PREDICTORS OF DROPOUT IN A LIFESTYLE INTERVENTION

- SUBSTUDY OF THE PREVIEW TRIAL

SINI ORPANA

MASTER'S THESIS

UNIVERSITY OF HELSINKI

DEPARTMENT OF FOOD AND ENVIRONMENTAL SCIENCES

DIVISION OF NUTRITION

JUNE 2018

Abstract

Faculty	Department
Faculty of Agriculture and Forestry	Department of Food and Environmental Sciences

Author

Sini Orpana

Title

Predictors of dropout in a lifestyle intervention. Substudy of the PREVIEW trial.

Subject

Nutrition Science

Level	Month and year	Number of pages
Master's thesis	6/2018	75

Abstract

Dropout is a serious problem in lifestyle interventions. Dropout is inadequately analyzed and inconsistently reported in weight loss and lifestyle intervention literature even though it is considered to be one of the most important aspects in assessing trial success. Attrition decreases the statistical power of the study, weakens the reliability and validity of the results, decreases generalizability of the outcomes, and deteriorates the accuracy of given recommendations. In addition, withdrawal leads to poor intervention outcomes and wasted resources. Dropout rates in weight loss intervention range from 7 to 90 percent.

The analysis of this study comprised of the first two years of a three-year intervention called PREVIEW. The participants analyzed in this thesis (N = 2088) were overweight or obese prediabetic adults with an age of 52 ± 11.5 years at the start of the trial. Participants came from six European and two Australasian countries. Two thirds of the participants were women and the majority were Caucasian. The final prediction model was computed with multivariate logistic regression.

Forty-four percent of the participants dropped out by year two. Dropouts were younger, had a higher baseline BMI, and lost less weight in the weight loss phase than the completers. Compared to Australia, the odds for dropping out were higher in the other countries analyzed. The odds for dropout in an ascending order were: Australia, Finland, New Zealand, the Netherlands, Bulgaria, the United Kingdom, Spain and Denmark. Compared to university education, the odds for dropping out were higher for participants with secondary vocational education, secondary school education, or education indicated as 'other'.

The findings of this thesis may be used in the development of similar trials and in public healthcare. In lifestyle intervention programs, special attention could be directed to individuals who are likely to drop out. In addition, as dropout is dependent of the explanatory variables, the results may also benefit future analysis of data from PREVIEW.

Keywords

Dropout (DO), attrition, withdrawal, lifestyle intervention, weight loss, weight maintenance (WM), completer, low-energy diet (LED), LED DO, early WM DO, late WM DO, body mass index (BMI), diabetes risk score (DRS)

Where deposited

The University of Helsinki, Department of Food and Environmental Sciences, Faculty of Nutrition Science

Additional information

Supervising professor: Professor Mikael Fogelholm

Supervisors: Associate Professor Thomas Meinert Larsen and postdoctoral researcher Pia Christensen

Tiivistelmä

Tiedekunta	Laitos
Maatalous-metsätieteellinen tiedekunta	Elintarvike- ja ympäristötieteiden laitos

Tekijä

Sini Orpana

Työn nimi

Keskeyttämisen ennustaminen elämäntapainterventiossa. PREVIEW-tutkimuksen alatutkimus.

Oppiaine

Ravitsemustiede

Työn laji	Aika	Sivumäärä
Pro gradu	6/2018	75

Tiivistelmä

Tutkimuksen keskeyttäminen on vakava ongelma elämäntapainterventioissa. Vaikka keskeyttäminen on yksi tärkein tutkimuksen onnistumisen mittareista, on se painonhallinta- ja elämäntapainterventiokirjallisuudessa riittämättömästi analysoitu ja epäjohdonmukaisesti raportoitu. Vetäytyminen vähentää tilastollista voimaa ja heikentää tulosten luotettavuutta, oikeellisuutta ja yleistettävyyttä ja annettujen suositusten tarkkuutta. Lisäksi keskeyttäminen aiheuttaa resurssien tuhlausta ja heikentää osallistujien onnistumista. Painonhallintainterventioiden keskeyttämisprosentit vaihtelevat välillä 7-90 %.

Tämä tutkimus käsitteli kolmivuotisen PREVIEW-tutkimuksen kahta ensimmäistä vuotta. Tutkimuksen osallistujat (N = 2088) olivat ylipainoisia tai lihavia esidiabeettisia aikuisia. Kuudesta Euroopan ja kahdesta Australaasian maasta tulevien osallistujien ikä oli tutkimuksen alussa 52 ± 11.5 vuotta. Kaksi kolmasosaa osallistujista oli naisia ja suurin osa valkoihoisia. Lopullinen ennustusmalli tehtiin logistisella monimuuttuja-regressiomallilla.

Keskeyttämisprosentti oli kahden vuoden kuluttua 44 %. Keskeyttäjät olivat nuorempia, heillä oli korkeampi BMI alussa ja he laihtuivat vähemmän laihdutusjaksolla tutkimuksen päättäjiin verrattuna. Australiaan verrattuna keskeyttäminen oli todennäköisempää muissa tutkimusmaissa. Vetosuhde keskeyttämiseen oli nousevassa järjestyksessä: Australia, Suomi, Uusi-Seelanti, Alankomaat, Bulgaria, Yhdistyneet Kuningaskunnat, Espanja ja Tanska. Yliopistotutkintoon verrattuna vetäytyminen oli todennäköisempää osallistujilla, joilla on alempi ammattikorkeakoulututkinto, ammattikoulututkinto tai tutkinto ’muu’.

Näitä tutkimustuloksia voidaan hyödyntää samantapaisten tutkimusten kehityksessä ja julkisessa terveydenhuollossa. Elämäntapainterventioissa voitaisiin suunnata erityishuomiota korkeamman vetäytymisriskin yksilöihin. Koska keskeyttäminen riippuu muista selittävistä tekijöistä, voidaan tämän tutkimuksen tuloksia hyödyntää PREVIEW-aineiston jatkotutkimuksissa.

Avainsanat

Keskeyttäminen (DO), vetäytyminen, elämäntapainterventio, painonpudotus, painonhallinta (WM), tutkimuksen päättäjä, vähäenerginen dieetti (LED), LED DO, alkuvaiheen WM DO, loppuvaiheen WM DO, painoindeksi (BMI), diabetesriskikerroin (DRS)

Säilytyspaikka

Helsingin yliopisto, Elintarvike- ja ympäristötieteiden laitos, Ravitsemustieteen tiedekunta

Muita tietoja

Valvova professori: professori Mikael Fogelholm

Ohjaajat: apulaisprofessori Thomas Meinert Larsen ja tutkijatohtori Pia Christensen

Preface

I would like to thank Professor Mikael Fogelholm from the University of Helsinki and Associate professor Thomas Meinert Larsen and Postdoc Pia Christensen from the University of Copenhagen for their awesome guidance during the process of this thesis. I would also like to give special thanks to Postdoctoral research fellow Maija Huttunen-Lentz from the University of Stuttgart for the consultation in the theoretical framework of the Health Belief Model.

Thanks for the advice in the statistical analysis of this thesis belong to Associate professor Christian Ritz from the University of Copenhagen and Postdoc Matti Koivisto from the Technical University of Denmark.

This thesis is dedicated to my parents Eeva Sajaniemi-Orpana and Aarne Orpana (1933-2013) and my husband Matti Koivisto. Thank you, mom, for your love, support, company and housing during my studies in Finland. Thank you, dad, for being present in my memories and my heart. I would also like to thank my parents-in-law, Rauha and Lauri Koivisto, for their support.

Above all, I would like to give very special thanks to my husband for the love and support he has given me throughout my studies, during the process of this thesis, and each day we have known each other. Matti, I have the utmost respect for you as a scientist, as a man and as a husband. You are my soulmate. I love you.

Table of contents

1 Introduction

Dropout, also referred to as attrition or withdrawal, is a serious problem in lifestyle interventions as it causes loss of statistical power, weakens the reliability and validity of results, decreases generalizability of outcomes, and deteriorates the accuracy of given recommendations. In addition, dropout is associated with poorer intervention outcomes, that is, unsuccessful weight loss and weight maintenance, and more importantly unadaptedness to new lifestyle skills. In addition to academic, societal and individual ramifications, attrition leads to wasted resources.

Withdrawal has been associated with, for example, younger age, lower socio-economic status (SES), poorer initial treatment response, not being married, immigrant status or non-Caucasian ethnicity. However, no clear and consistent set of predictors have been found. The reasons for this include the lack of a uniform definition of attrition, the inadequacy or lack of dropout analysis, the large variety of explanatory variables studied and the vast differences in study designs and populations.

The aim of this thesis is to identify features predicting dropout in the PREVIEW trial. Understanding the phenomenon of attrition increases academic knowledge, and can be utilized in the development of design, methods and tools of lifestyle interventions to come and in programs of public healthcare.

2 Literature review

2.1 Introduction

Dropout is a serious problem in lifestyle and dietary interventions aiming at weight loss and weight maintenance (Miller and Brennan 2015, Moroshko et al. 2011). Dropout is inadequately analyzed and inconsistently reported in weight loss and lifestyle intervention literature even though it is considered to be one of the most important aspects in assessing trial success. Attrition causes loss of statistical power, weakens the reliability and validity of results, decreases the generalizability of outcomes, and deteriorates the accuracy of given recommendations. Dropout is associated with poorer intervention outcomes both in weight loss and weight maintenance, and more importantly lifestyle skills being taught in the trials are not adopted. In addition to academic, societal and individual ramifications, attrition leads to wasted resources. The phenomenon of attrition is important to understand since it is the dropout rather than the study completer who will in the long run be the typical patient in the treatment of chronic illness (Baekeland and Lundwall 1975).

A systematic review on weight loss interventions reported attrition rates from 7 to 90 % depending on the type and setting of the intervention (Moroshko et al. 2011). Another review on lifestyle modification programs reported attrition rates from 20 to 82 % (Leung et al. 2017). To date, no reliable or consistent set of predictors of dropout (Moroshko et al. 2011) or adherence (Leung et al. 2017, WHO 2003) have been found (the association of dropout and adherence is discussed in chapter 2.2.4). The inconsistent findings might reflect differences in study population or design, different behavioral methods used in the studies, or differences in dropout definition and measurement (Miller and Brennan 2015). In addition, the inconclusive findings reflect the large number of variables analyzed and diversity of methodologies used across studies, and the small number of studies analyzing each variable (Moroshko et al. 2011).

All in all, dropout affects the quality of academic research and publishing, hinders the improvement of lifestyle skills and weight management of individual participants, wastes resources, and influences the societal recommendations of academic research.

2.2 The phenomenon of dropout

2.2.1 Definition of dropout

Even in studies of comparable design, methods and participants, the diversity, or in the worst case the lack of dropout definition, makes a just comparison of dropout rates between studies nearly impossible. Baekeland and Lundwall (1975) define a dropout as a) the participant who fails to return, b) the participant who refuses to return, or c) the participant who is expelled from the program due to lack of cooperation, poor response, adverse event or the like. Miller and Brennan (2015) define dropouts as participants who fail to complete all aspects of treatment, including the treatment follow-up. In addition, they continue by saying that defining dropouts as a homogenous group might lead to inconsistent findings as early and late dropouts might differ from one another.

2.2.2 Comparing definitions of dropout

Miller and Brennan (2015) reviewed definitions of attrition and attrition rates in 27 weight loss interventions and found that there is little consistency across studies. The studies compared were of varying length, for example, from 10 weeks to four years, and of varying intensity, for example, from 10 group meetings in two years versus a very low-calorie diet combined with weekly meetings for two years. Most of the studies analyzed were behavioral modification programs aiming at healthier lifestyles. They found that duration and research demands were not consistently related to attrition.

According to Miller and Brennan (2015), eleven studies defined attrition according to visit non-attendance, however most studies used different definitions. For example, in three studies missing the last visit, and in one study attending less than 75 % of visits,

was considered as attrition. Eight studies defined withdrawal according to time: In six studies, dropout was defined as withdrawing at any given time point before the end of trial, and in two studies cutoff times were used to define early and late dropouts.

In six studies not completing the entire or part of the treatments/tasks was used for defining dropouts. For example, in one study attrition was defined as noncompletion of both treatment and maintenance phases, and in another study as discontinuation of treatment before completing phase I of III. Two studies used more than one criteria in the definition: The other study defined it as either failing to attend certain number of visits or incompletion of required tasks, and the other as active withdrawal or visit non-attendance.

In sum, it is evident that with such a diverse set of definitions, dropout rates are inconsistent, fair comparison between studies cannot be done, studies with more stringent definitions subject to 'poorer' outcomes, and studies with looser definitions seem to be more effective (Miller and Brennan 2015). A coherent and uniform definition and measurement of attrition would allow an objective and just assessment of studies.

2.2.3 Factors associated with dropout

Depending on reference, the number of variables collected and the depth and extent of data analysis, predictors of dropout can be divided into two or three sets of features. According to Baekeland and Lundwall (1975), dropping out results from three vectors: 1) intrapsychic factors which include demographic, clinical and personality factors, 2) personality and attitudinal features of the treatment personnel, and 3) environmental factors, such as family attitudes, intervening life events and transportation problems. Today, strict protocols guide the behavior of treatment staff and as such attitudinal factors should not influence counselling, investigations or other visits attended by the participant. However, this phenomenon cannot be ruled out but might be impossible to analyze if specific data of staff is not collected.

According to Miller and Brennan (2015), attrition predictors can be divided into pre-treatment and post-treatment factors. These reasons can be categorized into practical/physical, program/treatment specific, and psychological barriers (Table 1). In addition to the variables related to participants, treatment related factors such as distance to clinic has also been explored as a predictor of dropout. However, findings on their ability to predict attrition have been inconsistent.

Pre-treatment factors which are gathered prior to trial include socio-demographic variables such as age, gender, education and marital status, psychosocial factors such as previous dieting attempts, body image, personality traits and social support, and anthropometric factors such as body mass index (BMI), waist-hip ratio, and muscle and fat percentage (Miller and Brennan 2015). While some of the pre-treatment factors are stable throughout the trial many of them are prone to change, for example, BMI and level of motivation. In these circumstances it is the change which might explain the dropout and not the initial measurement itself. Post-treatment factors for dropping out include self-reported reasons for dropping out.

Theoretically and empirically grounded psychological variables related to dropout include self-regulation, motivation, symptoms of depression, and weight loss expectations (Miller and Brennan 2015). Preliminary research suggests that these variables might predict attrition in a more consistent manner.

Table 1. Categories for reasons of dropout (Miller and Brennan 2015).

Categories	Reasons for dropping out
Practical/physical	Logistics, family, work, health
Program/Treatment related	Demands of the research, unsatisfactory results, dissatisfaction of treatment or staff
Psychological	Lack of motivation and self-confidence, resistance to change, feelings of abandonments

2.2.4 Dropout versus non-adherence

Dropout is not a direct synonym for non-adherence instead it represents the most severe form of non-adherence. WHO (2003) defines adherence as "the extent to which a person's behavior – taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a healthcare provider." For any adherence to occur, the symptoms of the condition to be treated must be severe enough to evoke the need to adhere. The condition must also be seen as resolvable and acute.

Dropout is the most common indicator used for analyzing adherence in lifestyle and weight loss studies (Stoutenberg et al. 2015). However, adherence is a multifactorial phenomenon, and reporting merely dropout rates can be considered inadequate. Hadžiabdić et al. (2014) suggested that attrition and weight loss should be analyzed simultaneously in order to achieve a more comprehensive overview of adherence (see chapter 2.4.5).

As with analyses of dropout, adherence analyses continue to focus on patient-related predictors of adherence, when it has previously been proven that treatment-related factors have a major effect on adherence (WHO 2003). Identifying participants who are more likely to adhere to treatment does not merely result in more successful weight loss, weight maintenance and adoption of new lifestyle skills but will at the same time decrease attrition, enable more efficient use of resources, and increase the quality of data analysis and preciseness of recommendations. Looking at this from another perspective, identifying features related to higher dropout risk enables the development of a more comprehensive set of tools directed at increasing adherence to the protocol and increasing retention in the treatment.

2.3 Dropout in two pooled systematic reviews

2.3.1 Background

In this chapter, the results of two systematic reviews are pooled together: 1) Factors associated with adherence to lifestyle modification (Leung et al. 2017), and 2) Predictors of attrition in weight loss interventions (Moroshko et al. 2011). In these reviews there were two overlapping studies, which is accounted for in the following tables. Most variables were analyzed in both reviews, and this is indicated in the tables. Factors associated with adherence were categorized into psychosocial, socio-demographic, behavioral, and physical factors.

Moroshko et al. (2011) found that older age and higher education were associated with increased retention. On the other hand, greater body dissatisfaction, poorer body image and mental health, higher number of previous dieting attempts, lower level of physical activity (PA), lower self-efficacy, lower social support, lower initial weight loss, and higher initial weight loss expectations were associated with increased attrition. In addition, greater travel distance, financial difficulties, and the need to pay for treatment were associated with higher attrition. Moroshko et al. (2011) concluded that psychological and behavioral variables seemed to be the most predictive variables of dropout.

According to Leung et al. (2017), limited evidence suggested that older age, higher education, healthier eating, a physically active lifestyle, greater initial weight loss, and greater stage-of-change at baseline predicted greater adherence. Stage-of-change refers to the change in diet and PA-related behavior. In contrast, depression, stress, a higher number of previous weight loss attempts, being unemployed and having a strong body shape concern associated with higher dropout. In this review, motivation, male gender, and self-efficacy were inconsistently associated with attrition.

