Eur J Health Econ (2016) 17:379–389
DOI 10.1007/s10198-015-0682-0

ORIGINAL PAPER

Does diabetes prevention pay for itself? Evaluation of the M.O.B.I.L.I.S. program for obese persons

Jan Häußler[1] · **Friedrich Breyer**[1]

Received: 28 May 2014 / Accepted: 4 March 2015 / Published online: 31 March 2015

Abstract In response to the growing burden of obesity, public primary prevention programs against obesity have been widely recommended. Several studies have estimated the cost-effectiveness of diabetes-prevention trials for different countries. Nevertheless, it is still controversial if prevention conducted in more real-world settings and among people with increased risk but not yet exhibiting increased glucose tolerance can be a cost-saving strategy to cope with the obesity epidemic. We examine this question in a simulation model based on the results of the M.O.B.I.L.I.S program, a German lifestyle intervention to reduce obesity, which is directed on the high-risk group of people who are already obese. The contribution of this paper is the use of 4-year follow-up data on the intervention group and a comparison with a control group formed by SOEP respondents as inputs in a Markov model of the long-term cost savings through this intervention due to the prevention of type 2 diabetes. We show that from the point of view of a health insurer, these programs can pay for themselves.

Keywords Diabetes prevention · Cost analysis · Markov modeling · Obesity

JEL Classification I12 · H51

✉ Jan Häußler
haeussler.jan@gmail.com

[1] Department of Economics, University of Konstanz, Fach 135, 78457 Constance, Germany

Introduction

One of the main challenges for population health in the developed world is the steady increase in obesity. In the US, in 2009/2010, 35.7 % of the adult population was obese [29], and in several European countries such as Spain and Germany, the corresponding percentages were between 14.7 and 23 % [5, 27]. The reasons for this trend are a combination of increased intake of calories and reduced physical activity [9] and the resulting energy imbalance not only leads to a continuous weight gain but also to severe chronic conditions such as type 2 diabetes. Globally, it is estimated that 438 million people (7.8 % of the adult population) will have developed type 2 diabetes by 2030 unless effective prevention programs are implemented [18]. Diabetes and other diseases emerging as consequences of excessive weight will also cause a sizable economic burden [15] through direct medical costs of treatment as well as indirect costs of illness such as disability and early retirement [21, 10].

In response to this growing burden of obesity, public primary prevention programs against obesity have been widely recommended [34]. The European guideline for the prevention of type 2 diabetes focuses on obesity and sedentary lifestyle as these are the main modifiable risk factors of the disease [30]. Several studies [26, 7, 17, 3] have estimated the cost-effectiveness of diabetes-prevention trials for different countries. Thus, the aim of these studies was to measure either the costs per T2D case prevented or per QALY gained (see reviews by [30, 20]). Furthermore, [25] found that among all different interventions recommended by the American Diabetes Association (ADA), evidence was strongest for the cost-effectiveness of intensive lifestyle modification among persons with impaired glucose tolerance (IGT).

Table 1 Number of participants by year of intervention and follow-up

Start intervention	Completion intervention	2011 Follow-up	2012 Follow-up
2005	2006	27	11
2006	2007	51	100
2007	2008	34	93
2008	2009	0	85

Overall, 401 participants with a resulting average follow-up period of 4.08 years and a total observational period of 5.08 years

Nevertheless, it is still controversial if prevention programs conducted in more real-world settings and among people with increased risk but not yet exhibiting IGT really pay for themselves from the perspective of the health care system. We examine this question in a simulation model based on the results of the M.O.B.I.L.I.S program, a German lifestyle intervention to reduce obesity, which is directed on the high-risk group of people who are already obese. The contribution of this paper is the use of a Markov model to measure long-term cost savings of this intervention due to prevention of type 2 diabetes.

This paper is organized as follows. "Data and methods" is devoted to a description of data and methods and first presents an overview of the M.O.B.I.L.I.S. intervention program and the analyzed data, followed by a characterization of the Markov model framework and the analysis of cost savings. "Results" presents the results of the medium-term simulation and some sensitivity analyses. In the "Discussion", we discuss our findings. Concluding remarks are offered in "Conclusions".

Data and methods

Overview of the M.O.B.I.L.I.S. intervention

The M.O.B.I.L.I.S. program is a lifestyle intervention to reduce obesity. The program has been implemented in more than 100 sites throughout Germany since 2004, and with by now over 6000 participants [24]. The program addresses obese adults (BMI 30–40 kg/m^2) who have at least one additional obesity-related risk factor such as type 2 diabetes, high blood pressure, or orthopedic problems, but who are still capable of light exercise. Individuals with type-1 diabetes or several other conditions (such as eating disorders, consumption of psychotropic drugs, or cancer treatment in the last 5 years) are not allowed to participate. The intervention takes place in a group setting (with 15–18 participants per group) and consists of a total of 60 group sessions over a time period of 12 months, with a higher frequency in the first 8 weeks. Forty of the sessions are devoted to light physical activity such as walking, aerobic, yoga, and, if possible, jogging, and the remaining 20 sessions to nutritional and behavioral counseling. Participants are also encouraged to utilize sports facilities offered in their communities.

