beer consumption related to measures of abdominal and general obesity? A systematic review and meta-analysis. *Nutr Rev*, 2013, 71(2): 67-87.
- (133) Traversy, G.; Chaput, J.P. Alcohol Consumption and Obesity: An Update. *Curr Obes Rep*, 2015, 4(1): 122-130.

- (134) Bergmann, M.M.; Schütze, M.; Steffen, A.; Boeing, H.; Halkjaer, J.; Tjønneland, A.; Travier, N.; Agudo, A.; Slimani, N.; Rinaldi, S.; Norat, T.; Romaguera, D.; Rohrmann, S.; Kaaks, R.; Jakobsen, M.U.; Overvad, K.; Ekelund, U.; Spencer, E.A.; Rodríguez, L.; Sánchez, M.J.; Dorronsoro, M.; Barricarte, A.; Chirlaque, M.D.; Orfanos, P.; Naska, A.; Trichopoulou, A.; Palli, D.; Grioni, S.; Vineis, P.; Panico, S.; Tumino, R.; Riboli, E.; Wareham, N.J.; Bueno-de-Mesquita, B.; May, A.; Peeters, P.H. The association of lifetime alcohol use with measures of abdominal and general adiposity in a large-scale European cohort. *Eur J Clin Nutr*, 2011, 65(10): 1079-87.
- (135) Sieri, S.; Krogh, V.; Saieva, C.; Grobbee, D.E.; Bergmann, M.; Rohrmann, S.; Tjønneland, A.; Ferrari, P.; Chloptsios, Y.; Dilis, V.; Jenab, M.; Linseisen, J.; Wallström, P.; Johansson, I.; Chirlaque, M.D.; Sanchez, M.J.; Niravong, M.; Clavel-Chapelon, F.; Welch, A.A.; Allen, N.E.; Bueno-de-Mesquita, H.B.; van der Schouw, Y.T.; Sacerdote, C.; Panico, S.; Parr, C.L.; Braaten, T.; Olsen, A.; Jensen, M.K.; Bingham, S.; Riboli, E.; Slimani, N. Alcohol consumption patterns, diet and body weight in 10 European countries. *Eur J Clin Nutr*, 2009, 63 Suppl 4: S81-100.
- (136) Shimokata, H.; Muller, D.C.; Andres, R. Studies in the distribution of body fat. III. Effects of cigarette smoking. *JAMA*, 1989, 261(8): 1169-1173.
- (137) Flegal, K.M.; Troiano, R.P.; Pamuk, E.R.; Kuczmarski, R.J.; Campbell, S.M. The influence of smoking cessation on the prevalence of overweight in the United States. *N Engl J Med*, 1995; 333(18): 1165–1170.
- (138) Huot, I.; Paradis, G.; Ledoux, M. Quebec Heart Health Demonstration Project Research Group. Factors associated with overweight and obesity in Quebec adults. *Int J Obes Relat Metab Disord*, 2004, 28(6): 766-774.
- (139) Hofstetter, A.; Schutz, Y.; Jequier, E.; Wahren, J. Increased 24-hour energy expenditure in cigarette smokers. *N Engl J Med*, 1986, 314(2): 79-82.
- (140) Williamson, D.F.; Madans, J.; Anda, R.F.; Kleinman, J.C.; Giovino, G.A.; Byers, T. Smoking cessation and severity of weight gain in a national cohort. *N Engl J Med*, 1991, 324: 739-745.
- (141) Ward, K.D.; Klesges, R.C.; Vander Weg, M.W. Cessation of smoking and body weight. In: Björntop P, ed. *International textbook of obesity*. Chichester, United Kingdom: Wiley & Sons Ltd, 2001: 323-36.
- (142) Bamia, C.; Trichopoulou, A.; Lenas, D.; Trichopoulos, D. Tobacco smoking in relation to body fat mass and distribution in a general population sample. *Int J Obes Relat Metab Disord*, 2004, 28: 1091-1096.
- (143) John, U.; Hanke, M.; Rumpf, H.J.; Thyrian, J.R. Smoking status, cigarettes per day, and their relationship to overweight and obesity among former and current smokers in a national adult general population sample. *Int J Obes Relat Metab Disord*, 2005, 29: 1289-1294.
- (144) Chiolerio, A.; Jacot-Sadowski, I.; Faeh, D.; Paccaud, F.; Cornuz, J. Association of cigarettes daily smoked with obesity in a general European adult population. *Obes Res*, 2007, 15(5): 1311-1318.
- (145) Kvaavik, E.; Meyer, H.E.; Tverdal, A. Food habits, physical activity and body mass index in relation to smoking status in 40-42 year old Norwegian women and men. *Prev Med*, 2004, 38: 1-5.

- (146) Bien, T.H.; Burge, R. Smoking and drinking: a review of the literature. *Int J Addict*, 1990, 25: 1429-1454.
- (147) Serdula, M.K.; Byers, T.; Mokdad, A.H.; Simoes, E.; Mendlein, J.M.; Coates, R.J. The association between fruit and vegetable intake and chronic disease risk factors. *Epidemiology*, 1996, 7: 161-165.
- (148) Thompson, R.L.; Margetts, B.M.; Wood, D.A.; Jackson, A.A. Cigarette smoking and food and nutrient intakes in relation to coronary heart disease. *Nutr Res Rev*, 1992, 5(1): 131-52.
- (149) Canoy, D.; Wareham, N.; Luben, R.; Welch, A.; Bingham, S.; Day, N.; Khaw, K.T. Cigarette smoking and fat distribution in 21,828 British men and women: a population-based study. *Obes Res*, 2005, 13(8): 1466-75.
- (150) Kim, J.H.; Shim, K.W.; Yoon, Y.S.; Lee, S.Y.; Kim, S.S.; Oh, S.W. Cigarette smoking increases abdominal and visceral obesity but not overall fatness: an observational study. *PLoS One*, 2012, 7(9): e45815. doi: 10.1371/journal.pone.0045815.
- (151) Patel, S.R.; Hu, F.B. Short sleep duration and weight gain: a systematic review. *Obesity*, 2008, 16 (3): 643-653.
- (152) Patel, S.R. Reduced sleep as an obesity risk factor. *Obes Rev*, 2009, 10 Suppl 2: 61-8.
- (153) Bixler, E. Sleep and society: an epidemiological perspective. *Sleep Med*, 2009, 10: S3-9.
- (154) Gangwisch, J.E.; Malaspina, D.; Boden-Albala, B.; Heymsfield, S.B. Inadequate sleep as a risk factor for obesity: analyses of the NHANES I. *Sleep*, 2005, 28: 1289-96.
- (155) Wu, Y.; Zhai, L.; Zhang, D. Sleep duration and obesity among adults: a meta-analysis of prospective studies. *Sleep Med*, 2014, 15(12): 1456-62.
- (156) Marshall, N.S.; Glozier, N.; Grunstein, R.R. Is sleep duration related to obesity? A critical review of the epidemiological evidence. *Sleep Med Rev*, 2008, 12(4): 289-98.
- (157) Hasler, G.; Buysse, D.J.; Klaghofer, R.; Gamma, A.; Ajdacic, V.; Eich, D.; Rössler, W.; Angst, J. The association between short sleep duration and obesity in young adults: a 13-year prospective study. *Sleep*, 2004, 27(4): 661-666.
- (158) Chaput, J.P.; Despres, J.P.; Bouchard, C.; Tremblay, A. Short sleep duration is associated with reduced leptin levels and increased adiposity: results from the Quebec Family Study. *Obesity*, 2007, 15: 253-261.
- (159) Taheri, S.; Lin, L.; Austin, D.; Young, T.; Mignot, E. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. *PLoS Med*, 2004, 1: e62.
- (160) Spiegel, K.; Leproult, R.; Colecchia, E.F.; L'Hermite-Balériaux, M.; Nie, Z.; Copinschi, G.; Van Cauter, E. Adaptation of the 24-h growth hormone profile to a state of sleep debt. *Am J Physiol Regul Integr Comp Physiol*, 2000, 279: R874-R883.
- (161) Spiegel, K.; Leproult, R.; Van Cauter, E. Impact of sleep debt on metabolic and endocrine function. *Lancet*, 1999, 354: 1435-1439.