2.3.2 Studies included in the reviews

Leung et al. (2017) studied nineteen lifestyle modification programs, out of which nine were randomized controlled trials, eight were pre- or post-interventions and two were longitudinal studies. The studies varied from eight weeks to six years. All studies concentrated on weight and had varying levels of attrition analysis. In the systematic review of Moroshko et al. (2011), sixty-one studies discussing attrition in weight loss interventions for overweight or obese adults were identified. Of these studies, forty-two were lifestyle/behavioral interventions, eight were surgical and seven were dietary studies. Single studies concentrated on exercise, pharmaceutical treatment or monetary incentives. The duration of the studies varied from seven weeks to four years.

2.3.3 Psychosocial factors

In most studies, there was no clear set of psychosocial factors predictive of dropout or non-adherence (Table 2, Leung et al. 2017, Moroshko et al. 2011). About half of the studies found no association between self-efficacy (47 %), motivation (50 %), stage-of-change (50 %) or anxiety (57 %) and attrition. However, stress, stage-of-change, and motivation showed conflicting results, for example, higher stress levels were associated with increased dropout in 57 %, but lower in 29 % of the studies. Lower self-efficacy (41 %), greater depression (50 %), and lower social support (57 %) were also associated with increased dropout.

Only two studies looked into body shape concerns, quality of life, self-esteem or perceived hunger, and other variables such as history of mental disorders, perfectionism, and disinhibition were analyzed in single studies only, thus these are not discussed here.

Table 2. Psychosocial factors associated with decreased adherence/increased attrition in reviews of Leung et al. (2017) and Moroshko et al. (2011).

Variable	Studies analyzing variable [N]	Direction of the correlation between variable and increased attrition (or decreased adherence).			Reference
		Negative [N (%)]	None [N (%)]	Positive [N (%)]	
Self-efficacy [1]	17	7 (41)	8 (47)	2 (12)	Leung et al. (2017), Moroshko et al. (2011)
Depression	16	1 (6)	7 (44)	8 (50)	Leung et al. (2017), Moroshko et al. (2011)
Motivation	8	2 (25)	4 (50)	2 (25)	Leung et al. (2017)
Stress	7	2 (29)	1 (14)	4 (57)	Leung et al. (2017), Moroshko et al. (2011)
Stage-of-change [2]	4	2 (50)	2 (50)	-	Leung et al. (2017)
Anxiety	7	-	4 (57)	3 (43)	Leung et al. (2017), Moroshko et al. (2011)
Social support	7	4 (57)	3 (43)	-	Leung et al. (2017), Moroshko et al. (2011)

[1] For example self-efficacy: In seven studies lower self-efficacy and in two studies higher self-efficacy were associated with increased attrition. In eight studies no asscociation was found.

[2] Stage-of-change refers to the change in diet and PA-related behavior.

2.3.4 Socio-demographic factors

Presented in Table 3, in at least half of the studies age (50 %), gender (68 %), employment status (67 %), education (71 %), SES (83 %), marital status (76 %), and ethnicity (50 %) were not found to be statistically significantly associated with dropout (Leung et al. 2017, Moroshko et al. 2011). Being younger (43 %), unemployed or having a part-time job (33 %), having lower education (29 %) or lower SES (17 %), being married (18 %) and being African-American or non-Caucasian (38 %) were associated with increased attrition in a few studies only. Gender showed conflicting results; no association was found in 68 % of the studies, and male and female genders each were found to predict dropout in 16 % of the studies.

Table 3. Socio-demographic factors associated with increased attrition or decreased adherence in reviews of Leung et al. (2017) and Moroshko et al. (2011).

Variable	Studies analyzing variable [N]	Studies in which no significant was association found [N (%)]	Studies [N (%)] and variable feature associated with increased attrition (or decreased adherence).	Reference
Age	46	23 (50)	20 (43): Younger	Leung et al. (2017), Moroshko et al. (2011)
			2 (4): Older	
			1 (2): Younger women	
Gender*	25	17 (68)	4 (16): Male	Leung et al. (2017), Moroshko et al. (2011)
			4 (16): Female	
Employment status	12	8 (67)	4 (33): Unemployed or part-time job	Leung et al. (2017), Moroshko et al. (2011)
Education	24	17 (71)	7 (29): Lower education	Leung et al. (2017), Moroshko et al. (2011)
SES	6	5 (83)	1 (17): Lower SES	Leung et al. (2017), Moroshko et al. (2011)
Marital status	17	13 (76)	3 (18): Married	Leung et al. (2017), Moroshko et al. (2011)
			1 (6): Not having a partner	
Ethnicity	8	4 (50)	3 (38): Non-Caucasian or African-American	Leung et al. (2017), Moroshko et al. (2011)
			1 (12): Immigrant status	

* For example gender: In 17 studies no association was found. In four studies male gender and in another four studies female gender was associated with increased attrition.

2.3.5 Behavioral factors

In almost half of the studies (44 %) eating behavior, and in more than half of the studies (60 %) PA was not associated with dropout (Table 4, Leung et al. 2017, Moroshko et al. 2011). Less healthy eating habits (50 %), less PA (40 %), and higher number of previous weight loss attempts (60 %) were associated with increased attrition. Smoking increased the risk of dropout in 56 % of the studies. None of the studies found an association between drinking habits and dropout.

Table 4. Behavioral factors associated with increased attrition or decreased adherence in reviews of Leung et al. (2017) and Moroshko et al. (2011).

Variable	Studies analyzing variable [N]	Direction of the correlation between variable and increased attrition (or decreased adherence).			Reference
		Negative [N (%)]	None [N (%)]	Positive [N (%)]	
Healthy eating behavior*	18	9 (50)	8 (44)	1 (6)	Leung et al. (2017), Moroshko et al. (2011)
Physical activity	15	6 (40)	9 (60)		Leung et al. (2017), Moroshko et al. (2011)
Previous weight loss attempts	15	1 (7)	5 (33)	9 (60)	Leung et al. (2017), Moroshko et al. (2011)
Smoking habit	9	-	4 (44)	5 (56)	Leung et al. (2017), Moroshko et al. (2011)
Drinking habit	5	-	5 (100)	-	Leung et al. (2017), Moroshko et al. (2011)

* For example healthy eating behavior: in nine studies less healthy eating and in one study healthier eating were associated with increased attrition (or decreased adherence). In eight studies no association was found.

2.3.6 Physical factors

Baseline weight/BMI/fat was not associated with attrition in most studies (67 %), and the rest showed somewhat conflicting results: Twenty-two percent of the studies found higher initial weight/BMI/fat, and 11 % found lower initial weight/BMI/fat to be associated with higher dropout (Table 5, Leung et al. 2017, Moroshko et al. 2011). In 88 % of the studies lower initial treatment response and in 67 % higher or unrealistic weight loss expectations were predictive of dropout. Age at the onset of obesity showed conflicting results. Only two studies looked into glucose intolerance and found impaired glucose tolerance to be associated with increased attrition. No association between blood pressure and dropout was found.

Table 5. Physical factors associated with decreased adherence/increased attrition in reviews of Leung et al. (2017) and Moroshko et al. (2011).

Variable	Studies analyzing variable [N]	Direction of the correlation between variable and increased attrition (or decreased adherence).			Reference
		Negative [N (%)]	None [N (%)]	Positive [N (%)]	
Initial weight/BMI/fat[1]	36	4 (11)	24 (67)	8 (22)	Leung et al. (2017), Moroshko et al. (2011)
				1 (3)[2]	
Initial weight loss/ treatment response	8	7 (88)	1 (12)		Leung et al. (2017), Moroshko et al. (2011)
Weight loss expectation	9	1 (11)	2 (22)	6 (67)[3]	Moroshko et al. (2011)
Age of onset of obesity	5	1 (20)	3 (60)	1 (20)	Moroshko et al. (2011)

[1] For example initial weight/BMI/fat: in four studies lower and in eight studies higher initial weight/BMI/fat mass were associated with increased attrition/decreased adherence. In 24 studies no association was found.

[2] In one study higher initial weight was associated with higher dropout only in women.

[3] Includes one study with 'unrealistic expectations'.

2.3.7 Conclusions

Lower self-efficacy, greater depression, and lower social support were associated with increased attrition. Being younger, unemployed or having a part-time job, having lower education or lower SES, being married, and being African-American or non-Caucasian were associated with increased dropout in a few studies only. Gender showed conflicting results. Unhealthy eating habits, less or no PA, higher number of previous weight loss attempts, and smoking were associated with increased attrition. Lower initial treatment response, and higher and unrealistic weight loss expectations also related to increased dropout risk.

2.4 Dropout in lifestyle interventions

This chapter presents a review on lifestyle interventions primarily aimed at prediabetics and/or overweight and obese adults. The purpose of this review is to sum up results of studies which are more comparable to the data used in this thesis.

A systematic review by Laws et al. (2012) focused on lifestyle interventions aimed at diabetes prevention. They found that most studies and reviews today concentrate largely on the efficacy and effectiveness of programs and thus focus on internal validity only. The importance of external validity should not be overlooked since it represents the generalizability of the findings to population beyond the study participants. Laws et al. (2012) concluded that although dropout rates were reported in almost all studies (29/31), subgroup analysis on attrition was lacking in most studies and merely four reported on factors predicting dropout. This underlines the inconsistency in reporting, the questionability of reporting quality, and the difficulty of study comparison. Besides Laws et al. (2012), also Moroshko et al. (2011) found that the majority of weight loss interventions neglect to report attrition rates and/or predictors of attrition.

Due to these shortcomings, the review in this chapter concentrates on summarizing attrition results from studies more precisely comparable to data analyzed in this thesis. Criteria for study inclusion were as follows:

1. Design: lifestyle intervention (diet and exercise included), aim at weight loss and/or maintenance.
2. Study participants: prediabetics and/or overweight and obese adults with no major chronic illnesses, such as diabetes or coronary heart disease.
3. Duration of study: ≥ 1 year.
4. Dropout analysis: preferably includes an in-depth dropout analysis, minimum requirement is, however, that dropout rate is reported or can be concluded from figures provided.

The literature search was performed mainly in PubMed during January and February of 2018. The key search terms included attrition, dropout, withdrawal, intervention, lifestyle, prediction, prevention, type 2 diabetes (T2D), overweight, obese, obesity, and randomized controlled trial, and combinations of the former. Restrictions in publishing years were not used.

After excluding most articles found, the search yielded 21 articles matching the above-mentioned criteria (see Appendix 1). The rejected articles included, for example, diet-only-interventions and observational intervention programs, participants with severe obesity and/or comorbidities and/or were less than a year in duration. A few studies were rejected due to missing dropout rates or figures from which these could have been concluded.

All studies were either lifestyle interventions or included a lifestyle intervention group, and all participants were prediabetic and/or overweight or obese. The duration of the studies varied from one to six years, and the number of participants ranged from 32 to 2553. The dropout rates varied from 4 to 59 %. Dropout rates were not directly reported in three articles; however, they could be concluded from figures given. Altogether thirteen articles included some analysis on factors associated with attrition; only statistically significant variables are discussed in the following sections. Fifteen studies included prediabetic participants, however, only six of these included dropout analyses.

As in the systematic reviews by Leung et al. (2017) and Moroshko et al. (2011), the definitions of dropout were not unified in these articles. For further information, see Appendix 2.

2.4.1 Predictors of dropout in reviewed articles

Psychosocial and behavioral factors

Greater depressive symptoms, binge eating and previous weight loss attempts were related to increased attrition in two studies (Table 6). According to single studies, increased stress and increased hunger-satiety scores were associated with higher dropout. Single studies analyzing psychological health score, quality of life score and body image satisfaction found a negative correlation between the variable and attrition, that is, lesser scores were associated with increased attrition. History of weight cycling was not associated with dropout.

Socio-demographic factors

Five studies found age to correlate negatively with attrition, and four studies found no association (Table 6). Three studies found no association between gender and dropout, yet another three found that male participants had a higher dropout risk. One study found marital status not to be related to attrition, however another study found that being divorced, separated or widowed associated with increased attrition. A single study analyzing SES found no association between the variable and dropout. One study found no correlation between education and dropout, and another study found that the less educated were more likely to drop out. Being unemployed or retired related to increased attrition in single studies, yet in a third study work status was not association with attrition. Lower income level was related to greater attrition in a single study.

Physical factors

Five studies found baseline BMI not to be related to dropout, however another five studies found that greater baseline BMI associated with increased dropout risk (Table 6). Greater early weight loss predicted longer retention time in one study. Single studies found that greater baseline weight, waist circumference and number of comorbidities were associated with increased attrition. Chronic pain was seen as a risk for dropout in one study. Two studies found that greater 2-h plasma glucose levels at

baseline predicted earlier dropout, however, one study found no association between them.

Treatment related factors

Single analyses found that exercise minutes, VO_{2max} and carbohydrate, energy and fiber intake correlated negatively with dropout, that is, for example, less time devoted to PA and lower energy intake increased dropout risk (Table 6). Not possessing a healthcare insurance and poorer group attendance related to increased attrition in single studies. One study analyzing clinic-related factors found that younger staff and large or small user populations associated with increased dropout.

Table 6. Factors associated with attrition in 13 of the 21 reviewed studies.

Factors with positive, negative or no association with increased attrition	Reference
Baseline BMI +, waist circumference +, unemployment +, retirement -.	Absetz et al. 2007
No association in baseline BMI. Age -.	Andersson and Rössner 1997
No association in baseline values (BMI, tricep skin-fold index, oxygen uptake, diastolic blood pressure, systolic blood pressure, cholestrol, triacylglyserides, hypertention medication).	Eriksson et al. 1991
Baseline depression symptoms +, age -, early weight loss -.	Fabricatore et al. 2009
Baseline BMI +, previous weight loss attempts +, not possessing health insurance +, BES[1] +, BDI[2] +, HSS[3] +, age -, education -, group attendance -.	Goode et al. 2016
BMI +, education -.	Hadžiabdić et al. 2014
Short-term predictors: Male +, age -, household income -, comorbidity +, chronic pain +, small/large site user population +, staff age -. Long-term predictors: Male +, age -, chronic pain +, small/large site user population +, staff age -, marital status (separated, divorces, widowed) +.	Jiang et al. 2016
No association in age, gender, baseline BMI, baseline fasting glucose or 2-h plasma glucose concentrations.	Mensink et al. 2003
No association in gender, age, BMI, history of weight cycling, education, work or marital status. Stress +.	Michelini et al. 2014
No association in age, gender and major risk factors.	Oldroyd et al. 2006
3-year results: Baseline BMI +, 2-h plasma glucose level +, VO2max -. 6-year results: BMI +, 2-h plasma glucose level +, VO2max - and SES - .	Roumen et al. 2007 and 2011
No association in age and BMI. Male +.	Sakane et al. 2011
Previous weight loss attempts +, weight outcome evaluation strictness +, baseline weight +, binge eating +, carbohydrate and energy intake at baseline -, quality of life scores -, exercise minutes -, fiber intake -, psychological health score -, body image satisfaction -.	Teixeira et al. 2004

No association means that the variable was not associated with dropping out on a statistically significant level.

\+ = positive association means increase in variable associates with increased attrition.

\- = negative association means decrease in variable associates with increased attrition.

[1] Binge Eating Scale, [2] Beck Depresion Inventory, [3] Hunger Satiety Scale

2.4.2 Reasons for withdrawal

Thirteen studies reported on reasons to drop out (Table 7). The most common reasons were medical (mentioned in 10 studies), fail to respond/loss of contact or unknown reason (8), death (6), refusal/unwillingness to continue (4), personal reasons (4), moving from the area (4), dissatisfaction (3), work commitment (3), lack of time (2), switching medical care/treatment program (2), lack of motivation (2), transportation issues (1), pregnancy (1) and family issues (1).

Table 7. Reasons for dropping out were reported in 13 of the 21 reviewed articles.

Reasons for dropout	Reference
7 personal problems (such as an ill wife or alcohol problems), 5 no wish to participate, 9 unknown reasons, 4 diseased, 3 moving from the area, and 1 going into another treatment program.	Andersson and Rössner 1997
6 switched medical care, 4 died, 2 moved away, and the others failed to respond to further invitations.	Eriksson et al. 1991
15 % unmotivated, 8 % health-related issues, 6,5 % unknown reasons, and 2,5 % dissatisaction with the program.	Hadžiabdić et al. 2014
6 for time commitment and 3 for medical reasons (arthritis symptoms, back pain, and thrombocytosis).	Liao et al. 2002
5 medical reason and 17 subjects refused to attend the annual visits.	Mensink et al. 2003
35 unspecified reasons, 4 for acute illness, 3 pregnancies and 2 for unforseen job difficulties.	Michelini et al. 2014
14 participants, reasons for withdrawing were family problems, work commitments or ill health.	Oldroyd et al. 2006
7 people refused follow-up, 29 left the area (for work), and 11 died	Pan et al. 1997
17 lost to follow-up, 9 unwilligness to continue and 2 died.	Ramachandran et al. 2006
10 medical reasons, 7 lack of time, 7 lack of motivation, 3 dissatisfaction, 2 unkonown reasons, 1 no response, 1 no transportation, and 1 death.	Roumen et al. 2007 and 2011
40 loss of contact, 20 reason related to the centers (closure of a center etc.), 18 personal reasons (moving etc.), and 5 medical reasons.	Sakane et al. 2011
35 % lack of time, 22 % dissatisfaction with the program/staff, 17 % personal life issues, and 17 % health limitations.	Teixeira et al. 2004
9 could not be contacted, 3 withdrew due to severe illness, 1 died, and 27 withdrew for personal reasons.	Tuomilehto et al. 2001

2.4.3 Incentives

In most articles (18/21) no incentives to increase participant retention were mentioned. In one study a discount card for the city's public leisure facilities was offered in order to induce PA (Oldroyd et al. 2006), however, it is questionable if this incentive actually decreased withdrawal as the dropout rate was 21 % at year one and 31 % at year two. In another study pedometers or exercise T-shirts were given after

the participant completed each assessment (Jiang et al. 2016). The attrition rate of this study was 43 % at year one. Given the high dropout rates in these studies, it seems unlikely that the incentives used biased retention.