What makes the M.O.B.I.L.I.S. intervention particularly interesting for a health economic evaluation is the fact that participation fees are to a large extent reimbursed by public health insurance. The participants have to pay participation fees of €785 (before 2008: €685) in advance but upon completion of the program all but a co-payment of €100 is returned by the sickness fund provided the person has participated in at least 75 % of all sessions. Thus, besides their intrinsic motivation to improve their health status, the participants have an additional monetary incentive to attend the program sessions regularly once they have started the program. In fact, of the first 5025 starters, 4336 (or 87 %) participated in the check-up scheduled at the end of the program and 3985 (or 92 %) of those or 79 % of all starters had fulfilled the condition of a 75 % participation rate.

The effectiveness of the intervention in reducing weight and enhancing the physical activity of the participants over the intervention period has been documented in the studies by [2, 24]. Frey et al. [11] also show that the intervention effects are persistent 1 year after program termination. Compared to the initial levels, average weight was reduced by 6.8 kg (±7.8) and weekly physical activity was significantly higher. Besides that, [14] find that compared to a quasi-experimental control group, the participants show significantly enhanced psychological variables (self-efficacy, strength of goal intention) at a 2-year follow-up.

4-Year follow-up

The present follow-up study conducted in 2011 and 2012 allows the evaluation of long-term effects of the M.O.B.I.L.I.S. program. With special focus on the prevention of type 2 diabetes as a crucial public health consequence of obesity, the follow-up study and the present evaluation only consider the subgroup of participants with no type 2 diabetes at the beginning of the intervention (86.5 %) [24]. The follow-up periods after completion of the intervention vary between 3 and 6 years with an average follow-up period of about 4 years after the

Table 2 Descriptive statistics of the M.O.B.I.L.I.S. sample development

Variable	Baseline	Completion	Follow-up
Weight (kg)	98.603 (12.246)	91.305 (13.046)	95.683 (14.299)
BMI (kg/m^2)	34.887 (2.692)	32.317 (3.509)	33.850 (3.862)
WHR	0.926 (0.078)	0.909 (0.078)	0.917 (0.082)
Heart rate (bpm)	78.005 (12.269)	73.990 (11.036)	70.972 (9.478)
Blood pressure (mmHg)[a]	132.5/85.8 (15.9/9.1)	126.1/82.3 (15.3/10.6)	132.1/82.5 (15.0/9.4)
HDL (mg/dl)	57.197 (13.666)	59.784 (14.322)	61.714 (15.742)
LDL (mg/dl)	133.418 (33.015)	131.811 (32.639)	133.952 (33.798)
Fasting blood sugar (mg/dl)	92.970 (13.152)	91.636 (12.737)	93.870 (16.490)
HbA1c (in %)	5.625 (0.341)	5.637 (0.281)	5.649 (0.435)
Physical activity[b]	22.879 (21.416)	37.031 (27.352)	34.305 (27.269)

$N = 401$, standard errors in *parenthesis*

[a] Systolic/diastolic

[b] Measured as MET per week, based on the evaluation of the physical activity questionnaire developed by [12]

completion of the program. To rule out seasonal effects, the participants were contacted in the same calendar month in which they had completed the program. A total of $N_0 = 958$ individuals that had completed the program had taken part in the 1-year follow-up and fulfilled the inclusion criteria (before the intervention: age 40–60 years, BMI 30–40 kg/m^2, no type 2 diabetes) were asked by mail and additional personal phone calls to participate in the study. A response rate of 42 % results in a study sample of $N = 401$ individuals, with an average age 49.8 years (±5.7), average BMI 34.9 kg/m^2 (±2.7), and a female share of 82.5 %. M.O.B.I.L.I.S. participants from the years 2005 (10 %), 2006 (38 %), 2007 (31 %), and 2008 (21 %). An overview of the respective follow-up lengths can be found in Table 1. All medical and anthropometric values of the sample are reliable, as they were measured and documented by physicians. Additional lifestyle-related items (physical activity, nutrition) were raised on the basis of a questionnaire.