- (162) Sivak, M. Sleeping more as a way to lose weight. *Obes Rev*, 2006, 7: 295-296.
- (163) Sawka, M.N.; Gonzalez, R.R.; Pandolf, K.B. Effects of sleep deprivation on thermoregulation during exercise. *Am J Physiol*, 1984, 246: R72-R77.
- (164) Kolka, M.A.; Martin, B.J.; Elizondo, R.S. Exercise in a cold environment after sleep deprivation. *Eur J Appl Physiol Occup Physiol*, 1984, 53: 282-285.
- (165) Dinges, D.F.; Pack, F.; Williams, K.; Gillen, K.A.; Powell, J.W.; Ott, G.E.; Aptowicz, C.; Pack, A.I. Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep*, 1997, 20: 267-277.
- (166) Hu, F.B. Sleep deprivation and obesity. In: *Obesity Epidemiology*, Oxford University Press, 2008, 16: 320-341.
- (167) Ford, D.E.; Kamerow, D.B. Epidemiologic study of sleep disturbances and psychiatric disorders. An opportunity for prevention? *JAMA*, 1989, 262(11): 1479-84.
- (168) Vgontzas, A.N.; Bixler, E.O.; Lin, H.M.; Prolo, P.; Mastorakos, G.; Vela-Bueno, A.; Kales, A.; Chrousos, G.P. Chronic insomnia is associated with nyctohemeral activation of the hypothalamic-pituitary-adrenal axis: clinical implications. *J Clin Endocrinol Metab*, 2001, 86(8): 3787-94.
- (169) Wareham, N.J.; van Sluijs, E.M.; Ekelund, U. Physical activity and obesity prevention: a review of the current evidence. *Proc Nutr Soc*, 2005, 64: 229-247.
- (170) Wing, R.R.; Hill, J.O. Successful weight loss maintenance. *Annu Rev Nutr*, 2001, 21: 323-341.
- (171) Koh-Banerjee, P.; Chu, N.F.; Spiegelman, D.; Rosner, B.; Colditz, G.; Willett, W.; Rimm, E. Prospective study of the association of changes in dietary intake, physical activity, alcohol consumption, and smoking with 9-y gain in waist circumference among 16 587 US men. *Am J Clin Nutr*, 2003, 78(4): 719-27.
- (172) Sternfeld, B.; Wang, H.; Quesenberry, C.P. Jr; Abrams, B.; Everson-Rose, S.A.; Greendale, G.A.; Matthews, K.A.; Torrens, J.I.; Sowers, M. Physical activity and changes in weight and waist circumference in midlife women: findings from the Study of Women's Health Across the Nation. *Am J Epidemiol*, 2004, 160(9): 912-22.
- (173) Hu, F.B. Physical activity, sedentary behaviors, and obesity. In: *Obesity Epidemiology*, Oxford University Press, 2008, 15: 301-319.
- (174) Blair, S.N.; LaMonte, M.J.; Nichaman, M.Z. The evolution of physical activity recommendation: how much is enough? *Am J Clin Nutr*, 2004, 79: 913S-920S.
- (175) Thorp, A.A.; Owen, N.; Neuhaus, M.; Dunstan, D.W. Sedentary Behaviors and Subsequent Health Outcomes in Adults - A Systematic Review of Longitudinal Studies, 1996–2011. *Am J Prev Med*, 2011, 41(2): 207-215.
- (176) Bowman, S.A. Television-viewing characteristics of adults: correlations to eating practices and overweight and health status. *Prev Chronic Dis*, 2006, 3(2): A38.

- (177) Thomson, M.; Spence, J.C.; Raine, K.; Laing, L. The association of television viewing with snacking behavior and body weight of young adults. *Am J Health Promot*, 2008, 22(5): 329-335.
- (178) Steyerberg, E.W. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Rotterdam: Springer, 2009.
- (179) Noordzij, M.; Dekker, F.W.; Zoccali C.; Jager, K.J. Measures of disease frequency: prevalence and incidence. *Nephron Clin Pract*, 2010, 115: c17-c20.
- (180) Moons, K.G.M.; Kengne, A.P.; Woodward, M.; Royston, P.; Vergouwe, Y.; Altman, D.G.; Grobbee, D.E. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio) marker. *Heart*, 2012, doi: 10.1136/heartjnl-2011-301246.
- (181) Lloyd-Jones, D.M. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation*, 2010, 121: 1768-1777.
- (182) Tripepi, G.; Heinze, G.; Jager, K.J.; Stel, V.S.; Dekker, F.W.; Zoccali, C. Risk prediction models. *Nephrol Dial Transplant*, 2013, 28: 1975-1980.
- (183) Harrell, F. E. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer, 2001.
- (184) Harrell, F.E.; Lee, K.L.; Mark, D.B. Tutorial in biostatistics: multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 1996, 15: 361-387.
- (185) Ishwaran H.; Kogalur, U.B.; Blackstone, E.H.; Lauer, M.S. Random survival forests. *Ann Appl Stat*, 2008, 2(3): 841-860.
- (186) Hsieh, E.; Gorodeski, E.Z.; Blackstone, E.H.; Ishwaran, H.; Lauer, M.S. Identifying important risk factors for survival in patient with systolic heart failure using random survival forests. *Circ Cardiovasc Qual Outcomes*, 2011, 4(1): 39-45.
- (187) Pencina, M.J.; D'Agostino, R.B. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statist Med*, 2004, 23: 2109-2123.
- (188) Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 1982, 143(1): 29-36.
- (189) Cook, N. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, 2007, 115: 928-935.
- (190) The University of North Carolina at Chapel Hill, Lecture 22: ROC curves, Fall 2010. Available online: <http://www.unc.edu/courses/2010fall/ecol/563/001/docs/lectures/lecture22.htm> (accessed on 10th September 2015).
- (191) Hosmer, D.W.; Lemeshow, S. *Applied logistic regression*. 2nd ed. Wiley, 2000, pp 156-164.
- (192) Gordis, L. *Epidemiology*. 4th ed, Philadelphia: Saunders, Elsevier, 2009.