In the third article three studies by the same author and same research facility were reviewed, and incentives were used in two of these studies (Goode et al. 2016). In the first one (SMART trial) participants were compensated for their time at the assessments (6, 12, 18, and 24-month), however, the compensation was not specified (Burke et al. 2009). In another study (PREFER trial) all subjects received $50 for each assessment attended, and a magazine subscription (Burke et al. 2006). Participants were also compensated for completing paper-and-pencil instruments and three-day food records, and for their time at the appointments (6, 12, 18, and 24-months). The compensation in the latter was not specified. With no incentives (SELF trial) the dropout rate was 20 % at 18 months (Burke et al. 2015), with monetary and other incentives (PREFER trial) the dropout rate was 18 % at 18 months, and with unspecified incentives (SMART trial) the dropout rate was 10 % at 24 months. Thus, it seems that in these trials incentives decreased attrition.

In conclusion, it is possible that strong incentives affect retention rates, thus further complicating the comparison of dropout rates across studies and the analysis of predictors of dropout.

2.4.4 Dropout rates and their increase during trial

The range of dropout in the studies reviewed varied from 4 to 59 %. Dropout rates were not directly reported in three articles, however could be concluded from figures provided. The average unweighted cumulative attrition rate was 23 % at year one, 26 % at year two and 35 % at year three.

Dropout is rarely stable over the course trials. In a two-year weight loss and weight maintenance lifestyle intervention the end of trial attrition rate was 51 % (Rothenberg

et al. 2015). (This study was left out from the review due to including participants with severe comorbidities.) The authors concluded that attrition was greatest between six and twelve months, after the transition from a very low-calorie-diet to regular food. This is an interesting finding since most lifestyle intervention studies fail to analyze the absolute attrition rates at different times during trial and report merely the final dropout rate.

The 21 studies included in this review lacked analysis of the absolute dropout rates at different times during trial. Only seven studies included specific attrition rates over the course of the trials (Table 8). It was found that on average (6/7), attrition was strongest during the first year of trial, and the absolute dropout rate declined towards the end of the trial. However, in one trial (Absetz et al. 2007) the cumulative rate of attrition at year one was 11 % and at year three was 33 %, which might reflect a stable dropout throughout the trial. In another study, the absolute dropout was highest in the first year, after which there was fluctuation rather than a steady decline (Penn et al. 2009). However, no proper conclusion of dropout rate development can be drawn from such limited data.

Table 8. Cumulative attrition rates at specific times in 7 of the 21 reviewed studies.

Cumulative attrition rates at specific times during trial	Reference
Year 1: 11 %, year 3: 33 %.	Absetz et al. 2007
Year 1: 23 %, year 2: 31 %.	Bourn et al. 1994
Month 6: 14 %, month 18: 18 %.	Goode et al. 2016
Year 1: 25 %, year 2: 27 %.	Mensink et al. 2003
Year 1: 21 %, year 2: 31 %.	Oldroyd et al. 2006
Year 1: 22 %, year 2: 31 %, year 3: 43 %, year 4: 45 % and year 5: 59 %.	Penn et al. 2009
Year 1: 15 %, year 2: 28 %.	Roumen et al. 2007

2.4.5 Association of weight and dropout

Attendance and completion correlate positively with weight loss outcomes, that is, dropouts tend to have poorer outcomes in regards to weight loss and weight maintenance (Moroshko et al. 2011). Factors associated with attrition may also predict the level of weight maintenance success. According to a review by Elfhag and Rössner (2004), successful weight maintenance was associated with more initial weight loss, reaching a self-determined weight goal, physically active lifestyle, regular meal rhythm, healthier eating, self-monitoring behavior, control of over-eating, internal motivation, and social support. In addition, psychological strength and stability, better coping strategies and ability handle stress, self-efficacy, autonomy, and responsibility were also associated with successful weight maintenance.

As with varying definitions of attrition, one challenge in comparing studies is the varying definitions of 'weight maintenance'. In Elfhag and Rössner's (2004) review, weight maintenance implied an initial weight loss that had been maintained for at least six months. In addition to positive predictors, they found that history of weight cycling, disinhibited eating and binge eating, higher level of hunger, negative emotions and stress-related eating, and passive problem solving were associated with weight gain.

Thus, when assessing the success of weight loss and lifestyle interventions, it is important to a) analyze both the weight outcomes and the dropout, and b) understand that both unsuccessful weight loss/weight maintenance and increased dropout may be explained by the same variables and are thus interconnected.

2.5 The Health Belief Model

2.5.1 Background

The Health Belief Model (HBM) is a widely used psychological approach explaining health-related behavior. The HBM was originally developed by Hochbaum (1958) and Rosenstock (1960, 1966), and later extended to apply to people's responses to symptoms (Kirscht 1974) and to their behavior in response to diagnosed illness (Becker et al. 1974).

The HBM is based on value-expectancy concepts, that is, a) the desire to avoid illness or recover from illness (value), and b) the belief that specific actions will prevent or ameliorate illness (expectation) (Strecher and Rosenstock 1997). The expectancy is further defined as the individual's estimate of personal susceptibility to illness, severity of illness, and the likelihood of being able to decrease threat by personal action.

Starting from 1952 Hochbaum assessed the readiness of more than 1200 adults invited to a free-of-charge tuberculosis screening with a mobile X-ray unit and demonstrated with considerable accuracy that the probability of a person to be screened was strongly associated with the perceived susceptibility and perceived benefits of that person (Strecher and Rosenstock 1997). Eighty-two percent of those who had a belief in their own susceptibility and a belief that early detection would be beneficial, volunteered for at least one scanning. On the other hand, only 21 % of those who exhibited neither of these beliefs volunteered for the X-ray.

2.5.2 Elements of the Health Belief Model

It is now believed that individuals will take action if they;

1) regard themselves as susceptible to the illness,
2) believe that serious consequences exist,
3) believe that certain available actions are beneficial in either reducing their susceptibility or severity of illness, and

4) believe that barriers of taking action are outweighed by its benefits (Becker et al. 1974, Kirscht 1974)

Perceived threat

Perceived severity refers to the subjective assessment of the severity of the illness and its potential medical, clinical and social consequences (Strecher and Rosenstock 1997). That is, the model proposes that individuals who consider the illness more severe, are more likely to take action to prevent it. Perceived susceptibility refers to the individual's subjective perception of the risk of getting ill, and includes the acceptance of diagnosis, personal estimates of susceptibility and susceptibility to illness in general. That is, individuals who perceive themselves as high-risk individuals are more likely to take action to prevent it. Altogether perceived severity and perceived susceptibility are referred to as perceived threat.

Cost-benefit analysis

Perceived benefits refer to the beliefs regarding the effectiveness of the available actions aimed at reducing the threat of illness (Strecher and Rosenstock 1997). Perceived barriers refer to the possible negative aspects of the former action, that is the health action itself may be time-consuming, upsetting, difficult, and unpleasant. The individual will consciously or un-consciously weight the costs of the action against the benefits, and only then decide if any actions are taken.

Cues to action

Although never studied empirically by Hochbaum, he suggested that the readiness to take action could only be potentiated by other factors, cues, such as bodily events or by environmental events, such as the media (Janz and Becker 1984). Janz and Becker (1984) concluded that these necessary triggers may be internal, such as physiological symptoms, or external, such as information, illness of a friend or family member, or a health warning label. The intensity of cues needed to produce action varies between individuals, and is affected by perceived threat, benefits and barriers.

Other variables

Demographic, sociopsychological and structural variables can affect the individual's perceptions and thus indirectly influence health behavior (Strecher and Rosenstock 1997). Especially sociodemographic factors, particularly education, are believed to influence perceived threat, benefits and barriers, and thus affect the actions taken. Self-efficacy was added to the HBM model in 1988. Self-efficacy refers to the conviction that one can successfully execute the behavior required to produce the outcome. From the elements discussed above a model for assessing health action was created (Figure 1).

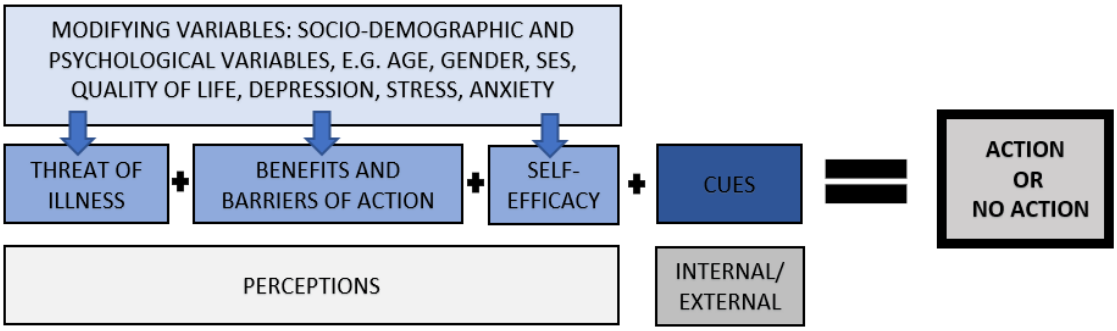

Figure 1. The Health Belief Model for assessing outcome action (modified from Janz and Becker 1984).

As an example, let us consider an obese prediabetic individual with moderate self-efficacy and a family history of T2D. According to the HBM, if this person a) perceives their personal threat of T2D to be high, b) believes that a regular meal-rhythm and exercising will lead to weight loss and reduced disease risk, and c) believes that the benefits of such actions outweigh the barriers, it is likely that they will take actions to prevent the disease from emerging. The decision to take action would indicate better adherence to program and decreased risk to drop out.

2.6 Summary of the literature findings

Dropout rates vary from 7 to 90 % in weight loss interventions (Moroshko et al. 2011). Attrition is inadequately analyzed and inconsistently reported in weight loss and lifestyle intervention literature (Miller and Brennan 2015, Moroshko et al. 2011). Dropout causes loss of statistical power, weakens the reliability, validity and accuracy of the results and recommendations, decreases generalizability of outcomes, and wastes resources. In addition, dropout is associated with poorer intervention outcomes in weight loss, weight maintenance and lifestyle skills. The phenomenon of attrition is important to understand since it is the dropout rather than the study completer who will in the long run be the typical patient in the treatment of chronic illness (Baekeland and Lundwall 1975).

Two systematic reviews found that lower self-efficacy, greater depression, and lower social support were associated with increased attrition (Leung et al. 2017, Moroshko et al. 2011). Being younger, having lower education, unhealthy eating habits, less or no PA, more previous weight loss attempts, and smoking were associated with increased attrition. In addition, lower initial treatment response, higher and unrealistic weight loss expectations and greater body dissatisfaction also related to higher dropout risk. Being unemployed or having a part-time job, lower SES, being married, and being African-American or non-Caucasian were associated with increased attrition in a few studies only. Gender showed conflicting results.

In the review of 21 lifestyle interventions, it was found that greater depressive symptoms, binge eating and higher number of previous weight loss attempts were associated with increased dropout. Younger age, male gender, greater baseline BMI/weight, lower initial treatment response or being divorced, separated or widowed were also related to increased dropout risk. Lower education, unemployment, retirement, and lower income level were associated with increased attrition in single studies.

3 Objectives

3.1 Objectives of this study

The main objective of this study was to analyze if dropout could be predicted using easily, objectively and affordably collected variables. Detecting participant features associated with dropout is valuable from at least two perspectives: a) potential participants in high attrition risk categories may be ruled out of treatment or study, and b) more precise tools may be developed in order to decrease withdrawal in all participant groups.

Reporting on attrition and analyzing predictors of dropout have only been done in a few studies of similar design before. Unlike general weight loss studies, most PREVIEW-comparable lifestyle interventions lack dropout analysis. One systematic review addressing external validity in diabetes prevention studies found that only four out of thirty-one studies included analysis of attrition (Laws et al. 2012). Thus, due to these shortcomings, and to add academic knowledge, this thesis focused on attrition. Unlike most studies using complete data sets only, this thesis included imputed values for the analysis of overall weight maintenance success and dropout.

The data used in this study included the first two years of a three-year intervention trial (PREVIEW), and due to this the word 'completer' refers to all participants who participated at least two years. The variables analyzed in the prediction of dropout included socio-demographic variables and weight data, that is, variables which can be easily collected and measured in any setting and in any country. In addition to finding possible predictive features, the results were compared to other similar studies.

The socio-demographic variables used in this thesis included age, gender, ethnicity, country, education, employment and marital status. The clinical variables included baseline BMI (and weight), low-energy diet (LED) weight loss percent (WL_{LED}), two-year weight loss percent ($WL_{2\text{-}YEAR}$), and diabetes risk score (DRS). The behavioral variables

included previous weight loss attempts, smoking and drinking habits. In addition to these, reasons for ending trial were looked into.

The research questions of this thesis were:

1. Are the examined variables (age, gender, ethnicity, country, education, employment status, marital status, baseline BMI (and weight), WL_{LED}, DRS, previous weight loss attempts (Y/N), number of previous weight loss attempts, number of successful weight loss attempts, smoking, and alcohol consumption) found similar across dropouts and completers?
2. Can dropout be predicted using the variables mentioned above?
    a. How well does the model predict the odds for dropping out?
    b. Which explanatory variables remain in the final model?
3. Does $WL_{2\text{-}YEAR}$ predict dropout?
4. What are the main reasons for ending trial?

3.2 Hypotheses

The following hypotheses were based on the findings in literature and were tested in the statistical analysis of this thesis.

1. It was hypothesized that the dropouts are more likely younger, men, less educated, single, unemployed, of ethnic minority, and smoke more than the completers. It was also hypothesized that multiple previous weight loss attempts predict earlier withdrawal.

2. It was hypothesized that completers lose more weight during the LED phase and maintain their weight more successfully in the weight maintenance phase than the dropouts.

4 Data and methods

4.1 Data

4.1.1 The PREVIEW trial

The study has been described in detail elsewhere (Fogelholm et al. 2017). In short, the PREVIEW trial was a randomized, controlled, multinational lifestyle intervention for prediabetic adults conducted over a period of three years through group meetings. PREVIEW began with a low-energy diet (LED) phase lasting for eight weeks. A minimum of 8 % weight reduction was required for the eligibility for the following 148-week weight maintenance phase.

The trial was setup in a 2 x 2 factorial design with two different diets and two different exercise regimes resulting in four study groups. The diets included a high-protein (25 E%), moderate-carbohydrate (45 E%), low-glycemic index diet (GI < 50) and a moderate-protein (15 E%), high-carbohydrate (55 E%), medium-GI diet (GI > 56). The exercise regimes used were high intensity (> 6 MET) and moderate intensity (3-6 MET), in which MET is the metabolic equivalent of task (energy expenditure compared to resting energy expenditure). Block randomization with stratification by gender and age group was used for the grouping of participants.

PREVIEW differed from previous studies alike in study population size, internationality, design and methods. The objective of PREVIEW was to identify the most effective lifestyle (diet and exercise) in the prevention of T2D. The primary endpoint of the study was the incidence of the disease.

4.1.2 Data collection and variables

Participants were recruited via multiple methods, for example, newspaper, radio and TV advertisements, and direct contact with healthcare providers (Fogelholm et al. 2017). Interested individuals were pre-screened in a phone-interview in which inclusion and exclusion criteria were queried, including the Finnish Diabetes Risk Score

assessment (Silventoinen et al. 2005). Potential participants were given written and oral information. Signed informed consent was required prior to laboratory screening which included measurements of weight, height, resting blood pressure, electrocardiography (for 55-year-olds and older), and an oral glucose tolerance test.

The inclusion criteria used in screening are presented in Table 9. In brief, participants eligible for the study were overweight or obese prediabetics between the ages of 25 to 70 years. Serious existing medical conditions, such as diabetes mellitus, significant cardiovascular disease and chronic renal impairment were used as exclusion criteria. More specific inclusion and exclusion criteria are described elsewhere (Fogelholm et al. 2017).

Table 9. The inclusion criteria used for participant screening in the PREVIEW trial.

Age	25-70 years
Weight	BMI ≥ 25 kg/m²
Prediabetes	The criteria from WHO/IDF (International Diabetes Foundation) for assessing prediabetes will be used as the formal inclusion criteria, i.e. having: • Impaired Fasting Glucose: Fasting venous plasma glucose concentration 5.6 - 6.9 mmol/l or • Impaired Glucose Tolerance: Venous plasma glucose concentration of 7.8 - 11.0 mmol/l at 2 h after oral administration of 75 g glucose (OGTT), with fasting plasma glucose less than 7.0 mmol/l. Due to potential between-lab variation (local assessments), HbA1c is not used as an inclusion criteria in the screening.
Informed consent	Required
Smoking	Smoking is allowed, provided subjects have not recently (within 1 month) changed habits. However, smoking status is monitored throughout the study and used as a confounding variable.
Motivation	Motivation and willingness to be randomized to any of the groups and to do his/hers best to follow the given protocol.
Other	Able to participate at clinical investigation days (CID) during normal working hours.

The data analyzed in this thesis comprised of the first two years of the three-year intervention, that is, from screening (visit 1) to visit 17 (CID6). Visit 2 (CID1) was baseline, and weight data collected at this visit was used as the baseline weight of participants. Additional data on 'last visit attended' and 'end of trial reason' were used for the categorizing of the participants. Data was retrieved from Open Clinica's

electronic case report forms and included all eight study sites: Bulgaria, Denmark, Finland, the Netherlands, Spain, the United Kingdom, Australia and New Zealand.