- (193) Parikh, R; Mathai, A.; Parikh, S.; Sekhar, G.C.; Thomas, R. Understanding and using sensitivity, specificity and predictive values. *Indian J Ophthalmol*, 2008, 56 (1): 45-50.
- (194) Youden, W.J. Index for rating diagnostic tests. *Cancer*, 1950, 3: 32-35.
- (195) Bewick, V.; Cheek, L.; Ball, J. Statistics review 13: receiver operating characteristic curves. *Crit Care*, 2004, 8: 508-512.
- (196) D'Agostino, R.B.; Griffith, J.L.; Schmidt, C.H.; Terrin, N. Measures for evaluating model performance. *Proceedings of the Biometrics Section. American Statistical Association, Biometrics Section: Alexandria, VA*, 1997; 253-258.
- (197) D'Agostino, R.B.; Nam, B.H. Evaluation of the performance of survival analysis models: discrimination and calibration measures. *Handbook of Statistics*, 2004, New York: Elsevier, 1-15.
- (198) May, S.; Hosmer, D.W. A simplified method of calculating an overall goodness-of-fit test for the Cox proportional hazards model. *Lifetime Data Anal*, 1998, 4: 109-120.
- (199) R Graphical Manual: Function for calibration plot and Hosmer-Lemeshow goodness of fit test. Available online: http://rgm3.lab.nig.ac.jp/RGM/R_rdfile?f=PredictABEL/man/plotCalibration.Rd&d=R_CC (accessed on 10th September 2015).
- (200) Diamond, G.A. What price perfection? Calibration and discrimination of clinical prediction models. *J Clin Epidemiol*, 1992, 45(1): 85-89.
- (201) Poirier, P.; Després, J.P. Waist circumference, visceral obesity, and cardiovascular risk. *J Cardiopulm Rehabil*, 2003, 23(3): 161-9.
- (202) Wajchenberg, B.L. Subcutaneous and visceral adipose tissue: their relation to the metabolic syndrome, *Endocr Rev*, 2000, 21(6): 697-738.
- (203) Klein, S. The case of visceral fat: argument for the defense. *J Clin Invest*, 2004, 113(11): 1530-2.
- (204) Goodpaster, B.H.; Krishnaswami, S.; Resnick, H.; Kelley, D.E.; Haggerty, C.; Harris, T.B.; Schwartz, A.V.; Kritchevsky, S.; Newman, A.B. Association between regional adipose tissue distribution and both type 2 diabetes and impaired glucose tolerance in elderly men and women. *Diabetes Care*, 2003, 26(2): 372-9.
- (205) Kanaya, A.M.; Harris, T.; Goodpaster, B.H.; Tylavsky, F.; Cummings, S.R. Adipocytokines attenuate the association between visceral adiposity and diabetes in older adults. *Diabetes Care*, 2004, 27(6): 1375-80.
- (206) Nielsen, S.; Guo, Z.; Johnson, C.M.; Hensrud, D.D.; Jensen, M.D. Splanchnic lipolysis in human obesity. *J Clin Invest*, 2004, 113(11): 1582-8.
- (207) Riboli, E.; Kaaks, R. The EPIC Project: rationale and study design. *European Prospective Investigation into Cancer and Nutrition, Int J Epidemiol*, 1997, 26 Suppl 1: S6-S14.
- (208) Riboli, E.; Hunt, K.J.; Slimani, N.; Ferrari, P.; Norat, T.; Fahey, M.; Charrondière, U.R.; Hémon, B.; Casagrande, C.; Vignat, J.; Overvad, K.; Tjønneland, A.; Clavel-Chapelon, F.; Thiébaud, A.; Wahrendorf,

- J.; Boeing, H.; Trichopoulos, D.; Trichopoulou, A.; Vineis, P.; Palli, D.; Bueno-De-Mesquita, H.B.; Peeters, P.H.; Lund, E.; Engeset, D.; González, C.A.; Barricarte, A.; Berglund, G.; Hallmans, G.; Day, N.E.; Key, T.J.; Kaaks, R.; Saracci, R. European Prospective Investigation into Cancer and Nutrition (EPIC): study population and data collection. *Public Health Nutr*, 2002, 5(6B): 1113-1124.
- (209) Boeing, H.; Korfmann, A.; Bergmann, M.M. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. *Ann Nutr Metab*, 1999, 43(4): 205-215.
- (210) Kroke, A.; Bergmann, M.M.; Lotze, G.; Jeckel, A.; Klipstein-Grobusch, K.; Boeing, H. Measures of quality control in the German component of the EPIC-study. European Prospective into Cancer and Nutrition. *Ann Nutr Metab*, 1999, 43: 216-224.
- (211) Klipstein-Grobusch, K.; Georg, T.; Boeing, H. Interviewer variability in anthropometric measurements and estimates of body composition. *Int J Epidemiol*, 1997, 26 Suppl 1: S174-180.
- (212) John, U.; Greiner, B.; Hensel, E.; Ludemann, J.; Piek, M.; Sauer, S.; Adam, C.; Born, G.; Alte, D.; Greiser, E.; Haertel, U.; Hense, H.W.; Haerting, J.; Willich, S.; Kessler, C. Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz Präventivmed*, 2001, 46: 186-194.
- (213) Holle, R.; Happich, M.; Lowel, H.; Wichmann, H.E. KORA -- a research platform for population based health research. *Gesundheitswesen*, 2005, 67 (Suppl 1): S9-25.
- (214) Chambless, L.; Cairns, V.; Herbold, M.; Döring, A.; Filipiak, B.; Schneller, H.; Viessmann, M.; Keil, U. MONICA-Projekt Region Augsburg, Survey sampling. München: GSF-Report 31/86, 1986.
- (215) Haerting, J.; Kluttig, A.; Greiser, K.H.; Nuding, S.; Werdan, K. A cohort study investigating risk factors for cardiovascular disease in an urban elderly East-German population (CARLA study). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, 2012, 55 (6-7): 795-800.
- (216) Greiser, K.H.; Kluttig, A.; Schumann, B.; Kors, J.A.; Swenne, C.A.; Kuss, O.; Werdan, K.; Haerting, J. Cardiovascular disease, risk factors and heart rate variability in the elderly general population: design and objectives of the CARdiovascular disease, Living and Ageing in Halle (CARLA) Study. *BMC Cardiovasc Disord*, 2005, 5: 33.
- (217) Krawczak, M.; Nikolaus, S.; von Eberstein, H.; Croucher, P.J.; El Mokhtari, N.E.; Schreiber, S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet*, 2006, 9: 55-61.
- (218) Nöthlings, U.; Krawczak, M. PopGen: Eine populationsbasierte Biobank mit Langzeitverfolgung der Kontrollkohorte. *Bundesgesundheitsbl*, 2012, 55: 831-835.
- (219) Bellach, B.M.; Knopf, H.; Thefeld, W. Der Bundes-Gesundheitssurvey 1997/98. *Gesundheitswesen*, 1998, 60 (Sonderheft 2): 59-68.
- (220) Gößwald, A.; Lange, M.; Dölle, R.; Hölling, H. Die erste Welle der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) – Gewinnung von Studienteilnehmenden, Durchführung der Feldarbeit und Qualitätsmanagement. *Bundesgesundheitsblatt*, 2013, 56 (5/6): 611-619.

- (221) Haftenberger, M.; Lahmann, P.H.; Panico, S.; Gonzalez, C.A.; Seidell, J.C.; Boeing, H.; Giurdanella, M.C.; Krogh, V.; Bueno-de-Mesquita, H.B.; Peeters, P.H.; Skeie, G.; Hjartaker, A.; Rodriguez, M.; Quirós, J.R.; Berglund, G.; Janlert, U.; Khaw, K.T.; Spencer, E.A.; Overvad, K.; Tjønneland, A.; Clavel-Chapelon, F.; Tehard, B.; Miller, A.B.; Klipstein-Grobusch, K.; Benetou, V.; Kiriazi, G.; Riboli, E.; Slimani, N. Overweight, obesity and fat distribution in 50-64-year old participants in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr*, 2002, 5(6B): 1147-1162.
- (222) Drewnowski A. Diet image: a new perspective on the food-frequency questionnaire. *Nutr Rev*, 2001, 59(11): 370-2.
- (223) Mahabir, S.; Baer, D.J.; Giffen, C.; Subar, A.; Campbell, W.; Hartman, T.J.; Clevidence, B.; Albanes, D.; Taylor, P.R. Calorie intake misreporting by diet record and food frequency questionnaire compared to doubly labeled water among postmenopausal women. *Eur J Clin Nutr*, 2006, 60, 561-565.
- (224) Cox, D.R. Regression models and life-tables. *J Roy Stat Soc B*, 1972, 34: 187-220.
- (225) Smith-Warner, S.A.; Spiegelman, D.; Ritz, J.; Albanes, D.; Beeson, W.L.; Bernstein, L.; Berrino, F.; van den Brandt, P.A.; Buring, J.E.; Cho, E.; Colditz, G.A.; Folsom, A.R.; Freudenheim, J.L.; Giovannucci, E.; Goldbohm, R.A.; Graham, S.; Harnack, L.; Horn-Ross, P.L.; Krogh, V.; Leitzmann, M.F.; McCullough, M.L.; Miller, A.B.; Rodriguez, C.; Rohan, T.E.; Schatzkin, A.; Shore, R.; Virtanen, M.; Willett, W.C.; Wolk, A.; Zeleniuch-Jacquotte, A.; Zhang, S.M.; Hunter, D.J. Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. *Am J Epidemiol*, 2006, 163 (11): 1053-64.
- (226) Borenstein, M.; Hedges, L.; Rothstein, H. Introduction to meta-analysis. 2007. Available online: <https://www.meta-analysis.com/downloads/Meta%20Analysis%20Fixed%20vs%20Random%20effects.pdf> (accessed on 4th September 2015).
- (227) Higgins, J.P.; Thompson, S.G.; Deeks, J. J; Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ*, 2003, 327(7414): 557–560.
- (228) Higgins, J.P.; Thompson, S.G. Quantifying heterogeneity in a meta-analysis. *Stat Med*, 2002, 21(11): 1539-58.
- (229) Kleinbaum, D.G.; Klein, M. *Survival Analysis - A Self-Learning Text*. 3rd ed, New York, Springer Science + Business Media, LLC, 2012.
- (230) Shim, J.-S.; Oh, K.; Kim, H.C. Dietary assessment methods in epidemiologic studies. *Epidemiol Health*, 2014, 36: e2014009.
- (231) Neuhouwer, M.L. Methodologies and tools for dietary intake assessment, In: *Handbook of Nutrition and Food*, 3rd ed, Boca Raton, CRC Press Taylor & Francis Group, LLC, 2014.
- (232) Thompson, F.E.; Subar, A.F. Dietary assessment methodology, 2001. Available online: http://appliedresearch.cancer.gov/diet/adi/thompson_subar_dietary_assessment_methodology.pdf (accessed on 4th September 2015).

- (233) Peters, T.M.; Shu, X.-O.; Moore, S.C.; Xiang, Y.B.; Yang, G.; Ekelund, U.; Liu, D.-K.; Tan, Y.-T.; Ji, B.-T.; Schatzkin, A.S.; Zheng, W.; Chow, W.H.; Matthews, C.E.; Leitzmann, M.F. Validity of a Physical Activity Questionnaire in Shanghai. *Med Sci Sports Exerc*, 2010, 42(12): 2222-2230.
- (234) Kaaks, R.; Slimani, N.; Riboli, E. Pilot phase studies on the accuracy of dietary intake measurements in the EPIC project: overall evaluation of result. *European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol*, 1997, 26 (Suppl 1): S26-36.
- (235) Margetts, B.M.; Pietinen, P. European Prospective Investigation into Cancer and Nutrition: validity studies on dietary assessment methods. *Int J Epidemiol*, 1997, 26 (Suppl 1): S1-5
- (236) Wong, S.L.; Shields, M.; Leatherdale, S.; Malaisson, E.; Hammond, D. Assessment of validity of self-reported smoking status. *Health Rep*, 2012, 23(1): 47-53.
- (237) Colditz, G.A.; Willett, W.C.; Stampfer, M.J.; London, S.J.; Segal, M.R.; Speizer, F.E. Patterns of weight change and their relation to diet in a cohort of healthy women. *Am J Clin Nutr*, 1990, 51(6): 1100-5.
- (238) Visscher, T.L.; Seidell, J.C. Time trends (1993-1997) and seasonal variation in body mass index and waist circumference in the Netherlands. *Int J Obes Relat Metab Disord*, 2004, 28(10): 1309-16.
- (239) Pencina, M.J.; D'Agostino, R.B. Sr; D'Agostino, R.B. Jr; Vasan, R.S. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*, 2008, 27(2): 157-172. discussion 207-212.
- (240) Vickers, A.J.; Cronin, A.M. Traditional statistical methods for evaluating prediction models are uninformative as to clinical value: towards a decision analytic framework. *Semin Oncol*, 2010, 37(1): 31-38.
- (241) Mori, T.A.; Bao, D.Q.; Burke, V.; Puddey, I.B.; Watts, G.F.; Beilin, L.J. Dietary fish as a major component of a weight-loss diet: effect on serum lipids, glucose, and insulin metabolism in overweight hypertensive subjects. *Am J Clin Nutr*, 1999, 70(5): 817-25.
- (242) Thorsdottir, I.; Tomasson, H.; Gunnarsdottir, I.; Gisladdottir, E.; Kiely, M.; Parra, M.D.; Bandarra, N.M.; Schaafsma, G.; Martínéz, J.A. Randomized trial of weight-loss-diets for young adults varying in fish and fish oil content. *Int J Obes (Lond)*, 2007, 31(10): 1560-6.
- (243) Hofbauer, K.G. Molecular pathways to obesity. *Int J Obes*, 2002, 26, Suppl 2, S18-S27.
- (244) Leibel, R.L.; Rosenbaum, M.; Hirsch, J. Changes in energy expenditure resulting from altered body weight. *New Engl J Med*, 1995, 332: 621-628.
- (245) Hofbauer, K.G.; Huppertz, C. Pharmacotherapy and evolution. *Trends Ecol Evol*, 2002, 17: 328-334.
- (246) Suter, P.M.; Maire, R.; Vetter, W. Alcohol consumption: a risk factor for abdominal fat accumulation in men. *Addict Biol*, 1997, 2: 101-103.