The socio-demographic variables analyzed in this thesis included age, gender, ethnicity, country, education, employment and marital status. The biological/ anthropometrical variables included weight, BMI, DRS, WL_{LED} and $WL_{2\text{-}YEAR}$. Other variables included smoking, drinking and previous weight loss attempts. The differences in these variables between participant groups were analyzed. In addition, the variables were further analyzed as possible explanatory factors in the prediction model for dropout.

4.1.3 Study flow

The intervention was carried out as a fading visit approach, that is, the frequency of meetings and supervision gradually declined towards the end of the study (Table 10). The purpose of this approach was to accustom the participants back to normal living conditions. In the first six months (screening and until week 26) there were 11 visits, and in the second half of the first year there were three visits. In the second year there were three visits altogether.

Table 10. The number, timing and details of visits in the first two years of the three-year intervention.

Number of visits		11											3			3		
Visit No		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Time [months]			0		1		2		3	4	5	6	8	10	12	15	18	24
Time [weeks]			0	2	4	6	8	10	12	16	20	26	32	44	52	64	78	104
Prescreening	x																	
Screening		x																
LED phase				x	x	x	x											
CID			1				2					3			4		5	6
Weight data			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

4.2 Categorizing participants

Altogether 2326 participants were eligible to enter the LED phase. Of these two participants were removed due to lack of screening data (height needed for BMI), 103 were removed due to lack of baseline weight (weight needed for BMI and weight loss percent) and 133 were removed due to missing end of trial information (needed for group categorizing). Thus, the total number of participants analyzed in this thesis was 2088.

4.2.1 Dichotomous outcome status

In the PREVIEW protocol a completer was officially defined as a participant who remains in the trial for the full three-year-intervention period, or who is diagnosed with T2D before the end of trial (Fogelholm et al. 2017). Since the data used in this thesis comprised of the first two years of the intervention, completers were defined as participants who did not end the trial before the two-year mark (visit 17), and those who were diagnosed with T2D. Dropouts were defined as participants whose last visit was visit 16 or less. Participants whose last visit was less than visit 17 (CID6) and who lacked the end of trial information were ruled out of the analysis, as there could be no just conclusion on their outcome status. This categorization resulted in a dichotomous status of completers and dropouts.

4.2.2 Precise outcome status

Participants in this study were further divided into LED dropouts (last visit less or equal to visit 6), early weight maintenance dropouts (last visit between visits 7 and 11), late weight maintenance dropouts (last visit between visits 12 and 16), and completers (last visit at least visit 17). This categorization resulted in four outcome categories, that is, LED dropouts (LED DO), early weight maintenance dropouts (early WM DO), late weight maintenance dropouts (late WM DO), and completers (Figure 2). These specific categories enabled a more precise analysis of the participants as the difference between the weight loss phase and weight maintenance is rather extreme.

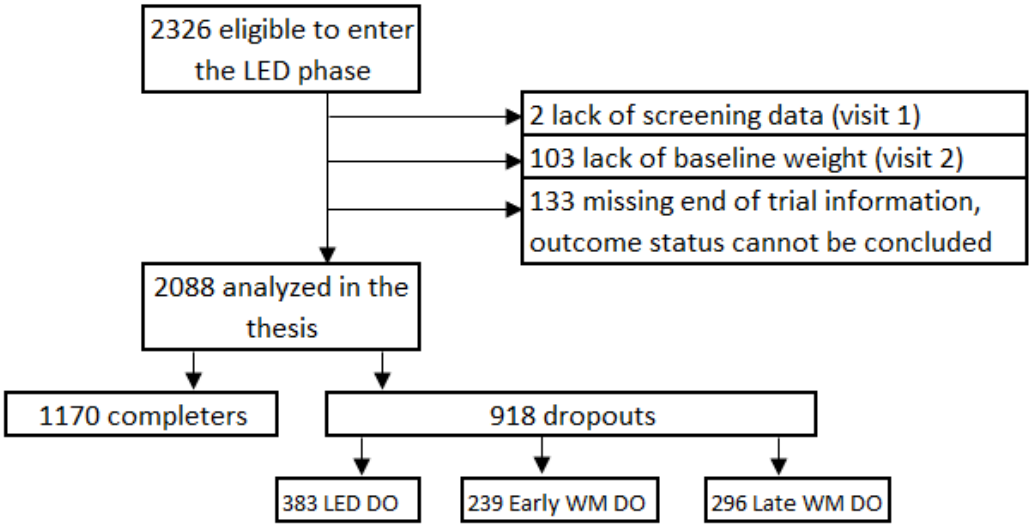

Figure 2. Flow chart of the participants analyzed in this thesis.

4.3 Methods

4.3.1 Statistical analysis

The statistical analyses were computed with IBM Statistical Package for Social Sciences (SPSS, version 24). The normality of all continuous variables was examined using normality tests. The distribution of non-normal continuous variables was examined using Mann-Whitney U test for the dichotomous outcome status of participants and Kruskal-Wallis test for the more precise outcome statuses. The frequency of categorical variables across all outcome statuses was examined using Pearson Chi-Square test.

The prediction model for the dichotomous outcome was carried out with logistic regression model. This was done first separately for individual variables after which all significant variables were examined together. For the multivariate model, two methods were used; 1) the enter method in which all selected variables were forced into the model, and 2) the forward stepwise method with likelihood ratio in which only significant variables remain in the model.

4.3.2 Imputation of missing weight data

For the analysis of weight maintenance success, missing weight data was imputed. Imputation was only carried out for missing weight data at visit 17. Participants included in the weight maintenance phase, that is, early WM DOs, late WM DOs and completers, were considered here. Five different methods of imputation were used. Thus, five different variables for weight loss percent at year two ($WL_{2\text{-YEAR}}$) were calculated with weights at the baseline (visit 2) and at visit 17. These new imputation-based variables reflected the overall weight loss success of the participants.

In baseline observation carried forward (BOCF) the baseline weight (visit 2 weight) was imputed for all those who were missing visit 17 weight. In last observation carried forward (LOCF), the last observed weight was imputed for all missing visit 17 weights. In the two LOCF+ methods, two different weight gains were expected: 1) 0.2 kilos per month absent (Teixeira et al. 2004), and 2) 0.3 kilos per month absent (Fabricatore et al. 2009). In these methods, weight gain was calculated by multiplying the kilos indicated by the months the participant was absent, and weight at visit 17 was calculated as the sum of the participants last observed weight and the weight gain.

The last method was data-based. In this method, the average observed weight loss percent (in the maintenance phase) was calculated from existing data for each group (early WM DO, late WM DO and completers). The imputed weight was calculated by multiplying last observation of any given participant by the group's average weight loss percent. Thus, five imputation-based variables were created: $WL_{2\text{-YEAR BOCF}}$, $WL_{2\text{-YEAR LOCF}}$, $WL_{2\text{-YEAR LOCF+0.2}}$, $WL_{2\text{-YEAR LOCF+0.3}}$, and $WL_{2\text{-YEAR DATA-BASED}}$.

5 Results

5.1 Baseline characteristics

5.1.1 Participant statuses

Altogether 2088 participants were included in the analysis. The number of dropouts was 918, making the dropout rate 44 % at year two. Eighteen percent of participants were LED DOs, 11 % were early WM DOs and 14 % were late WM DOs (Figure 3).

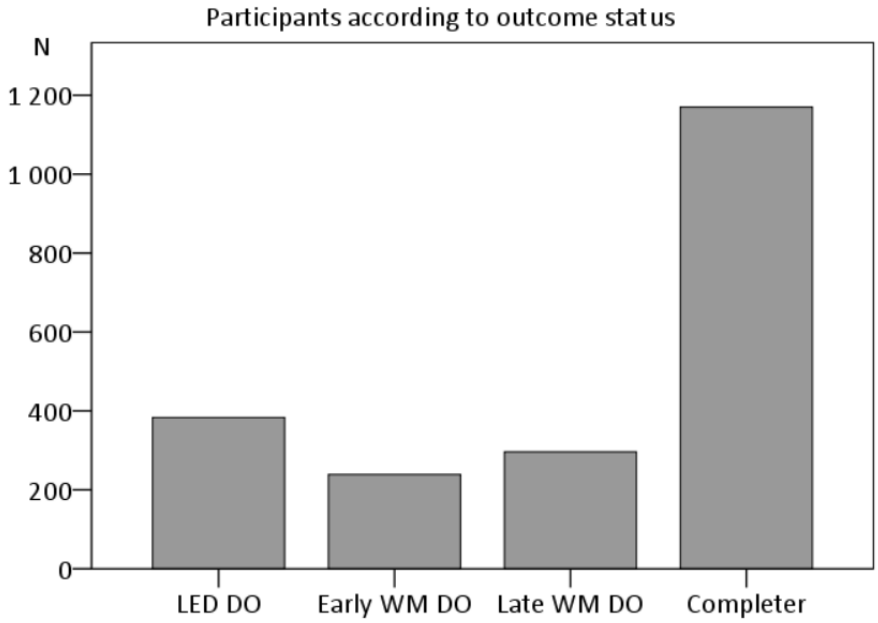

Figure 3. Participants according to their outcome status.

5.1.2 Socio-demographic features

There were some missing values in certain variables, and thus the number of respondents ($N_{variable}$) is indicated only if it deviates from 2088. The baseline characteristics are presented in tables in Appendix 3.

About two thirds (67 %) of participants were women. The age at the start of trial was 52 ± 11.5 years (Figure 4) with a range from 25 to 70 years. The gap between the ages from 46 to 54 stemmed from the original inclusion criteria. The inclusion criteria for

age was from 18 to 45 and from 55 to 70 years for about first two years, after which ages from 46 to 54 were also included.

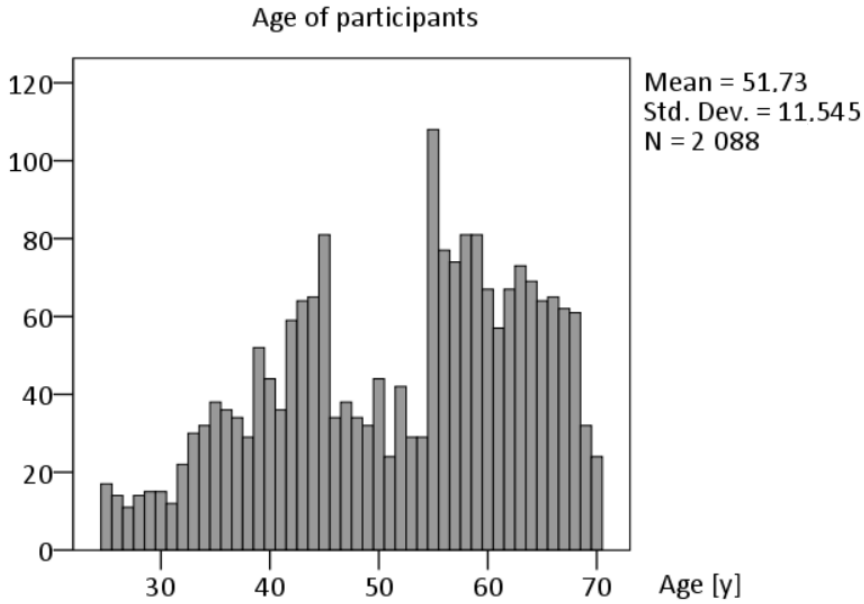

Figure 4. The age histogram of the participants.

Most participants were Caucasian (88 %). Participants who identified themselves as Pacific Islanders or Polynesian or as both are referred to as Pacific-Polynesian, and participants who identified themselves as Maori or Maori and another ethnicity are referred to as Maori/Maori-mixed. Besides Caucasians each of the remaining ethnicity represented on average 2 % of participants. That is, there were 60 Asian, 47 Pacific-Polynesian, 44 Maori-Maori-mixed, 41 Hispanic, 37 Black, and 5 Arabic participants. In addition, 25 participants were from other diverse and mixed ethnicities, referred to as 'other'.

The participants came from eight countries: Denmark (N = 353), New Zealand (N = 285), Bulgaria (N = 280), Finland (N = 276), Spain (N = 269), the UK (N = 254), the Netherlands (N = 202), and Australia (N = 169).

Thirty-eight percent had a university degree, 18 % had higher vocational education, 19 % had secondary vocational education, 14 % had secondary school, 2 % had

primary/junior school, less than 1 % had no formal education, and 4 % had education indicated as 'other'. About three percent were missing, $N_{education}$ = 2003.

There were 12 categories for employment status, and only the most significant ones are mentioned here. Sixty-one percent were in paid-work, 18 % were wholly retired, 5 % were unemployed, 4 % were looking after the home, and entrepreneurs, 'doing something else' and in full-time education each represented 2 %. Less than three percent were missing, $N_{employment}$ = 2028.

Sixty-one percent of participants were married, 12 % were single (never married), 11 % were divorced, 7 % were in civil partnership, 4 % were widowed, 2 % were separated, and 2 % were category 'other'. About two percent were missing, $N_{marital\ status}$ = 2043.

5.1.3 Biometric characteristics

The biometric variables examined included baseline BMI, baseline weight, height, weight loss (kg), WL_{LED} and diabetes risk score (DRS). The baseline characteristics of all participants are shown in Table 11. The baseline weight was 100 ± 21.3 kg, height was 168 ± 9.3 cm, baseline BMI was 35 ± 6.5, weight loss was 11 ± 3.8 kg, WL_{LED} was 11 ± 3.1 %, and DRS was 16 ± 3.6. The minimum weight loss (minus 5 kg) represented a five-kilo weight gain in the LED phase and was an example of serious non-adherence to the program. Baseline weight, BMI, WL_{LED} and DRS were not normally distributed.

Table 11. Selected baseline characteristics and weight loss in the LED phase.

Variables	N	Min	Max	Mean	SD
Weight [kg]	2088	58	238	100	21.3
Height [cm]	2088	139	198	168	9.3
BMI [kg/m²]	2088	25	77	35	6.5
Weight loss [kg]	1887	-5	25	11	3.8
WL_{LED} [%]	1887	-6	20	11	3.1
Diabetes risk score	2076	4	65	16	3.6

5.2 Characteristics associated with dropout

The characteristics statistically significantly associated with dropout were age, baseline BMI, WL_{LED}, DRS, ethnicity, country, employment status, education, marital status, smoking, alcohol use and $WL_{2\text{-}YEAR}$. Variables that were not statistically significantly associated with dropout are not presented. These included gender, three variables related to previous weight loss attempts and two variables related to amount of alcohol used. All original and modified variables according to both status categories (dichotomous and precise) are presented in tables in Appendix 4.

The results of the precise outcome categories (LED DO, early WM DO, late WM DO and completer) are not discussed in this thesis. In sum, the distributions of the continuous variables were not the same between the precise groups (Kruskal-Wallis test $p < 0.001$ for age, baseline BMI and WL_{LED}, and $p < 0.05$ for DRS). In addition, the frequencies of the categorical variables were not the same across the precise groups (Pearson Chi-Square test $p < 0.001$ for ethnicity, country, education, marital status, smoking and alcohol use, and $p < 0.05$ for employment status). This indicates that even across specific dropout categories there are differences between participants. However, the precise categories are not further discussed here as the main focus of this thesis was the prediction of dropout as a dichotomous variable (see chapter 5.3). Precise dropout categories should be analyzed in more detail in future studies.

5.2.1 Continuous variables

Age, baseline BMI, weight $loss_{LED}$ and diabetes risk score

The distribution of age, baseline BMI, WL_{LED} and DRS was not the same across completers and dropouts (Mann-Whitney U test $p < 0.001$ for all others except for DRS $p < 0.05$). Dropouts were statistically significantly younger than the completers (49 ± 12.2 vs. 54 ± 10.4, $p < 0.001$), had a higher baseline BMI (37 ± 7.3 vs. 34 ± 5.6, $p < 0.001$), lost less weight in the LED phase (9 ± 3.7 vs. 12 ± 2.3, $p < 0.001$), and had a lower DRS (16 ± 3.3 vs. 16 ± 3.8, $p < 0.05$) than the completers.

Age, baseline BMI, WL_{LED} and DRS were statistically significantly associated with dropout in logistic regression. All variables were analyzed in separate models, and the results are combined in Table 12. The Nagelkerke R^2 for dropout prediction varied from 0.4 % (DRS) to 17.4 % (WL_{LED}), that is, out of the examined four variables the weight loss percent achieved in the LED phase was the strongest predictor of dropout.

For every one-percent increase in WL_{LED}, the odds for dropout decreased 24.3 % ($p < 0.001$), that is, the likelihood of dropping out was less the more weight was lost. For every one-year increase in age, the odds for dropout decreased 4.1 % ($p < 0.001$), that is, the likelihood of dropping out was lesser the older the participant was. For every one-unit increase in baseline BMI, the odds for dropout increased 6.7 % ($p < 0.001$), that is, the likelihood of dropping out was greater with greater baseline BMI. And last but not least, for every one-unit increase in DRS, the odds for dropout decreased 3 % ($p < 0.05$), that is, the likelihood of dropout was lower with higher DRS.

Table 12. Explaining dropout with age, baseline BMI, WL_{LED} and DRS in separate logistic regression models.

Variables in the equation, analyzed individually	Nagelkerke R^2	B	S.E.	Wald	df	p	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
WL_{LED}	0.174	-0.279	0.020	200.094	1	0.000	0.757	0.728	0.787
Age	0.072	-0.042	0.004	109.812	1	0.000	0.959	0.951	0.966
Baseline BMI	0.053	0.065	0.007	76.887	1	0.000	1.067	1.051	1.082
Diabetes risk score	0.004	-0.031	0.013	5.775	1	0.016	0.970	0.946	0.994

5.2.2 Categorical variables

Ethnicity, country, employment, education, marital status, smoking and alcohol use

A few variable modifications were computed: 1) Ethnicity: Due to the fact that most participants were Caucasian (1829 vs. 259) and due to similarity in the dropout behavior of the other ethnicities, Caucasians were compared to non-Caucasians. 2) Employment status: Due to the large number of different levels of employment, the variable was analyzed according to three groups: in paid work/entrepreneurs, partly/wholly retired, and 'other'. 3) Marital status: Since the majority of participants were married (1269 vs. 774) and due to similarity in the dropout behavior of the other statuses, married participants were compared to 'others'.

All variables were analyzed in separate logistic regression models, and the results are combined in Table 13. The Nagelkerke R^2 for predicting dropout varied from 1.0 % (ethnicity) to 9.9 % (country). The frequency of none of the variables was the same across completers and dropouts (for all variables Pearson Chi-Square test $p < 0.001$).

Fifty-eight percent of Caucasians and 45 % non-Caucasians were completers, that is, more than half of Caucasians were completers and more than half of non-Caucasians were dropouts. Compared to Caucasians, non-Caucasians had 68 % higher odds to drop out ($p < 0.001$).

As for variable country, 75 % of Finns, 73 % of Australians, 66 % of Dutch, 56 % of New Zealanders, 56 % of Brits, 54 % of Danes, 48 % of Spaniards, and 30 % of Bulgarians were completers. In the logistic regression model, participants in Bulgaria had 6.4, in Spain 2.9, in Denmark 2.4, in New Zealand 2.2 and in the UK 2.2 times higher odds for dropout compared to Australia (all with $p < 0.001$). Finland and the Netherlands did not show any significant difference to Australia.

Fifty-seven percent of participants in paid work/entrepreneurs, 66 % of wholly/partly retired and 45 % of category 'other' were completers. Compared to participants in paid

work/entrepreneurs, participants wholly/partly retired had 32 % lower odds ($p < 0.05$), and 'others' had 63 % ($p < 0.001$) higher odds for dropout.

As for education, for example, 65 % of participants with higher vocational, 58 % with university education, 49 % with secondary school and 33 % with no formal education were completers. In the logistic regression model, participants with higher vocational education had 24 % lower odds, secondary vocational education had 31 % higher odds, and secondary school education had 46 % higher odds for dropout compared to university education (all with $p < 0.05$). The other education levels did not show statistically significant difference to university education.

Sixty-one percent of married participants and 48 % of participants not married were completers. Compared to married participants, the 'others' had 71 % higher odds for dropout ($p < 0.001$).

Fifty-nine percent of non-smokers, 50 % of participants who smoke less than weekly and 35 % of participants who smoke at least daily were completers. Compared to non-smokers, participants smoking at least daily had 2.6 times higher odds for dropout ($p < 0.001$). Less than weekly smokers showed no difference to non-smokers.

Sixty percent of participants who use alcohol and 48 % of abstinent participants were completers. Compared to participants who reported not to use alcohol, alcohol users had 38 % lower odds for dropout ($p < 0.001$). That is, according to this data smokers and non-drinkers were more likely to drop out.

Table 13. Explaining dropout with ethnicity, country, employment, education, marital status, smoking and use of alcohol in separate logistic regression models.

Variables in the equation, analyzed individually	Nagelkerke R^2	B	S.E.	Wald	df	p	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Ethnicity (Non-Caucasian)	0.010	0.517	0.134	14.957	1	0.000	1.677	1.290	2.178
Country (base = Australia)	0.099			144.438	7	0.000			
Bulgaria		1.861	0.217	73.224	1	0.000	6.430	4.198	9.847
Denmark		0.860	0.204	17.757	1	0.000	2.364	1.584	3.527
Finland		-0.104	0.223	0.219	1	0.640	0.901	0.582	1.395
New Zealand		0.781	0.211	13.703	1	0.000	2.184	1.444	3.302
Spain		1.081	0.213	25.847	1	0.000	2.946	1.943	4.469
The Netherlands		0.335	0.229	2.143	1	0.143	1.398	0.893	2.191
The United Kingdom		0.792	0.215	13.577	1	0.000	2.208	1.449	3.366
Employment status (base = In paid work…)	0.022			32.485	2	0.000			
Wholly/partly retired		-0.385	0.122	10.034	1	0.002	0.680	0.536	0.863
Other		0.489	0.123	15.817	1	0.000	1.631	1.282	2.075
Education (base = Univ…)	0.017			25.079	6	0.000			
Higher vocational educ.		-0.279	0.129	4.679	1	0.031	0.757	0.588	0.974
Secondary vocational educ.		0.267	0.123	4.695	1	0.030	1.306	1.026	1.664
Secondary school		0.377	0.136	7.703	1	0.006	1.458	1.117	1.903
Primary/Junior school		0.133	0.296	0.200	1	0.654	1.142	0.639	2.040
No formal education		1.030	0.869	1.406	1	0.236	2.802	0.510	15.390
Other		0.314	0.229	1.884	1	0.170	1.369	0.874	2.142
Marital status (Other)	0.022	0.538	0.092	34.041	1	0.000	1.712	1.429	2.051
Smoking (base = No)	0.027			40.406	2	0.000			
Yes, but less than weekly		0.361	0.244	2.189	1	0.139	1.434	0.889	2.313
Yes, at least daily		0.961	0.153	39.211	1	0.000	2.613	1.935	3.530
Drink alcohol (Yes)	0.016	-0.479	0.095	25.421	1	0.000	0.619	0.514	0.746

5.2.3 Imputed variables

Only the logistic regression model for imputation-based variables is considered in this section. Results of the $WL_{2\text{-YEAR}}$ variables according to different statuses can be found in Appendix 4. All variables generated from imputed weight data were analyzed separately and the results are combined in Table 14. For more information on the imputation methods see chapter 4.3.2. The number of participants analyzed here was $N_{WL\ 2\text{-YEAR}} = 1705$.

With LOCF ($WL_{2\text{-YEAR LOCF}}$) and the data-based ($WL_{2\text{-YEAR DATA-BASED}}$) methods, it seemed that participants who lost more weight had higher odds to drop out (42 % and 57 %, respectively, both with $p < 0.001$). These imputation methods cannot be considered just and logical as both overestimate the weight maintenance success of dropouts. Typically, after an energy-restricted-diet weight tends to increase with time. Since the participation of dropouts is cut short, their last observed weight seems lower than that of the completers. The interpretation of the data-based method is the same, as it is merely a modification of the LOCF method. In conclusion, these methods overestimate the weight maintenance success of dropouts, and thus, these results have to be interpreted with extreme caution.

As for the BOCF ($WL_{2\text{-YEAR BOCF}}$), the results relate to the findings of baseline BMI. As greater baseline BMI predicted increased odds for dropout, so does lower $WL_{2\text{-YEAR BOCF}}$. The more weight a participant lost the less likely they were to drop out. The odds for dropout decreased 27 % ($p < 0.001$) for every one percent increase in weight lost.

According to both LOCF+ methods ($WL_{2\text{-YEAR LOCF+0.2}}$ and $WL_{2\text{-YEAR LOCF+0.3}}$), the less weight a participant lost the more likely they were to drop out. Since these methods might be closer to real life weight development, the results could be considered more reliable than the first two methods discussed. According to 'LOCF+ 0.2 kg/month absent' method, for every one-percent increase in weight loss the odds for dropout decreased

3.2 % ($p < 0.001$). According to 'LOCF+ 0.3 kg/month absent' method, for every one-percent increase in weight loss the odds for dropout decreased 6.9 % ($p < 0.001$).

Table 14. Explaining dropout with five imputed $WL_{2\text{-YEAR}}$ variables in separate logistic regression models.

Variables in the equation, analyzed individually	Nagelkerke R^2	B	S.E.	Wald	df	p	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
$WL_{2\text{-YEAR LOCF}}$	0.021	0.041	0.008	25.067	1	0.000	1.042	1.026	1.059
$WL_{2\text{-YEAR DATA-BASED}}$	0.027	0.056	0.010	32.027	1	0.000	1.057	1.037	1.078
$WL_{2\text{-YEAR BOCF}}$	0.332	-0.311	0.021	216.784	1	0.000	0.733	0.703	0.764
$WL_{2\text{-YEAR LOCF+0.2}}$	0.013	-0.033	0.009	14.946	1	0.000	0.968	0.952	0.984
$WL_{2\text{-YEAR LOCF+0.3}}$	0.057	-0.072	0.009	63.123	1	0.000	0.931	0.914	0.947

In sum, according to two imputation methods the more weight a participant lost the more likely they were to drop out, and according to three imputation methods the more weight a participant lost the less likely they were to drop out. Due to the fact that these conclusions are drawn from fabricated data the results must be considered cautiously. In addition, note that none of the $WL_{2\text{-YEAR}}$ variables are considered in the multivariate logistic model as they are not based on measured data.

5.2.4 Reasons for ending trial

The reasons for ending trial are presented in Figure 5 ($N_{End\ reason}$ = 1798). Besides trial completion (34 %), the most common reasons for ending trial included withdrawal for personal reasons (22 %), lost-to-follow up (11 %), other reasons (9 %), significant non-compliance with the protocol (5 %) and diagnosis of T2D (3 %). The most common reasons in category other were not achieving the required 8 % weight loss, time restrains, medical reasons, moving from the area and pregnancy.

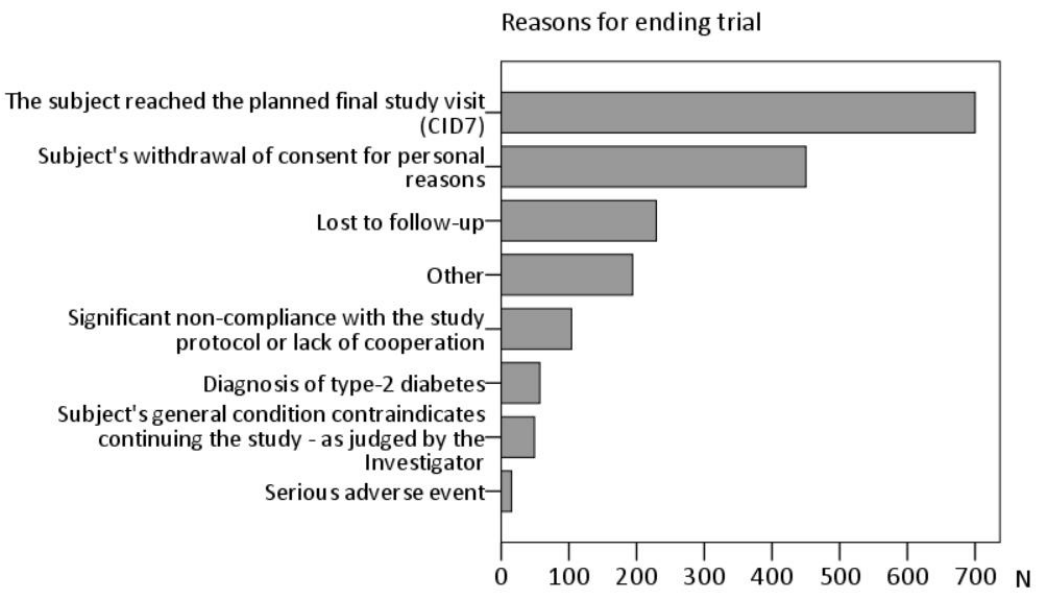

Figure 5. The reasons for ending trial.

5.2.5 Summary of the single explanatory variable models

In conclusion, dropouts were younger, had a higher baseline BMI, lost less weight in the LED phase, had a lower diabetes risk score, were more likely to be non-Caucasian, daily smokers and abstinent from alcohol. Compared to Australians, Bulgarians, Danes, Spaniards, New Zealanders and Brits were more likely to drop out, and there was no significant difference between Australians and Finns, and Australians and Dutch.

As for the employment status, compared to participants in paid work/entrepreneurs, participants wholly/partly retired were less likely, and participants with employment status 'other' were more likely to drop out. Compared to married participants, participants not married were more likely to drop out. In addition, compared to university education, participants with higher vocational education were less likely to drop out and participants with secondary vocational education or secondary school education were more likely to drop out.

The most common reasons for ending trial prematurely were withdrawal for personal reasons, lost-to-follow up and significant non-compliance with the study protocol. Reasons identified as 'other' included not achieving the required weight loss, time restrains, medical reasons, moving from the area and pregnancy.

5.3 Prediction model for dropout

Dropout can be predicted with five variables

The final model computed with forward likelihood ratio method resulted in a model including five predictive variables: age, baseline BMI, WLLED, country and education (Table 15). All these explanatory variables were statistically significant even though some of the individual dummy variables in education were not significant.

The Nagelkerke R^2 for predicting dropout with age, baseline BMI, WLLED, country and education was 29 %. For every one-unit increase in age the odds for dropout decreased 3 %, in baseline BMI increased 4.5 %, and in WLLED decreased 25 % (for all $p < 0.001$). Compared to Australia, the odds for participants to drop out were higher in other countries: Finland and New Zealand 1.8 times ($p < 0.05$), the Netherlands 2.0 times ($p < 0.05$), Bulgaria 3.1 times ($p < 0.001$), the UK 3.3 times ($p < 0.001$), Spain 4.7 times ($p < 0.001$), and Denmark 4.9 times ($p < 0.001$). Compared to university education, participants with secondary vocational education or secondary school had 44 % ($p < 0.05$) and participants with education 'other' had 87 % ($p < 0.05$) higher odds to drop out.

Table 15. The final logistic regression model in which dropout is predicted using age, baseline BMI, WLLED, country and education.

Variables in the Equation		B	S.E.	Wald	df	p	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
	Age	-0.030	0.005	32.192	1	0.000	0.970	0.960	0.981
	Baseline BMI	0.044	0.010	20.913	1	0.000	1.045	1.025	1.064
	WLLCD	-0.288	0.023	151.772	1	0.000	0.750	0.716	0.785
Country	Base = Australia			69.688	7	0.000			
	Bulgaria	1.136	0.282	16.238	1	0.000	3.115	1.792	5.413
	Denmark	1.586	0.268	34.877	1	0.000	4.883	2.885	8.264
	Finland	0.592	0.285	4.304	1	0.038	1.807	1.033	3.161
	New Zealand	0.607	0.273	4.931	1	0.026	1.835	1.074	3.135
	Spain	1.537	0.276	31.111	1	0.000	4.652	2.710	7.984
	The Netherlands	0.676	0.308	4.800	1	0.028	1.965	1.074	3.597
	The United Kingdom	1.207	0.282	18.269	1	0.000	3.343	1.922	5.813
Education	Base = University education			16.941	6	0.010			
	Higher vocational education	-0.124	0.166	0.561	1	0.454	0.883	0.638	1.222
	Secondary vocational education	0.368	0.161	5.216	1	0.022	1.444	1.054	1.980
	Secondary school	0.367	0.173	4.510	1	0.034	1.444	1.029	2.026
	Primary/Junior school	0.279	0.361	0.599	1	0.439	1.322	0.652	2.680
	No formal education	1.386	1.004	1.905	1	0.168	4.001	0.559	28.652
	Other	0.627	0.285	4.841	1	0.028	1.871	1.071	3.270

Compared to the individual analyses of the explanatory variables in the previous sections, the coefficients were generally similar. Variables age, baseline BMI, and WL_{LED} each had the same sign in the B coefficient and were of similar magnitude.

Considering the explanatory variable country, all but one country retained the sign in the B coefficient compared to having country as the single explanatory variable. The sign for B for Finland changed from negative to positive and become significant. In addition, the Netherlands changed from being insignificant to significant. Thus, all countries compared to Australia were significant in predicting dropout in this final multivariate model. The odds for dropout changed most dramatically for Bulgaria (from 6.4 to 3.1), Denmark (from 2.4 to 4.9), and Finland (from 0.9 to 1.8). These changes show that considering multiple explanatory variables in the regression model is important as it changes the interpretation of the individual explanatory variables.

Variable education did not show significant difference compared to having it as the single explanatory variable in the model. All the B coefficients retained their sign and were approximately of similar magnitude. Higher vocational education lost its significance, and thus compared to university education showed no difference in regards to dropout in the final model. Education 'other' become significant, thus showing a difference compared to university education. Hence, the interpretation of education in the multivariate model can be considered more logical than the interpretation of the single explanatory model in which none of the lowest education levels (primary/junior school, no formal education, and 'other') showed difference to university degree but the second highest education did.

A prediction model for dropout with a forced enter method is not discussed here due to its method and is only presented in Appendix 5.

5.4 Country analysis

Due to the changes observed in the interpretation of country variable between the single and multivariate models, and due to great differences in the odds ratios between countries in the multivariate model (Table 16), a more thorough analysis was executed and is presented in this section.

First, between the models (single versus multivariate model) can the other significant variables, i.e., age, baseline BMI, WLLED and education, explain the rather great changes observed in the odds ratios for Bulgaria, Denmark and Finland?

Second, in the final multivariate model, can cumulative disadvantageous features of the other variables, i.e., ethnicity, employment, marital status, DRS, smoking and alcohol use, explain the highest dropout odds (Denmark, Spain, the UK and Bulgaria)?

Table 16. Country coefficients in the single and multivariate model.

Variables in the equation	Country in the single explanatory variable model			Country in the multivariate model		
	B	p	Exp(B)	B	p	Exp(B)
Base = Australia		0.000			0.000	
Bulgaria	1.861	0.000	6.430	1.136	0.000	3.115
Denmark	0.860	0.000	2.364	1.586	0.000	4.883
Finland	-0.104	0.640	0.901	0.592	0.038	1.807
New Zealand	0.781	0.000	2.184	0.607	0.026	1.835
Spain	1.081	0.000	2.946	1.537	0.000	4.652
The Netherlands	0.335	0.143	1.398	0.676	0.028	1.965
The United Kingdom	0.792	0.000	2.208	1.207	0.000	3.343

5.4.1 What explains the changes in dropout odds?

Since the dropout odds for Bulgarians more than halved (from 6.4 to 3.1), and for Danes and Finns more than doubled (from 2.4 to 4.9 and from 0.9 to 1.8) these countries were looked into more precisely. Only the significant variables, that is, age, baseline BMI, WLLED and education, are presented as the remained in the final model.