- (247) Colditz, G.A.; Giovannucci, E.; Rimm, E.B.; Stampfer, M.J.; Rosner, B.; Speizer, F.E.; Gordis, E.; Willet, W.C. Alcohol intake in relation to diet and obesity in women and in men. *Am J Clin Nutr*, 1991, 54: 49-55.
- (248) Schulze, M.B.; Hoffmann, K.; Boeing, H.; Linseisen, J.; Rohrmann, S.; Möhlig, M.; Pfeiffer, A.F.; Spranger, J.; Thamer, C.; Häring, H.U.; Fritsche, A.; Joost, H.G. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. *Diabetes Care*, 2007, 30(3): 510-5.
- (249) Lindström, J.; Tuomilehto, J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care*, 2003, 26(3): 725-31.
- (250) Wilson, P.W.; Meigs, J.B.; Sullivan, L.; Fox, C.S.; Nathan, D.M.; D'Agostino, R.B. Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med*, 2007, 167(10): 1068-74.
- (251) Chen, J.; Pee, D.; Ayyagari, R.; Graubard, B.; Schairer, C.; Byrne, C.; Benichou, J.; Gail, M.H. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J Natl Cancer Inst*, 2006, 98(17): 1215-26.
- (252) Gail, M.H. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst*, 2008, 100(14): 1037-41.
- (253) Bach, P.B.; Kattan, M.W.; Thornquist, M.D.; Kris, M.G.; Tate, R.C.; Barnett, M.J.; Hsieh, L.J.; Begg, C.B. Variations in lung cancer risk among smokers. *J Natl Cancer Inst*, 2003, 95(6): 470-8.
- (254) Spitz, M.R.; Hong, W.K.; Amos, C.I.; Wu, X.; Schabath, M.B.; Dong, Q.; Shete, S.; Etzel, C.J. A risk model for prediction of lung cancer. *J Natl Cancer Inst*, 2007, 99(9): 715-26.
- (255) Driver, J.A.; Gaziano, J.M.; Gelber, R.P.; Lee, I.M.; Buring, J.E.; Kurth, T. Development of a risk score for colorectal cancer in men. *Am J Med*, 2007, 120(3): 257-63.
- (256) Ma, E.; Sasazuki, S.; Iwasaki, M.; Sawada, N.; Inoue, M.; Shoichiro Tsugane; Japan Public Health Center-based Prospective Study Group. 10-Year risk of colorectal cancer: development and validation of a prediction model in middle-aged Japanese men. *Cancer Epidemiol*, 2010, 34(5): 534-41.
- (257) Spencer, E.A.; Appleby, P.N.; Davey, G.K.; Key, T.J. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr*, 2002, 5(4): 561-5.
- (258) Niedhammer, I.; Bugel, I.; Bonenfant, S.; Goldberg, M.; Leclerc, A. Validity of self-reported weight and height in the French GAZEL cohort. *Int J Obes Relat Metab Disord*, 2000, 24(9): 1111-1118.
- (259) Park, J.Y.; Mitrou, P.N.; Keogh, R.H.; Luben, R.N.; Wareham, N.J.; Khaw, K.T. Effects of body size and sociodemographic characteristics on differences between self-reported and measured anthropometric data in middle-aged men and women: the EPIC-Norfolk study. *Eur J Clin Nutr*, 2011, 65(3): 357-67.

- (260) DiPietro, L. Physical activity, body weight, and adiposity: an epidemiologic perspective. *Exerc Sport Sci Rev*, 1995, 23: 275-303.
- (261) Gail, M.H.; Costantino, J.P. Validating and improving models for projecting the absolute risk of breast cancer. *J Natl Cancer Inst*, 2001, 93(5): 334-5.

Appendix

App. Table 1. General characteristics of the study population for the minimum and selection models

	Study population				
	EPIC-Potsdam	EPIC-Heidelberg	SHIP	KORA	BGS98/DEGS
N	15,465	12,229	1,318	1,671	1,521
Duration of follow-up (y)	8.6 (0.9)	8.6 (0.7)	11.2 (0.8)	7.2 (0.3)	11.9 (1.0)
Socio-demography					
Age at baseline (y)	47.3 (8.0)	48.6 (7.4)	39.7 (10.8)	43.0 (10.5)	39.2 (10.6)
Men (%)	37.2	44.6	45.1	49.9	47.6
Education (%)					
no/primary school	12.9	22.0	13.6	5.0	6.8
secondary/professional school	45.5	41.6	73.1	74.9	73.7
university	41.6	36.5	13.3	20.1	19.5
Anthropometry					
Weight					
At baseline (kg)	68.7 (10.8)	69.7 (11.1)	71.5 (12.0)	71.2 (11.6)	70.7 (11.6)
At follow-up (kg)	72.4 (11.8)	72.8 (12.6)	75.8 (13.2)	73.3 (12.6)	74.4 (13.3)
Absolute change (kg)	3.7 (5.0)	3.1 (6.1)	4.3 (6.3)	2.1 (4.7)	3.7 (6.4)
Annual change (g/year)	429 (588)	359 (723)	388 (560)	299 (661)	310 (536)
Change (% of baseline weight)	5.5 (7.5)	4.5 (9.1)	6.4 (9.0)	3.1 (6.7)	5.4 (9.2)
Height					
At baseline (cm)	167.8 (8.5)	169.7 (8.6)	170.7 (8.9)	169.7 (9.3)	170.8 (9.1)
BMI					
At baseline (kg/m ²)	24.3 (2.6)	24.1 (2.7)	24.4 (2.8)	24.6 (2.7)	24.1 (2.7)
At follow-up (kg/m ²)	25.6 (3.1)	25.2 (3.4)	26.1 (3.4)	25.2 (3.0)	25.5 (3.4)
Obese at follow-up (%)	7.6	5.9	14.0	5.9	9.3
Waist					
At baseline (cm)	80.6 (9.8)	82.2 (10.1)	80.6 (10.1)	84.1 (9.7)	82.8 (10.0)
At follow-up (cm)	89.9 (10.4)	90.8 (12.2)	86.3 (11.0)	87.0 (10.7)	87.5 (11.2)
Absolute change (cm)	9.3 (5.6)	8.6 (8.7)	5.7 (6.5)	3.0 (5.6)	4.7 (7.0)
Annual change (mm/year)	10.9 (6.6)	10.1 (10.2)	5.0 (5.8)	4.2 (7.8)	3.9 (5.9)
Change (% of baseline waist)	11.9 (7.4)	10.8 (10.9)	7.3 (8.3)	3.7 (6.8)	5.9 (9.0)
Abdominal obese at follow-up (%)	32.7	32.3	18.8	18.9	22.2
Residual waist					
Absolute change (cm)					
75 th percentile	2.6	3.5	2.3	2.2	2.4
90 th percentile	5.2	7.2	4.9	4.4	5.2
Annual change (mm/year)					
75 th percentile	3	4	2	3	2
90 th percentile	6	8	4	6	4
Lifestyle factors					
Sports					
0 h/week	56.3	37.5	41.1	25.4	33.7
>0 to <2 h/week	18.9	20.9	35.3	49.8	41.6
≥2 h/week	24.8	41.6	23.6	24.8	24.7
Smoking habits (%)					
Current non-smoking	79.2	77.6	66.1	70.9	66.2
Current smoking	20.8	22.4	33.9	29.1	33.8
Alcohol consumption					
0 g/day	2.4	4.2	28.7	23.2	12.5
>0 to <6 g/day	37.5	31.8	22.0	23.0	47.7
≥6 to <18 g/day	32.9	29.3	22.3	19.7	22.2
≥18 g/day	27.2	34.6	27.0	34.1	17.6
Dietary factors					
Fruits and vegetables					
<1 portion/day	14.3	34.5	56.4	52.4	51.2
≥1 portion to <2 portions/day	17.5	26.4	25.8	23.5	19.0
≥2 portions/day	68.2	39.1	17.8	24.1	29.9

Appendix

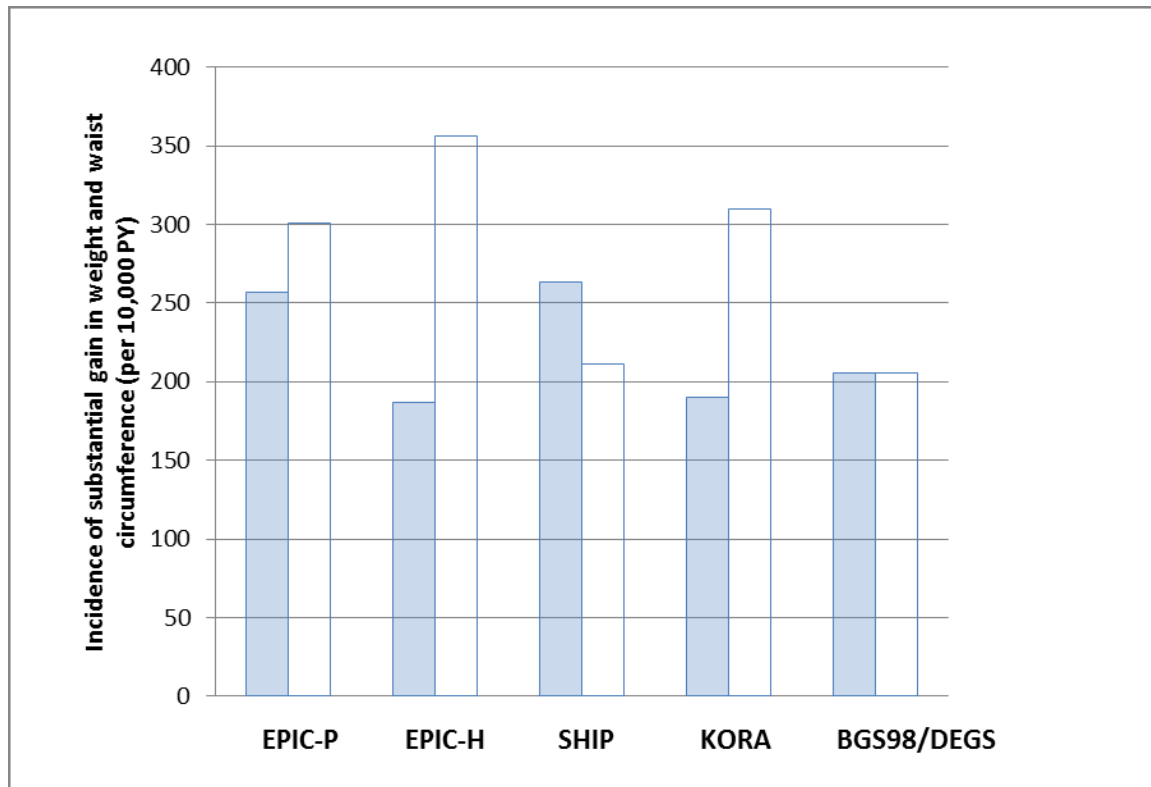
Meat					
<4-6 portions/week	12.2	58.3	20.9	39.0	27.8
≥4-6 portions/week to <1 portion/day	12.0	20.6	19.2	27.4	31.6
≥1 portion/day	75.8	21.1	59.9	33.7	40.6
Whole grain bread¹					
<2-3 portions/week	49.6	9.4	6.4	30.2	33.4
≥2-3 portions/week to <4-6 portions/week	17.1	15.4	7.5	33.2	26.0
≥4-6 portions/week	33.3	75.3	86.1	36.6	40.6
Chocolate					
<1 portion/week	22.4	29.5	42.6	37.7	27.4
≥1 portion/week to < 2-3 portions/week	23.8	19.6	17.6	22.2	24.1
≥2-3 portions/week	53.8	50.9	39.8	40.1	48.6
Cake and cookies					
<1 portion/week	15.6	19.3	32.1	32.0	21.2
≥1 portion/week to <2-3 portions/week	33.3	31.1	29.6	30.0	33.5
≥2-3 portions/week	51.2	49.7	38.3	38.0	45.3
Soft drinks					
<2-3 portions/month	70.8	69.9	56.6	56.3	40.6
≥2-3 portions/month to 4-6 portions/week	23.1	20.2	28.5	30.9	38.1
≥4-6 portions/week	6.2	9.9	15.0	12.9	21.3
Fish					
<2-3 portions/month	17.9	38.6	34.8	32.7	22.8
≥2-3 portions /month to <1 portion/week	24.0	23.2	30.1	28.1	25.2
≥1 portion/week	58.1	38.2	35.1	39.1	52.1
Further health related factors					
Sleep disorders (%)	13.4	11.9	16.8	17.2	16.0

¹ In EPIC-Heidelberg and SHIP intake of non-white bread was used instead of whole grain bread.

App. Table 2. Follow-up times, cases of SWG and SWCG during total follow-up and within the first 5 years across cohorts for the maximum models

	Study population				
	EPIC-Potsdam	EPIC-Heidelberg	SHIP	KORA	BGS98/DEGS
Follow-up time (PY)	124,239	101,986	14,155	11,851	17,976
SWG					
Cases during total follow-up (%)	3,188 (21.8)	1,902 (15.9)	373 (29.5)	225 (13.6)	370 (24.5)
Cases during first 5 years of follow-up (%)	896 (6.1)	487 (4.1)	64 (5.1)	89 (5.4)	67 (4.4)
SWCG					
Cases during total follow-up (%)	3,739 (25.6)	3,629 (30.4)	299 (23.7)	367 (22.2)	369 (24.4)
Cases during first 5 years of follow-up (%)	2,058 (14.1)	2,431 (20.4)	85 (6.7)	225 (13.6)	111 (7.4)

Follow-up times are presented in person-years (PY). Cases are presented in absolute numbers including proportions in parentheses.