The means and standard deviations of age, baseline BMI and WL$_{LED}$ according to country are presented in Table 17. The distributions of the variables were not the same across countries (for all Kruskal-Wallis test $p < 0.001$).

According to the model, younger age increased the odds for dropout. Bulgarians were younger than any other nation ($p < 0.001$) except for Spaniards and New Zealanders. This partly explains the decrease in their dropout odds (from 6.4 to 3.1). Danes were statistically significantly older than Spaniards, Bulgarians and New Zealanders. Finns were older than any other nationality ($p < 0.001$) except for the Dutch to whom there was no statistical difference. Older age of Danes and Finns partly explains the increase in their dropout odds (from 2.4 to 4.9) and (from 0.9 to 1.8) moving from the single to the multivariate model.

Higher baseline BMI increased risk for dropping out. Bulgarians had statistically higher baseline BMI than Finns, Dutch, Spaniards and Danes, partly explaining the more than halving of their dropout odds in the multivariate model. Danes had statistically lower baseline BMI only compared to Bulgarians and New Zealanders, and no difference to Australians, Spaniards or Brits. On its own this does not explain the increase in their dropout odds. Finns had the second lowest baseline BMI with the Dutch having the lowest and with no significant difference to the Spaniards. Low baseline BMI partly explains the increase in dropout odds for Finns.

WL$_{LED}$ had the strongest effect on dropout with one-unit increase decreasing dropout odds by 25 % in the multivariate model. Bulgarians had the lowest WL$_{LED}$ compared to any other nationality ($p < 0.001$) and this variable may on its own explain the halving of their dropout odds between the models. Danes had the second highest WL$_{LED}$ and thus this might explain the increase in their dropout odds. Finns had higher weight loss only compared to Bulgarians and Dutch, that is, this would not on its own explain the increase in the dropout odds for Finns.

Table 17. Mean and standard deviation of age according to country.

Country		Age [y]		Baseline BMI [kg/m²]			WLED [%]	
	N*	Mean	SD	Mean	SD	N**	Mean	SD
Australia	169	53	10.5	36	6.8	152	11	2.6
Bulgaria	280	47	11.9	37	8.3	237	7	3.5
Denmark	353	54	10.7	35	5.2	344	12	2.5
Finland	276	58	9.0	33	4.6	260	12	2.1
New Zealand	285	47	11.3	39	7.5	246	11	2.7
Spain	269	47	10.3	34	6.2	244	12	2.7
The Netherlands	202	57	10.0	32	4.1	187	11	3.1
The United	254	52	12.1	35	5.7	217	12	3.0
Total	2088	52	11.5	35	6.5	1887	11	3.1

* for age and baseline BMI

** for WLED

The frequency of different educations across countries is presented in Table 18. The frequencies were not the same across countries (Pearson Chi-Square test $p < 0.001$). Compared to university education, the odds for dropping was higher for participants with secondary vocational, secondary school and 'other' education in the multivariate model.

Bulgaria had more university educated participants than any other country except Australia, and thus an increase in their dropout odds would have be expected when moving from the single explanatory variable model to the multivariate model. However, as the opposite occurred, education does not explain the change in the odds ratio for Bulgaria. In contrast, Denmark had the lowest level of university educated participants compared to any other country except to the UK and the Netherlands, and thus a decrease in their dropout odds would have be expected. Again however, as the opposite occurred, education cannot explain the change in the odds ratios for Danes. The percentage of university educated Finns equaled the average of all participants, and thus the increase in their dropout odds is not explained by education. In the multivariate model the significance of education as a whole was lower than that of the other explanatory variables ($p < 0.05$ versus $p < 0.001$), hence it does not have equal power to explain the changes observed in the odds ratios.

Table 18. Education levels percentagewise according to country.

Education	Australia	Bulgaria	Denmark	Finland	New Zealand	Spain	The Netherlands	The United Kingdom	Total
University education	59	65	20	39	49	38	17	30	39
Higher vocational education	20	6	22	18	14	14	32	31	19
Secondary vocational education	4	20	36	30	8	21	22	8	20
Secondary school	15	7	6	10	24	19	12	29	15
Primary/Junior school	0	1	5	2	0	5	6	0	2
No formal education	0	0	1	0	1	0	0	0	0
Other	2	1	11	1	3	2	12	2	4
Total	100	100	100	100	100	100	100	100	100

In sum, age partly explained the changes in the odds ratios between the models for Bulgarians, Danes and Finns. Baseline BMI partly explained only the changes for Finns and Bulgarians, but not for Danes. WLLED might on its own have explained the changes observed for Bulgarians, and partly for Danes, however, could not account for the changes observed for Finns. Education could not explain any of the changes observed in the odds ratios, as the directions of the changes were opposite to what occurred.

5.4.2 What explains the highest dropout odds?

The possible cumulation of disadvantageous features in ethnicity, employment, marital status, DRS, smoking and alcohol use are presented in this section (Tables 19 and 20). Age, baseline BMI, WLLED and education remained in the final model and thus their confounding effect was accounted for. Thus, they are not analyzed in this section.

The frequency of ethnicity, employment, marital status, smoking and alcohol use was not the same (for all Pearson Chi-Square test $p < 0.001$), and the distribution of DRS was not the same (Kruskal-Wallis test $p < 0.001$) across countries. The analysis in this section emphasizes the highest odds for dropout, that is, Denmark (4.9), Spain (4.7), the UK (3.3) and Bulgaria (3.1).

The odds for dropping was higher for non-Caucasians. The percentage of non-Caucasians might partly explain the odds for the UK and Spain but certainly not for Denmark and Bulgaria. Also, it is noteworthy that almost half of the participants from New Zealand were non-Caucasians, however, the odds for New Zealanders to drop out were one of the lowest (1.8).

Compared to working or retired, participants with employment status ‘other’ were more likely to drop out. The share of ‘others’ might partly explain the dropout odds for Bulgaria and Spain, however, is unlikely to explain that of Denmark and the UK. Again, it is noteworthy that the highest percentage of ‘others’ was observed in the Netherlands, which had one of the lowest (2.0) dropout odds.

The odds for dropping out was higher for participants who were not married. The higher percentage of participants not married might partly explain the odds for Bulgaria and Denmark. However, this did not seem to be the case for Spain and the UK.

Smoking at least daily increased the odds for dropping out. The percentage of non-smokers was about or above 90 % for all other countries except for Bulgaria and Spain. Thus, this might partly explain their dropout odds, however, it does not explain the odds for Denmark and the UK.

Abstaining from alcohol increased the odds for dropout. Spain and Bulgaria had the highest number of non-alcohol users, 58 % and 42 % respectively, which might partly explain their higher odds for dropout. However, Denmark and the UK had lower than average number of non-alcohol users, which does not explain their dropout odds.

Table 19. Ethnicity, employment status, marital status, smoking and alcohol use percentagewise according to country.

Variables		Australia	Bulgaria	Denmark	Finland	New Zealand	Spain	The Netherlands	The United Kingdom	Total
Ethnicity	Caucasian	83	100	98	100	54	85	99	81	88
	Non-Caucasian	17	0	2	0	46	15	1	19	12
	Total	100	100	100	100	100	100	100	100	100
Employment status	In paid work/entrepreneur	69	65	67	54	79	70	44	64	65
	Wholly/partly retired	19	15	22	36	5	8	26	21	19
	Other	11	20	11	10	16	21	31	15	17
	Total	100	100	100	100	100	100	100	100	100
Marital status	Married	66	55	53	67	62	66	74	61	62
	Other	34	45	47	33	38	34	26	39	38
	Total	100	100	100	100	100	100	100	100	100
Smoking	No	96	66	89	89	91	82	91	94	87
	Yes, but less than weekly	2	2	5	5	2	6	4	1	3
	Yes, at least daily	2	32	7	7	7	12	4	5	10
	Total	100	100	100	100	100	100	100	100	100
Alcohol use	No	25	42	25	19	36	58	24	26	32
	Yes	75	58	75	81	64	42	76	74	68
	Total	100	100	100	100	100	100	100	100	100

The mean DRS according to country is presented in Table 20. Lower DRS increased the odds to drop out. Brits had lower DRS than any other nationality except for the New Zealanders, which might partly explain their dropout odds. DRS did not explain the dropout odds for Bulgaria, Spain and Denmark as their DRS was close to the average. In the pairwise comparison, these countries showed no difference to three other countries. Again, it is noteworthy that the Dutch had the lowest DRS, however, they had one the lowest dropout odds (2.0).

Table 20. Means and standard deviations of DRSs according to country.

Country	N	Mean	SD
Australia	169	17.5	3.4
Bulgaria	269	15.7	2.2
Denmark	353	16.1	3.2
Finland	276	18.1	4.2
Spain	285	15.2	2.8
New Zealand	269	16.6	3.3
The Netherlands	201	11.5	3.4
The United Kingdom	254	14.9	2.3
Total	2076	15.8	3.6

In sum, there was little evidence that the variables analyzed here explain the higher dropout odds for Denmark, Spain, the UK and Bulgaria. The share of non-Caucasians and having lower DRS might have explained some small part of the odds (3.3) for Brits. The percentage of non-Caucasians, lower DRS, higher number of smokers, lower number of non-alcohol users and having employment status 'other' might have partly explained the higher dropout odds (4.7) for Spaniards. And smoking, employment status, marital status, alcohol use and DRS might have partly explained the higher dropout odds (3.1) for Bulgarians. However, the highest odds for dropping out, 4.9 for Danes, cannot be explained with these variables as the only 'disadvantage' was seen in marital status. This analysis goes to show that there is much more to be investigated for a better understanding of the differences in dropout across countries.

5.5 Recommendations for further studies

The country samples in this study were not entirely comparable. The characteristics of the samples, be it age, education, marital status etc., were a) not evenly reflective of the populations of the individual countries, and b) not the same between countries. That is, there is a possibility that the country samples were either cumulated with beneficial characteristics associated with increased retention or disadvantageous characteristics associated with higher dropout. However, the statistical methods used in the analysis took into account the effects of multiple explanatory variables, and as such the study of significance of the other explanatory variables not included in the final model was analyzed. These variables would have been included in the final model had they provided significant explanatory power; however, it is known that in stepwise modelling the resulting model depends on the order in which the individual explanatory variables are included to and removed from the model. Thus, there remains some uncertainty in the final model resulting from the stepwise procedure.

The differences observed between countries in age, education, baseline BMI and WL$_{LED}$ were accounted for in the final multivariate model as they remained statistically significant and were thus included in the model. The possible cumulative effects of the

other variables, not significant enough to be incorporated in the regression model, was also analyzed. However, this analysis showed no clear explanation to the differences observed in the odds for dropping out between countries. This suggests that further analysis of participants, countries and sites is warranted.

As comparing countries, sites or staff was not the main focus of this thesis, no in-depth analysis of these areas was carried out. However, as Baekeland and Lundwall wrote in 1975, dropping out results from three vectors: 1) participant, 2) treatment personnel, and 3) environmental variables. Thus, truly understanding the phenomenon of dropout would require the analysis of all three vectors (see chapter 2.2.3). Further analysis should include the study of a) the cumulation of other disadvantageous features affecting dropout, b) psychological, social and behavioral variables of the participants, c) cultural differences between countries, d) other environmental factors, and e) paradata on sites and staff. Here paradata refers to, for example, data on flow and means of recruitment, accessibility of study site, number of participants per staff, and socio-demographic and behavioral variable of staffs, such as age, education, attitudinal issues and preciseness of following protocol.

6 Discussion

6.1 Findings versus literature

6.1.1 Five variables predict dropout

Age, baseline BMI, WLLED, country and education could be used for predicting dropout, as these explanatory variables remained significant in the final multivariate model. The Nagelkerke R^2 for predicting dropout with these variables was 29 %. The reference of this section includes the 21 studies found to be comparable to data used in this thesis (see chapter 2.4), and the pooled results of two systematic reviews (Leung et al. 2017, Moroshko et al. 2011).

In this thesis younger age predicted dropout. This is in line with the results of four similar studies (Andersson and Rössner 1997, Goode et al. 2016, Jiang et al. 2016, Fabricatore et al. 2009) and with 43 % of the pooled systematic reviews (Leung et al. 2017, Moroshko et al. 2011). On the other hand, half of the pooled studies found no association between age and dropping out, and two studies found that older age predicted dropout. According to this cumulative evidence it seems likely that a) either age does not predict dropout, or b) younger age does but only in certain samples. Confounding factors, participant recruitment regime, study design and data analysis might affect the inconsistency seen in the results.

In this thesis higher baseline BMI increased the odds to drop out. Higher baseline BMI/weight was also associated with attrition in six similar studies (Absetz et al. 2007, Goode et al. 2016, Hadžiabdić et al. 2014, Roumen et al. 2007, Roumen et al. 2011, Teixeira et al. 2004) and in 23 % of the pooled systematic reviews (Leung et al. 2017, Moroshko et al. 2011). On the other hand, most studies (67 %) in the pooled analysis found no association between baseline BMI/weight and dropping out, and 11 % found that lower baseline BMI was associated with dropout. Thus, it is still unclear if baseline BMI can be used universally as a dropout predictor. According to the results of this study and others, it seems that baseline BMI might predict dropout in at least studies in which BMI is an issue, that is, in weight loss and lifestyle interventions aimed at

overweight and obese individuals. Again, there are a lot of confounding factors, such as age, SES, and study design.

An increase in WLLED was found to be associated with decreased odds to drop out. Lower initial weight loss or treatment response was also associated with dropout in the study of Fabricatore et al. (2009) and in 88 % of the pooled reviews (Leung et al. 2017, Moroshko et al. 2011). Only one study in the systematic review by Moroshko et al. (2011) found that there was no association between the variables. Thus, there seems to be very convincing cumulative evidence that lower initial weight loss or poorer early treatment response does in fact predict increased dropout.

Some of the lower level educations were associated with increased odds for dropout, which is in line with two reviewed studies (Goode et al. 2016, Hadžiabdić et al. 2014) and with 29 % of the of the pooled reviews (Leung et al. 2017, Moroshko et al. 2011). However, most studies (71 %) found no association between education and dropout (Leung et al. 2017, Moroshko et al. 2011). According to this, lower education may predict dropout in certain settings only, and more investigation into education and other socio-economic variables is warranted. Education level is the most important socio-economic status indicator, as it is not prone to change unlike job titles and salaries. The share of individual SES classes and the magnitude of difference between the classes vary greatly across countries. Thus, certain countries are not comparable to one another as they reflect very differently constructed societies. The predictive power of such a variable might be difficult to assess.

According to the present study, compared to Australians, dropping out was more likely in ascending order for Finns, New Zealanders, Dutch, Bulgarians, Brits, Spaniards and Danes. Leung et al. (2017) and Moroshko et al. (2011) did not report on country differences and none of the comparable studies reviewed in this thesis included country comparison. Most weight loss and lifestyle intervention studies are conducted on a national level, and thus comparable studies for this section could not be found.

6.1.2 Reasons for ending trial

Besides trial completion, the most common reasons for ending trial included withdrawal for personal reasons, lost-to-follow up, 'other' reasons, significant non-compliance with the protocol and diagnosis of T2D. The most common reasons in category 'other' were not achieving the required 8 % weight loss, time restrains, medical reasons and moving from the area.

Personal reasons were also reported to be common in the studies of Andersson and Rössner (1997), Sakane et al. (2011), Teixeira et al. (2004) and Tuomilehto et al. (2001). Eriksson et al. (1991), Ramachandran et al. (2006), Sakane et al. (2011) and Tuomilehto et al. (2001) mentioned loss of contact as one of the most common reasons. Refusal to attend (Andersson and Rössner 1997, Mensink et al. 2003, Pan et al. 1997, Ramachandran et al. 2006), and unknown reasons (Andersson and Rössner 1997, Hadžiabdić et al. 2014, Michelini et al. 2014, Roumen et al. 2007 and 2011) were mentioned in four studies.

As in this study, medical reasons were found to be a very common reason for ending trial in other studies. Ten out of thirteen studies reported on medical reasons (Eriksson et al. 1991, Hadžiabdić et al. 2014, Liao et al. 2002, Mensink et al. 2003, Michelini et al. 2014, Oldroyd et al. 2006, Roumen et al. 2007, Sakane et al. 2011, Teixeira et al. 2004, Tuomilehto et al. 2001).

Lack of time was mentioned in two (Roumen et al. 2007 and 2011, Teixeira et al. 2004) and moving from the area in four of the comparable studies (Andersson and Rössner 1997, Eriksson et al. 1991, Pan et al. 1997, Sakane et al. 2011).

That is, the reasons for ending trial prematurely in the PREVIEW trial seem to be in line with earlier studies.

6.2 Hypotheses versus results

The first hypotheses stated that 'dropouts are more likely younger, men, less educated, single, unemployed, of ethnic minority, and smoke more than the completers'. It was also hypothesized that multiple previous weight loss attempts would predict earlier withdrawal. The second hypotheses stated that 'completers lose more weight during the low-energy diet phase and maintain their weight more successfully in the weight maintenance phase than the dropouts'.

In the single explanatory variable model dropouts were more likely to be younger, to some level less educated, not married, unemployed (rather than working or retired), non-Caucasian (in most of the countries studied this is equal to ethnic minority), and daily smokers. However, no association between gender and previous weight loss attempts and dropping out was found. Losing less weight in the dieting phase also predicted dropout. According to imputed weight data, it seemed that the less success participants had in maintaining their reduced weight the more likely they were to drop out. In the multivariate model younger age, higher baseline BMI, lower education and poorer weight loss during LED predicted dropout.