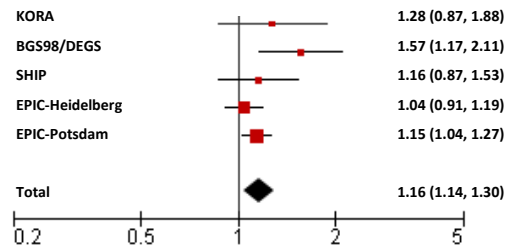
App. Figure 1. Incidence rates of SWG and SWCG (per 10,000 person-years) for the maximum models

Incidence rate = (Cases during total follow-up/person-years) * 10,000
SWG ... blue bars, SWCG ... white bars

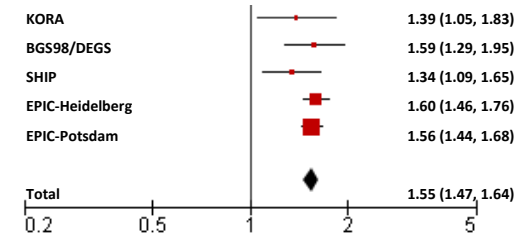
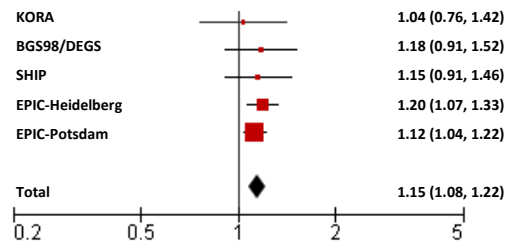
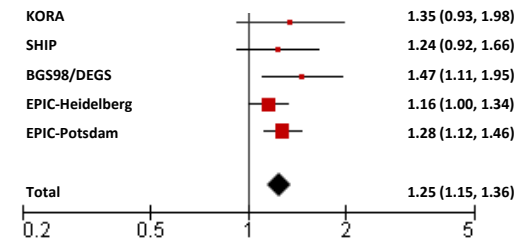
App. Figure 2. Associations of predictors with SWG and SWCG in the selection models

a, Associations with risk for SWG

Sex (female vs. male)

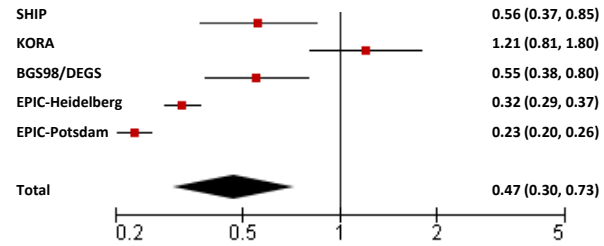
Heterogeneity: $\text{Chi}^2 = 6.55$ ($P = 0.16$), $I^2 = 39\%$

Smoking (vs. none smoking)

Heterogeneity: $\text{Chi}^2 = 2.98$ ($P = 0.56$), $I^2 = 0\%$ ≥ 2 -3 portions soft drinks/month to < 4 -6 portions/week (vs. < 2 -3 portions/month)Heterogeneity: $\text{Chi}^2 = 1.24$ ($P = 0.87$), $I^2 = 0\%$ ≥ 4 -6 portions/week (vs. < 2 -3 portions/month)Heterogeneity: $\text{Chi}^2 = 2.69$ ($P = 0.61$), $I^2 = 0\%$

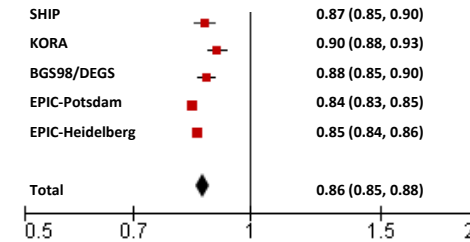
b, Associations with risk for SWCG

Sex (female vs. male)



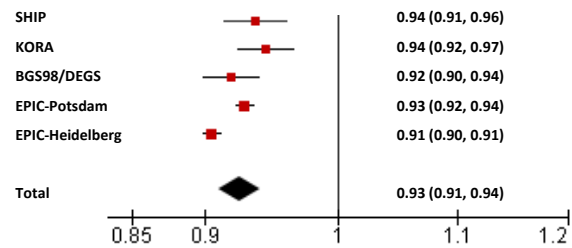
Heterogeneity: $\text{Chi}^2 = 84.79$ ($P < 0.0001$), $I^2 = 95\%$

Waist circumference (per cm)



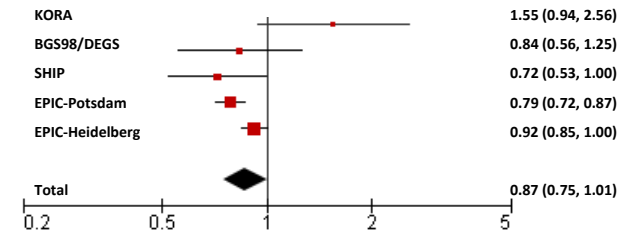
Heterogeneity: $\text{Chi}^2 = 33.60$ ($P < 0.0001$), $I^2 = 88\%$

Body height (per cm)



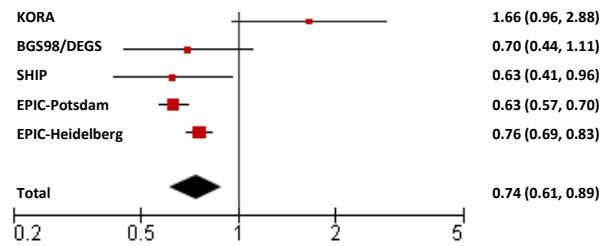
Heterogeneity: $\text{Chi}^2 = 35.89$ ($P < 0.0001$), $I^2 = 89\%$

Secondary/Professional school (vs. None/Primary school)

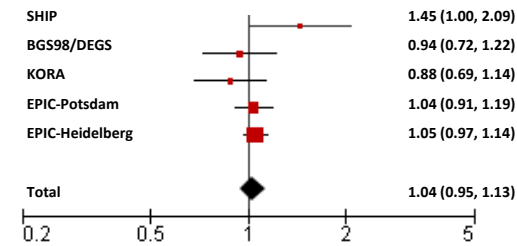


Heterogeneity: $\text{Chi}^2 = 12.53$ ($P = 0.01$), $I^2 = 68\%$

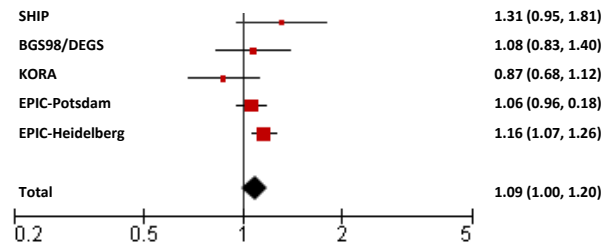
University (vs. None/Primary school)



Heterogeneity: $\text{Chi}^2 = 17.08$ ($P = 0.002$), $I^2 = 77\%$

 ≥ 4 -6 portions meat/week to < 1 portions/day (vs. < 4 -6 portions/week)

Heterogeneity: $\text{Chi}^2 = 5.36$ ($P = 0.25$), $I^2 = 25\%$

 ≥ 1 portion meat/day (vs. < 4 -6 portions/week)

Heterogeneity: $\text{Chi}^2 = 6.47$ ($P = 0.17$), $I^2 = 38\%$

App. Table 3. Predictive performance of the risk scores for the maximum models derived in SHIP, KORA, BGS98/DEGS and validated in EPIC-Potsdam and EPIC-Heidelberg

Outcome	Derivation sample						Validation sample 1		Validation sample 2	
	SHIP		KORA		BGS98/DEGS		EPIC-Potsdam		EPIC-Heidelberg	
	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT
SWG	0.73 (0.67, 0.79)	0.013	0.72 (0.67, 0.78)	0.341	0.81 (0.76, 0.85)	0.004	0.71 (0.70, 0.73)	<0.001	0.68 (0.65, 0.70)	<0.001
SWCG	0.73 (0.68, 0.78)	<0.001	0.66 (0.63, 0.69)	<0.001	0.71 (0.66, 0.75)	0.710	0.66 (0.65, 0.67)	<0.001	0.70 (0.69, 0.71)	<0.001

Prediction formula: $P(\text{SWG}, 5y) = 1 - 0.9604 + \exp^{[RSi - 394.10] / 100}$

Prediction formula: $P(\text{SWCG}, 5y) = 1 - 0.9228 + \exp^{[RSi - 1245.08] / 100}$

Meta-analytically combined ROC (95% CI) (SWG): 0.76 (0.70, 0.81)

Meta-analytically combined ROC (95% CI) (SWCG): 0.69 (0.65, 0.74)

App. Table 4. Predictive performance of the risk scores for the minimum models derived in SHIP, KORA, BGS98/DEGS and validated in EPIC-Potsdam and EPIC-Heidelberg

Outcome	Derivation sample						Validation sample 1		Validation sample 2	
	SHIP		KORA		BGS98/DEGS		EPIC-Potsdam		EPIC-Heidelberg	
	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT
SWG	0.70 (0.65, 0.76)	0.197	0.70 (0.64, 0.75)	0.027	0.79 (0.74, 0.84)	0.033	0.67 (0.65, 0.68)	<0.001	0.68 (0.66, 0.71)	<0.001
SWCG	0.73 (0.68, 0.78)	<0.001	0.65 (0.62, 0.69)	<0.001	0.70 (0.66, 0.75)	0.435	0.66 (0.65, 0.67)	<0.001	0.70 (0.69, 0.71)	<0.001

Prediction formula: $P(\text{SWG}, 5y) = 1 - 0.9596 + \exp^{[RSi - 452.87] / 100}$

Prediction formula: $P(\text{SWCG}, 5y) = 1 - 0.9212 + \exp^{[RSi - 1246.69] / 100}$

Meta-analytically combined ROC (95% CI) (SWG): 0.73 (0.67, 0.80)

Meta-analytically combined ROC (95% CI) (SWCG): 0.69 (0.65, 0.74)

App. Table 5. Predictive performance of the risk scores for the selection models derived in SHIP, KORA, BGS98/DEGS and validated in EPIC-Potsdam and EPIC-Heidelberg

Outcome	Derivation sample						Validation sample 1		Validation sample 2	
	SHIP		KORA		BGS98/DEGS		EPIC-Potsdam		EPIC-Heidelberg	
	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT	aROC (95% CI)	HLT
SWG	0.70 (0.63, 0.76)	0.034	0.66 (0.61, 0.72)	0.098	0.77 (0.73, 0.82)	0.062	0.67 (0.65, 0.69)	<0.001	0.69 (0.67, 0.70)	<0.001
SWCG	0.71 (0.65, 0.76)	<0.001	0.65 (0.62, 0.68)	<0.001	0.70 (0.65, 0.74)	0.573	0.66 (0.65, 0.67)	<0.001	0.71 (0.70, 0.72)	<0.001

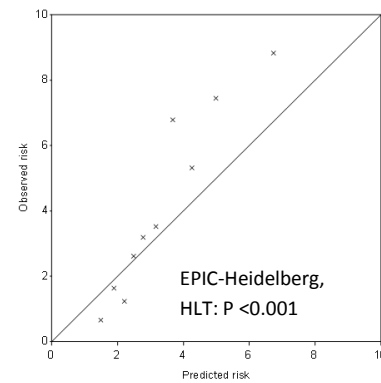
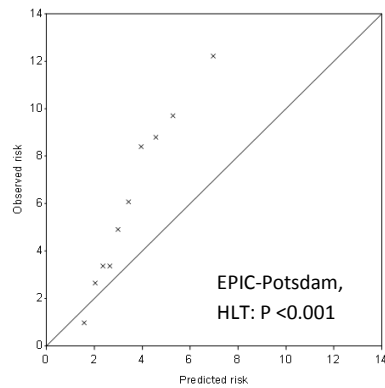
Prediction formula: $P(\text{SWG}, 5y) = 1 - 0.9569 + \exp^{[\text{RSi} - 616.73] / 100}$

Prediction formula: $P(\text{SWCG}, 5y) = 1 - 0.9189 + \exp^{[\text{RSi} - 1277.09] / 100}$

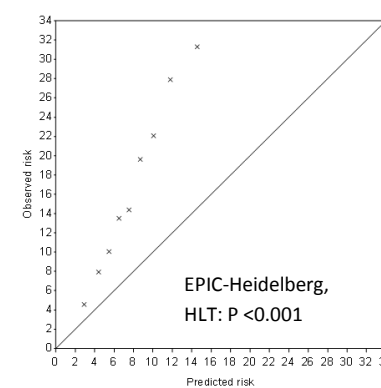
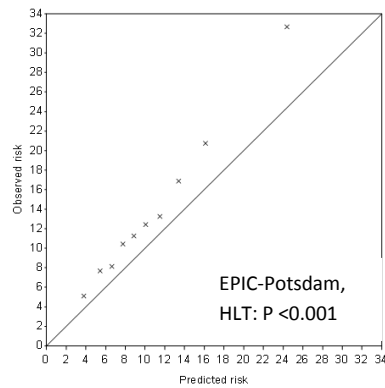
Meta-analytically combined ROC (95% CI) (SWG): 0.71 (0.64, 0.78)

Meta-analytically combined ROC (95% CI) (SWCG): 0.68 (0.64, 0.72)

SWG:



SWCG:



App. Figure 3. Calibration plots for SWG (upper row) and SWCG (lower row) for the selection model derived in SHIP, KORA, BGS98/DEGS and validated in EPIC-Potsdam and EPIC-Heidelberg

Danksagung

Die vorliegende Dissertation wurde in der Abteilung Epidemiologie am Deutschen Institut für Ernährungsforschung Potsdam-Rehbrücke erstellt. An dieser Stelle möchte ich mich bei all denjenigen bedanken, die mich während der Anfertigung der Dissertation unterstützt und motiviert haben.

Besonderer Dank gilt meinem Betreuer Herrn Prof. Dr. Heiner Boeing, der mich in den letzten Jahren in meiner wissenschaftlichen Entwicklung begleitet und gefördert hat. Außerdem möchte ich mich bei Prof. Dr. Reinhard Busse für die Möglichkeit der Promotion an der Technischen Universität Berlin und das konstruktive Gespräch bedanken.

Vielen Dank an meine Kolleginnen und Kollegen, insbesondere Frau Dr. Annika Steffen für ihre thematische und statistische Expertise und die hilfreichen Anmerkungen zum Manuskript.

Sehr herzlich möchte ich mich bei meinen Eltern und meinen Brüdern bedanken, die mir immer unterstützend zur Seite standen und dadurch maßgeblich zum Gelingen dieser Arbeit beigetragen haben.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass die im Fachbereich Management im Gesundheitswesen der Technischen Universität Berlin eingereichte Dissertation mit dem Titel „Derivation and validation of German risk scores predicting substantial gain in weight and waist circumference“ von mir selbständig angefertigt wurde und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet wurden. Weiterhin versichere ich, die Arbeit an keiner anderen Hochschule oder Fachhochschule eingereicht zu haben.

Berlin, 30. Oktober 2015

Ursula Bachlechner