In conclusion, to most parts the hypotheses were proven valid.

6.3 Justification and limitation of studied variables

The limitations of this study include the lack of investigation of psychological and social variables. In addition, out of behavioral variables, only previous dieting, smoking and alcohol use were analyzed. That is, some data collected in PREVIEW, such as quality of life, perceived stress, mood, self-efficacy, outcome expectation, motives, food and physical inactivity temptations, and social support for diet and exercise, were not analyzed. This was due to three main reasons: 1) One research center in the study group (Stuttgart, Germany) is assigned to analyze psychological and behavioral features due to their specific expertise, 2) limited economic and time resources of this thesis work, and 3) due to justification of selected variables in the following.

The justification of analyzing mostly socio-demographic and weight variables was that this data could be collected rather easily, objectively and affordably. For the collection and analysis of the variables in this thesis, there was no need for expensive equipment, laboratory testing, highly trained staff or professional pre-analysis. That is, these types of variables can be collected in almost any setting, and in any country. Since obesity and obesity-related diseases are increasing worldwide, a set of affordably, objectively and easily collectable and analyzable variables predictive of dropout in lifestyle interventions is needed. However, variables that require self-reporting, for example, education, smoking and alcohol use, are always subject to untruthful reporting and might bias results.

6.4 Strengths and weaknesses

The strengths of this study include the number of participants and countries, the duration of the trial, and the straightforwardness of the dropout prediction model. The variables used in this thesis could be collected in any setting regardless of a trained staff, and as such preliminary analysis on high dropout risk individuals could be made at the early stages of the trial. This indicates that special attention could be directed at individuals who might not commit to the program.

There are no other studies of similar design addressing multiple country comparison and as such the results of this study are rather unique. Although precise comparability of the country samples is difficult to assess, the results indicate that there is difference between participant retention across countries. However, since the country samples are not entirely reflective of the populations studied, the generalizability of the results is weakened. In addition, due to differences between the samples, no definite conclusions on trial success between countries can be drawn without further analysis.

Related to the above, gender, education, employment status, and age of the participants in this study were not representative of population averages. Gender and employment status were not found to be significant in the final prediction model

unlike age and education. However, it can always be argued that a different sample would produce different results, and thus study repetition and further analysis is warranted.

Psychological, social and environmental variables were not analyzed in this thesis, and only previous dieting, smoking and alcohol use were studied out of behavioral variables. In addition, analysis of site and staff paradata was not conducted. Further analysis on these areas might shed more light on the dropout phenomenon.

Dropout is important to understand since it is the dropout rather than the study completer who will in the long run be the typical patient in the treatment of chronic illness (Baekeland and Lundwall, 1975). Thus, considering public healthcare policies and specific lifestyle programs affecting entire nations, it is worthwhile to emphasize the use of methods which increase retention of the masses and at the same time use specific tools to increase the retention of the dropout risk groups.

The findings of this thesis may also be utilized in the development of design, methods and tools for similar trials to come. In addition, as dropout is not independent of the explanatory variables, the results may also benefit the future analysis of the PREVIEW data, in which these same explanatory variables may be used in modelling.

7 Conclusions

The dropout rate at year two was 44 %. Dropouts were more likely to be younger, have higher baseline BMI, lose less weight in the weight reduction phase, and be less educated. These explanatory variables were statistically significantly associated with dropout, and the results were to most parts in line with previous studies. These results were also in line with the study hypotheses.

Compared to Australians, dropping out was more likely in an ascending order for Finns, New Zealanders, Dutch, Bulgarians, Brits, Spaniards and Danes. Differences between countries persisted even in the multivariate logistic regression model which accounted for a multitude of confounding factors. Investigating the differences between countries was an addition to previous literature.

Poorer weight maintenance success may also predict dropout; however, the analysis was based on imputed values and thus its accuracy is difficult to assess.

The presented analyses show that dropout can be predicted using socio-demographic and weight-related variables which can be collected easily, rather objectively and affordably in any setting. That is, special attention can be directed at early stages to individuals who might not commit to the program offered. However, further investigation of psychological, social, behavioral, treatment, study site and cultural factors is recommended.

References

Absetz P, Valve R, Oldenburg B, Heinonen H, Nissinen A, Fogelholm M, et al.. Type 2 diabetes prevention in the "real world": one-year results of the GOAL Implementation Trial. Diabetes Care. 2007 Oct;30(10): 2465-2470.

Andersson I, Rössner S. Weight development, drop-out pattern and changes in obesity-related risk factors after two years treatment of obese men. Int J Obes Relat Metab Disord 1997;21(3):211-216.

Baekeland F, Lundwall L. Dropping out of treatment: A critical review. Psychol Bull 1975;82(5):738-783.

Becker MH, Drachman RH, Kirscht JP. A new approach to explaining sick-role behavior in low-income populations. Am J Public Health 1974;64(3): 205–216.

Bourn DM, Mann JI, McSkimming BJ, Waldron MA, Wishart JD. Impaired Glucose Tolerance and NIDDM: Does a Lifestyle Intervention Program Have an Effect? Diabetes Care 1994; 17(11): 1311-1319.

Burke LE, Choo J, Music E, Warziski M, Styn MA, Kim Y, et al. PREFER study: A randomized clinical trial testing treatment preference and two dietary options in behavioral weight management — rationale, design and baseline characteristic. Contemporary Clinical Trials 2006;27:34-48.

Burke LE, Styn MA, Glanz K, Ewing LJ, Elci OU, Conroy MB, et al. SMART trial: A randomized clinical trial of self-monitoring in behavioral weight management-design and baseline findings. Contemporary Clinical Trials 2009;30:540-551.

Burke LE, Ewing LJ, Ye L, Styn M, Zheng Y, Music E, et al. The SELF trial: A self-efficacy-based behavioral intervention trial for weight loss maintenance. Obesity 2015;23:2175-2182.

Diabetes Prevention Program Research Group. Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin. N Engl J Med 2002;346(6):393-403.

DiMatteo MR, Lepper HS, Croghan TW. Depression is a risk factor for noncompliance with medical treatment: meta-analysis of the effects of anxiety and depression on patient adherence. Arch Intern Med 2000;160(14):2101-2107.

Eriksson KF, Lindgärde F. Prevention of type 2 (non-insulin-dependent) diabetes mellitus by diet and physical exercise. The 6-year Malmö feasibility study. Diabetologia 1991;34(12):891-898.

Fabricatore AN, Wadden TA, Moore RH, Butryn ML, Heymsfield SB, Nguyen AM. Predictors of attrition and weight loss success: Results from a randomized controlled trial. Behav Res Ther 2009;47(8):685-691.

Fogelholm M, Larsen TM, Westerterp-Plantenga M, Macdonald I, Martinez JA, Boyadjieva N, et al. PREVIEW: Prevention of Diabetes through Lifestyle Intervention and Population Studies in Europe and around the World. Design, Methods, and Baseline Participant Description of an Adult Cohort Enrolled into a Three-Year Randomised Clinical Trial. Nutrients 2017;20;9(6).

Goode RW, Ye L, Sereika SM, Zheng Y, Mattos M, Acharya SD, et al. Socio-demographic, anthropometric, and psychosocial predictors of attrition across behavioral weight-loss trials. Eat Behav 2016;20:27-33.

Hadžiabdić MO, Mucalo I, Hrabač P, Matić T, Rahelić D, Božikov V. Factors predictive of drop-out and weight loss success in weight management of obese patients. J Hum Nutr Diet 2014;28:24-32.

Hochbaum GM. Public Participation in Medical Screening Programs: A Socio-psychological Study. Public Health Service Publication No. 572. Washington, DC: Government Printing Office, 1958.

Janz NK, Becker MH. The Health Belief Model: A Decade Later. Health Education & Behavior 1984;11(1): 1–47 (p. 4).

Jiang L, Yang J, Huang H, Johnson A, Dill EJ, Beals J, et al. Derivation and Evaluation of a Risk-Scoring Tool to Predict Participant Attrition in a Lifestyle Intervention Project. Prev Sci 2016;17(4):461-471.

Kaplan RM, Simon HJ. Compliance in Medical Care: Reconsideration of Self-Predictions. Annals of Behavioral Medicine 1990;12(2):66–71.

Kirscht JP. The Health Belief Model and Illness Behavior. Health Education Monographs winter 1974, pp.387-408.

Kosaka K, Noda M, Kuzuya T. Prevention of type 2 diabetes by lifestyle intervention: a Japanese trial in IGT males. Diabetes Res Clin Pract 2005;67(2):152-162.

Leung AWY, Chan RSM, Sea MMM, Woo J. An Overview of Factors Associated with Adherence to Lifestyle Modification Programs for Weight Management in Adults. Int J Environ Res Public Health 2017;16;14(8).

Liao D, Asberry PJ, Shofer JB, Callahan H, Matthys C, Boyko EJ, et al. Improvement of BMI, body composition, and body fat distribution with lifestyle modification in Japanese Americans with impaired glucose tolerance. Diabetes Care 2002;25(9):1504-1510.

Mensink M, Blaak EE, Corpeleijn E, Saris WH, de Bruin TW, Feskens EJ. Lifestyle intervention according to general recommendations improves glucose tolerance. Obes Res 2003;11(12):1588-1596.

Michelini I, Falchi AG, Muggia C, Grecchi I, Montagna E, De Silvestri A, et al. Early dropout predictive factors in obesity treatment. Nutr Res Pract 2014;8(1):94-102.

Miller BM, Brennan L. Measuring and reporting attrition from obesity treatment programs: A call to action! Obes Res Clin Pract 2015;9(3):187-202.

Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. Obes Rev 2011;12(11):912-934.

Oldroyd JC, Unwin NC, White M, Mathers JC, Alberti KG. Randomised controlled trial evaluating lifestyle interventions in people with impaired glucose tolerance. Diabetes Res Clin Pract 2006;72(2):117-127.

Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, et al. Effects of Diet and Exercise in Preventing NIDDM in People With Impaired Glucose Tolerance: The Da Qing IGT and Diabetes Study. Diabetes Care 1997; 20(4): 537-544.

Penn L, White M, Oldroyd J, Walker M, Alberti KG, Mathers JC. Prevention of type 2 diabetes in adults with impaired glucose tolerance: the European Diabetes Prevention RCT in Newcastle upon Tyne, UK. BMC Public Health 2009;16;9:342.

Ramachandran A, Snehalatha C, Mary S, Mukesh B, Bhaskar AD, Vijay V, Indian Diabetes Prevention Programme (IDPP). The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1). Diabetologia 2006;49(2):289-297.

Rosenstock IM. What Research in Motivation Suggests for Public Health. Am J Public Health Nations Health 1960;50(3 Pt 1): 295–302.

Rosenstock IM. Why People Use Health Services. Milbank Q. 1966;44:94-127.

Roumen C, Corpeleijn E, Feskens EJ, Mensink M, Saris WH, Blaak EE. Impact of 3-year lifestyle intervention on postprandial glucose metabolism: the SLIM study. Diabet Med 2008;25(5):597-605.

Sakane N, Sato J, Tsushita K, Tsujii S, Kotani K, Tsuzaki K, et al. Prevention of type 2 diabetes in a primary healthcare setting: three-year results of lifestyle intervention in Japanese subjects with impaired glucose tolerance. BMC Public Health 2011;11(1):40.

Silventoinen, K, Pankow J, Lindström J, Jousilahti P, Hu G, Tuomilehto J. The validity of the Finnish Diabetes Risk Score for the prediction of the incidence of coronary heart disease and stroke, and total mortality. Eur. J. Cardiovasc. Prev. Rehabil 2005;12:451–458.

Strecher VJ, Rosenstock IM. The Health Belief Model. In Baum A, Newman S, Weinman J, West R, McManus C, editors. Cambridge Handbook of Psychology, Health and Medicine. Cambridge University Press; 1997. p. 113-117.

Teixeira PJ, Going SB, Houtkooper LB, Cussler EC, Metcalfe LL, Blew RM, et al. Pretreatment predictors of attrition and successful weight management in women. Int J Obes Relat Metab Disord 2004;28(9):1124-1133.

Stoutenberg M, Stanzilis K, Falcon A. Translation of lifestyle modification programs focused on physical activity and dietary habits delivered in community settings. Int J Behav Med 2015;22(3):312-327.

WHO, Noncommuniclable Diseases department. Adherence to long-term therapies: evidence for action. Geneva, Switzerland. WHO Press. January 2003. 198 p. Available from: http://www.who.int/chp/knowledge/publications/adherence_report/en/.

Appendix 1

Table A1. List of 21 reviewed articles.

Article title	Reference
Type 2 diabetes prevention in the real world.	Absetz et al. 2007
men.	Andersson and Rössner 1997
Impaired glucose tolerance and NIDDM: Does a lifestyle intervention program have an effect?	Bourn et al. 1994
Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin.	Diabetes Prevention Program Research Group 2002
Prevention of Type 2 (non-insulin-dependent) diabetes mellitus by diet and physical exercise. The 6-year Malmö feasibility study.	Eriksson et al. 1991
Predictors of attrition and weight loss success: Results from a randomized controlled trial.	Fabricatore et al. 2009
Socio-demographic, anthropometric, and psychosocial predictors of attrition across behavioral weight-loss trials.	Goode et al. 2016
Factors predictive of drop-out and weight loss success in weight management of obese patients.	Hadžiabdić et al. 2014
Derivation and evaluation of a risk-scoring tool to predict participant attrition in a lifestyle intervention project.	Jiang et al. 2016
Prevention of type 2 diabetes by lifestyle intervention: a Japanese trial in IGT males.	Kosaka et al. 2004
Improvement of BMI, body composition, and body fat distribution with lifestyle modification in Japanese Americans with impaired glucose tolerance.	Liao et al. 2002
Lifestyle intervention according to general recommendations improves glucose tolerance.	Mensink et al. 2003
Early dropout predictive factors in obesity treatment.	Michelini et al. 2014
Randomised controlled trial evaluating lifestyle interventions in people with impaired glucose tolerance.	Oldroyd et al. 2006
Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance.	Pan et al. 1997
Prevention of type 2 diabetes in adults with impaired glucose tolerance: the European Diabetes Prevention RCT in Newcastle upon Tyne, UK.	Penn et al. 2009
The Indian Diabetes Prevention Programme shows that lifestyle modification and metformin prevent type 2 diabetes in Asian Indian subjects with impaired glucose tolerance (IDPP-1).	Ramachandran et al. 2006
Impact of 3-year lifestyle intervention on postprandial glucose metabolism: the SLIM study.	Roumen et al. 2007 and 2011
Prevention of type 2 diabetes in a primary healthcare setting: Three-year results of lifestyle intervention in Japanese subjects with impaired glucose tolerance.	Sakane et al. 2011
Pretreatment predictors of attrition and successful weight management in women.	Teixeira et al. 2004
Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance.	Tuomilehto et al. 2001

Appendix 2

Attrition definitions in studies comparable to the PREVIEW trial

Five studies had specific definitions of attrition (Table A2). These included non-completion of a) specific number of visits, b) a considerable number of visits in a timeframe, c) specific number of consecutive visits, and d) the final assessment of the trial. The remaining eleven studies referred to attrition or indicated non-completion as simply withdrawal, not completing the trial, or lost to or unavailable to follow-up.

Table A2. Definitions of attrition or non-completion in 16 of the 21 reviewed studies.

Definitions of attrition/completion or indication for non-completion	Reference
Participants who were lost to the 3-year follow-up.	Absetz et al. 2009
Men who did not participate for a considerable number of meetings (over a six month period) were defined as drop-outs. Also included analysis of according to attendance: men were divided into two groups after one year; those who attended at least 50% of the meeting during this first year and those who attended less.	Andersson and Rössner 1997
Withdrawal, not specified.	Bourn et al. 1994
Withdrawal, not completing the 5-year trial.	Eriksson et al. 1991
Attrition was defined as failure to complete the week 52 assessment visit.	Fabricatore et al. 2009
Participants who completed the final trial assessment were considered completers, and those who failed to attend the final assessment were non-completers. Attrition was defined as non-completion of the final end-of-trial assessment.	Goode et al. 2016
Withdrawal, not specified.	Hadžiabdić et al. 2014
Program attrition for a participant was defined as not completing all 16 DPP curriculum sessions. Long-term attrition, or loss to follow-up was defined as participant becoming inactive, but not by diabetes conversion, death, or pregnancy.	Jiang et al. 2016
Not completing the 6-month interval 4-year follow-up.	Kosaka et al. 2004
In standard care group, we defined as dropout patients those who did not attend two consecutive visits, whereas in CBT group, a dropout was defined as a subject who did not attend four meetings.	Michelini et al. 2014
Lost-to-follow up, not specified.	Pan et al. 1997
Leaving study early, otherwise not specified.	Penn et al. 2009
Completers were defined as completing the 3-years.	Roumen et al. 2007
Completers were defined as completing the 3-year follow-up.	Sakane et al. 2011
Non-completion, not completing the 16-month trial.	Teixeira et al. 2004
Withdrawal, not specified.	Tuomilehto et al. 2001

Another five articles included definition of adherence (Table A3). Three studies included dietary and physical activity records/measurements in adherence assessment, and one study included the afore mentioned and achieving a weight loss goal. In the last study adherence was assessed as the percentage of appointments attended.

Table A3. Definitions of adherence in 5 of the 21 reviewed studies.

Definition of adherence	Reference
Adherence measured according to achieving weight, exercise and dietary goals.	DPP Research Group 2002
Maximal aerobic capacity (VO2max) at 6 and 24 months compared with baseline was used to estimate adherence to the exercise protocol. Three-day food records at 3, 6, 9, 12, and 24 months were used to determine adherence to the prescribed diet. Dropping out, not specified.	Liao et al. 2002
Adherence was definined with dietary and PA adherence, e.g. total fat intake or fiber goal, or attending at least 1 h/wk in supervised exercise during two years.	Mensink et al. 2003
Adherence was measured as percentage attended appointments. No definition of attrition.	Oldroyd et al. 2006
Adherence was measured according to PA and dietary records. Attrition according to attending follow-up, otherwise not specified.	Ramachandran et al. 2006

As can be seen from the above, comparing studies of similar design in regards to attrition is arduous since no single precise and uniform definition of attrition or adherence is being used. In addition to this, some studies lack a specific definition altogether, and dropout is referred to in general terms, such as non-completion or lost to follow-up.

Appendix 3

Categorical baseline characteristics of participants

Table A4. Gender of participants.

Gender	N	Percent
Female	1402	67
Male	686	33
Total	2088	100

Table A5. Ethnicity of participants.

Ethnicity	N	Percent
Caucasian	1829	88
Asian	60	3
Pacific-Polynesian	47	2
Maori/Maori-mixed	44	2
Hispanic	41	2
Black	37	2
Other	25	1
Arabic	5	0
Total	2088	100

Table A6. Site country of participants.

Country	N	Percent
Australia	169	8
Bulgaria	280	13
Denmark	353	17
Finland	276	13
New Zealand	285	14
Spain	269	13
The Netherlands	202	10
The United Kingdom	254	12
Total	2088	100

Table A7. Education of participants.

Education	N	Percent
University education	790	38
Higher vocational education	385	18
Secondary vocational education	400	19
Secondary school	302	14
Primary/Junior school	49	2
No formal education	6	0
Other	85	4
Missing	71	3
Total	2088	100

Table A8. Employment status of participants.

Employment status	N	Percent
In full-time education (not paid for by an employer)	35	2
On government training/ employment program	31	1
In paid work	1264	61
Waiting to take up paid work already accepted	14	1
Unemployed	96	5
Permanently sick or disabled	27	1
Wholly retired from work	366	18
Looking after the home	76	4
Helping a family member	17	1
Doing something else	41	2
Entrepreneur	47	2
Partly retired	14	1
Missing	60	3
Total	2088	100

Table A9. Marital status of participants.

Marital status	N	Percent
Married	1269	61
In Civil Partnership	151	7
Separated (after being married)	41	2
Divorced	221	11
Widowed	78	4
Single (never married)	246	12
Other	3	0
Cohabitation	25	1
Has a partner/engaged	9	0
Missing	45	2
Total	2088	100

Appendix 4

Characteristics associated with attrition

Table A10. Age according to different outcome statuses.

Status	N	Min	Max	Mean	SD
Completer	1170	25	70	54	10.4
Dropout	918	25	70	49	12.2
Completer	1170	25	70	54	10.4
Late WM DO	296	25	70	50	11.8
Early WM DO	239	25	70	48	12.4
LED DO	383	25	70	48	12.4

Total 2088

Table A11. Baseline BMI according to different outcome statuses.

Status	N	Min	Max	Mean	SD
Completer	1170	25	70	34	5.6
Dropout	918	25	77	37	7.3
Completer	1170	25	70	34	5.6
Late WM DO	296	25	77	37	7.2
Early WM DO	239	25	64	37	6.8
LED DO	383	25	77	36	7.6

Total 2088

Table A12. WL_{LED} according to different outcome statuses.

Status	N	Min	Max	Mean	SD
Completer	1163	7	20	12	2.3
Dropout	724	-6	20	9	3.7
Completer	1163	7	20	12	2.3
Late WM DO	287	1	20	11	2.6
Early WM DO	230	1	19	10	2.7
LED DO	207	-6	18	6	3.4

Total 1887

Table A13. DRS according to different outcome statuses.

Status	N	Min	Max	Mean	SD
Completer	1166	4	65	16	3.8
Dropout	910	4	26	16	3.3
Completer	1166	4	65	16	3.8
Late WM DO	296	9	26	16	3.2
Early WM DO	238	4	26	16	3.3
LED DO	376	4	26	15	3.2

Total 2076

Table A14. Ethnicities according to different outcome statuses.

Ethnicity	Completer		Dropout		Total	
	N	%	N	%	N	%
Caucasian	1054	58	775	42	1829	100
Asian	31	52	29	48	60	100
Black	21	57	16	43	37	100
Arabic	3	60	2	40	5	100
Hispanic	15	37	26	63	41	100
Other	10	40	15	60	25	100
Maori/Maori-mixed	18	41	26	59	44	100
Pacific-Polynesian	18	38	29	62	47	100
Total	1170	56	918	44	2088	100

Ethnicity	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
Caucasian	1054	58	240	13	217	12	318	17	1829	100
Asian	31	52	10	17	4	7	15	25	60	100
Black	21	57	3	8	3	8	10	27	37	100
Arabic	3	60	2	40	0	0	0	0	5	100
Hispanic	15	37	7	17	5	12	14	34	41	100
Other	10	40	5	20	4	16	6	24	25	100
Maori/Maori-mixed	18	41	11	25	3	7	12	27	44	100
Pacific-Polynesian	18	38	18	38	3	6	8	17	47	100
Total	1170	56	296	14	239	11	383	18	2088	100

Table A15. Countries according to different statuses.

Country	Completer		Dropout		Total	
	N	%	N	%	N	%
Australia	124	73	45	27	169	100
Bulgaria	84	30	196	70	280	100
Denmark	190	54	163	46	353	100
Finland	208	75	68	25	276	100
New Zealand	159	56	126	44	285	100
Spain	130	48	139	52	269	100
The Netherlands	134	66	68	34	202	100
The United Kingdom	141	56	113	44	254	100
Total	1170	56	918	44	2088	100

Country	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
Australia	124	73	17	10	4	2	24	14	169	100
Bulgaria	84	30	4	1	70	25	122	44	280	100
Denmark	190	54	49	14	64	18	50	14	353	100
Finland	208	75	36	13	13	5	19	7	276	100
New Zealand	159	56	79	28	14	5	33	12	285	100
Spain	130	48	53	20	43	16	43	16	269	100
The Netherlands	134	66	13	6	13	6	42	21	202	100
The United Kingdom	141	56	45	18	18	7	50	20	254	100
Total	1170	56	296	14	239	11	383	18	2088	100

Table A16. Employment statuses according to different outcome statuses.

Employment status	Completer		Dropout		Total	
	N	%	N	%	N	%
In full-time education	12	34	23	66	35	100
On government training…	16	52	15	48	31	100
In paid work	713	56	551	44	1264	100
Waiting to start work (accepted)	5	36	9	64	14	100
Unemployed	42	44	54	56	96	100
Permanently sick or disabled	15	56	12	44	27	100
Wholly retired from work	238	65	128	35	366	100
Looking after the home	35	46	41	54	76	100
Helping a family member	8	47	9	53	17	100
Doing something else	17	41	24	59	41	100
Entrepreneur	30	64	17	36	47	100
Partly retired	12	86	2	14	14	100
Total	1143	56	885	44	2028	100

Employment status	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
In full-time education	12	34	7	20	5	14	11	31	35	100
On government training…	16	52	2	6	5	16	8	26	31	100
In paid work	713	56	189	15	144	11	218	17	1264	100
Waiting to start work (accepted)	5	36	4	29	1	7	4	29	14	100
Unemployed	42	44	14	15	14	15	26	27	96	100
Permanently sick or disabled	15	56	2	7	3	11	7	26	27	100
Wholly retired from work	238	65	41	11	34	9	53	14	366	100
Looking after the home	35	46	13	17	10	13	18	24	76	100
Helping a family member	8	47	5	29	2	12	2	12	17	100
Doing something else	17	41	4	10	5	12	15	37	41	100
Entrepreneur	30	64	7	15	4	9	6	13	47	100
Partly retired	12	86	0	0	1	7	1	7	14	100
Total	1143	56	288	14	228	11	369	18	2028	100

Table A17. Education according to different outcome statuses.

Education	Completer		Dropout		Total	
	N	%	N	%	N	%
University educ.	461	58	329	42	790	100
Higher vocational educ.	250	65	135	35	385	100
Secondary vocational educ.	207	52	193	48	400	100
Secondary school	148	49	154	51	302	100
Primary/Junior school	27	55	22	45	49	100
No formal education	2	33	4	67	6	100
Other	43	51	42	49	85	100
Total	1138	56	879	44	2017	100

Education	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
University educ.	461	58	100	13	88	11	141	18	790	100
Higher vocational educ.	250	65	50	13	26	7	59	15	385	100
Secondary vocational educ.	207	52	58	15	54	14	81	20	400	100
Secondary school	148	49	64	21	33	11	57	19	302	100
Primary/Junior school	27	55	8	16	3	6	11	22	49	100
No formal education	2	33	1	17	1	17	2	33	6	100
Other	43	51	10	12	18	21	14	16	85	100
Total	1138	56	291	14	223	11	365	18	2017	100

Table A18. Marital statuses according to different outcome statuses.

Marital status	Completer		Dropout		Total	
	N	%	N	%	N	%
Married	778	61	491	39	1269	100
In Civil Partnership	68	45	83	55	151	100
Separated (after marriage)	19	46	22	54	41	100
Divorced	113	51	108	49	221	100
Widowed	36	46	42	54	78	100
Single (never married)	115	47	131	53	246	100
Other	2	67	1	33	3	100
Cohabitation	14	56	11	44	25	100
Has a partner/engaged	5	56	4	44	9	100
Total	1150	56	893	44	2043	100

Marital status	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
Married	778	61	167	13	134	11	190	15	1269	100
In Civil Partnership	68	45	27	18	22	15	34	23	151	100
Separated (after marriage)	19	46	5	12	9	22	8	20	41	100
Divorced	113	51	34	15	21	10	53	24	221	100
Widowed	36	46	17	22	3	4	22	28	78	100
Single (never married)	115	47	40	16	40	16	51	21	246	100
Other	2	67	0	0	0	0	1	33	3	100
Cohabitation	14	56	3	12	2	8	6	24	25	100
Has a partner/engaged	5	56	1	11	2	22	1	11	9	100
Total	1150	56	294	14	233	11	366	18	2043	100

Table A19. Smoking according to dichotomous outcome status.

Smoking	Completer		Dropout		Total	
	N	%	N	%	N	%
No	1050	59	732	41	1782	100
Yes, but less than weekly	35	50	35	50	70	100
Yes, at least daily	73	35	133	65	206	100
Total	1158	56	900	44	2058	100

Smoking	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
No	1050	59	255	14	184	10	293	16	1782	100
Yes, but less than weekly	35	50	13	19	9	13	13	19	70	100
Yes, at least daily	73	35	27	13	41	20	65	32	206	100
Total	1158	56	295	14	234	11	371	18	2058	100

Table A20. Alcohol use according to dichotomous outcome status.

Drinking	Completer		Dropout		Total	
	N	%	N	%	N	%
No	320	48	343	52	663	100
Yes	839	60	557	40	1396	100
Total	1159	56	900	44	2059	100

Drinking	Completer		Late WM DO		Early WM DO		LED DO		Total	
	N	%	N	%	N	%	N	%	N	%
No	320	48	100	15	90	14	153	23	663	100
Yes	839	60	195	14	144	10	218	16	1396	100
Total	1159	56	295	14	234	11	371	18	2059	100

Table A21. $WL_{2\text{-YEAR}}$ (LOCF method) according to dichotomous outcome status.

Status	N	Min	Max	Mean	SD
Completer	1170	-17	35	6	6.9
Dropout	535	-15	41	8	5.0
Total	1705				

Table A22. $WL_{2\text{-YEAR}}$ (data-based method) according to dichotomous outcome status.

Status	N	Min	Max	Mean	SD
Completer	1170	-17	35	6	6.5
Dropout	535	8	9	8	0.5
Total	1705				

Table A23. $WL_{2\text{-YEAR}}$ (BOCF method) according to dichotomous outcome status.

Status	N	Min	Max	Mean	SD
Completer	1170	-17	35	6	6.9
Dropout	535	0	0	0	0
Total	1705				

Table A24. $WL_{2\text{-YEAR}}$ (LOCF+0.2 method) according to dichotomous outcome status.

Status	N	Min	Max	Mean	SD
Completer	1170	-17	35	6	6.9
Dropout	535	-17	38	5	5.0
Total	1705				

Table A25. $WL_{2\text{-YEAR}}$ (LOCF+0.3 method) according to dichotomous outcome status.

Status	N	Min	Max	Mean	SD
Completer	1170	-17	35	6	7.0
Dropout	535	-18	36	3	5.1
Total	1705				

Appendix 5

Prediction model for dropout with logistic regression model (enter method)

Table A26. Model summary table of the dropout prediction model (enter method).

Model Summary		Cox &	Nagelkerk
Step1	-2 Log likelihood	Snell R	e R Square
	1899.406[a]	0.226	0.307

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Table A27. Variables in the prediction model for dropout with forced enter method.

Variables in the Equation		B	S.E.	Wald	df	p	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	Age	-0.032	0.007	20.413	1	0.000	0.969	0.956	0.982
	Baseline BMI	0.043	0.010	18.437	1	0.000	1.044	1.024	1.064
	WLLED	-0.285	0.024	138.166	1	0.000	0.752	0.717	0.789
	DRS	-0.014	0.019	0.581	1	0.446	0.986	0.950	1.023
Ethnicity	Base = Caucasian			6.075	7	0.531			
	Asian	-0.229	0.374	0.373	1	0.541	0.796	0.382	1.657
	Black	-0.895	0.565	2.508	1	0.113	0.408	0.135	1.237
	Arabic	0.577	1.118	0.266	1	0.606	1.780	0.199	15.937
	Hispanic	-0.107	0.419	0.065	1	0.799	0.899	0.395	2.043
	Other	0.719	0.527	1.859	1	0.173	2.053	0.730	5.772
	Maori/Maori-mixed	-0.211	0.469	0.203	1	0.652	0.810	0.323	2.029
	Pacific-Polynesian	0.329	0.442	0.553	1	0.457	1.390	0.584	3.307
Country	Base = Australia			61.862	7	0.000			
	Bulgaria	1.021	0.305	11.189	1	0.001	2.777	1.526	5.052
	Denmark	1.536	0.284	29.311	1	0.000	4.647	2.665	8.104
	Finland	0.528	0.300	3.098	1	0.078	1.696	0.942	3.053
	New Zealand	0.610	0.303	4.058	1	0.044	1.840	1.017	3.329
	Spain	1.582	0.299	27.909	1	0.000	4.865	2.705	8.749
	The Netherlands	0.659	0.346	3.623	1	0.057	1.933	0.981	3.810
	The United Kingdom	1.328	0.301	19.396	1	0.000	3.772	2.089	6.810
Employment status	Base = In full-time education...			10.940	11	0.448			
	On government	-1.002	0.662	2.291	1	0.130	0.367	0.100	1.344
	In paid work	-1.043	0.437	5.705	1	0.017	0.352	0.150	0.829
	Waiting to start work (accepted)	-0.244	0.824	0.088	1	0.767	0.783	0.156	3.937
	Unemployed	-1.064	0.517	4.242	1	0.039	0.345	0.125	0.950
	Permanently sick or disabled	-1.317	0.705	3.493	1	0.062	0.268	0.067	1.066
	Wholly retired from work	-0.792	0.475	2.783	1	0.095	0.453	0.179	1.149
	Looking after the home	-0.711	0.524	1.840	1	0.175	0.491	0.176	1.372
	Helping a family member	-0.870	0.738	1.388	1	0.239	0.419	0.099	1.781
	Doing something else	-0.829	0.600	1.910	1	0.167	0.436	0.135	1.414
	Entrepreneur	-0.639	0.559	1.309	1	0.253	0.528	0.176	1.578
	Partly retired	-1.205	0.919	1.721	1	0.190	0.300	0.049	1.814
Education	Base = University educ.			14.047	6	0.029			
	Higher vocational educ.	-0.103	0.170	0.366	1	0.545	0.902	0.646	1.259
	Secondary vocational educ.	0.368	0.165	4.993	1	0.025	1.444	1.046	1.994
	Secondary school	0.312	0.179	3.038	1	0.081	1.367	0.962	1.942
	Primary/Junior school	0.211	0.374	0.317	1	0.574	1.234	0.593	2.570
	No formal education	1.348	1.039	1.683	1	0.194	3.850	0.502	29.495
	Other	0.620	0.289	4.594	1	0.032	1.859	1.054	3.277
Marital status	Base = Married			9.596	8	0.295			
	In Civil Partnership	0.430	0.221	3.777	1	0.052	1.538	0.996	2.374
	Separated (after marriage)	0.499	0.465	1.153	1	0.283	1.648	0.662	4.101
	Divorced	0.260	0.183	2.008	1	0.156	1.297	0.905	1.858
	Widowed	0.390	0.296	1.737	1	0.188	1.476	0.827	2.636
	Single (never married)	-0.018	0.183	0.009	1	0.924	0.983	0.686	1.407
	Other	-19.774	25748.709	0.000	1	0.999	0.000	0.000	
	Cohabitation	-0.867	0.633	1.877	1	0.171	0.420	0.122	1.453
	Has a partner/engaged	0.101	0.747	0.018	1	0.892	1.107	0.256	4.789
Smoking	Base = No			4.275	2	0.118			
	Yes, but less than weekly	0.009	0.297	0.001	1	0.976	1.009	0.564	1.806
	Yes, at least daily	0.420	0.204	4.256	1	0.039	1.522	1.021	2.296
Alcohol use	Drink alcohol (yes)	-0.016	0.131	0.016	1	0.901	0.984	0.762	1.271
	Constant	2.603	0.789	10.877	1	0.001	13.498		

a. Variable(s) entered on step 1: Age, baseline BMI, WLLCD, DRS, Ethnicity, Country, Employment status, Education, Marital status, Smoking, Alcohol use.