The descriptive statistics in Table 2 provide an overview of the sample characteristics before and after the intervention, as well as for the 4-year follow-up. The obesity measures [BMI, waist-to-hip ratio (WHR)] and reported physical activity have a common pattern over time. Compared to the baseline, the intervention still has a positive effect at the follow-up, though there is a clear rebound effect when we regard the development after completion of the program. However, looking at the other medical risk indicators, the interpretation is not as clear. For blood pressure, LDL cholesterol, and fasting blood sugar, the initial improvements after the intervention vanish, as the follow-up levels meet the baseline values. HDL cholesterol and HbA1c show a rising time trend, which might be due to general age effects [6] overlapping with possible effects of the intervention.

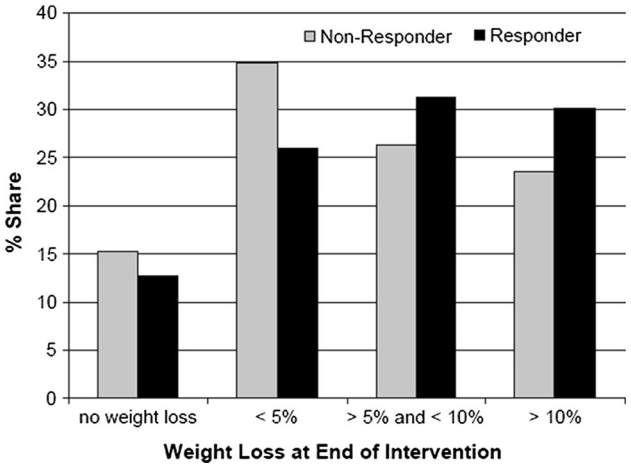

Fig. 1 Influence of initial program success

In order to check for possible self-selection effects in the follow-up sample, we compare the responders to the non-responders with respect to their initial weight loss during the program. The differences are not very striking, though we cannot exclude a self-selection bias for participation in the study. Among responders, 61.4 % had experienced a weight loss of more than 5 % while participating in the program, whereas the respective number among the non-responders was only 49.9 % (see Fig. 1 for details). The response rate of those with an initial weight loss of less than 5 % was 35.7 %, compared to 46.9 % for those who lost more than 5 % during the program.

Control group

For an informative evaluation of the economic benefits of the intervention based on the development of the risk factors, we need to account for the general age-related

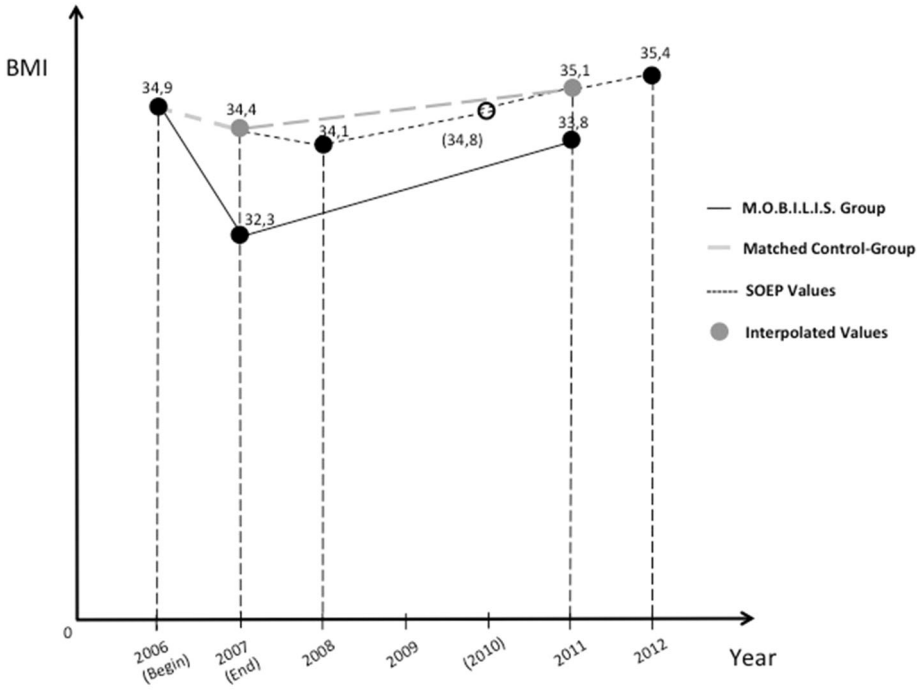

Fig. 2 BMI development in M.O.B.I.L.I.S. and the SOEP group. Average BMI values reported for the intervention group (adjusted for 2006, 2007, 2011). Computed values for the matched control-group (2006, 2007, 2011) and the reported SOEP values (2006, 2008, 2010, 2012)

trend in the population, which is achieved by including a control group. Due to the real-world implementation of the M.O.B.I.L.I.S. program and the retrospective study design, we are lacking an original randomized control group. We overcome this problem with the aid of an artificial control group, formed from a subgroup of $N_0 = 1308$ individuals from the German socio-economic panel (SOEP) in 2006 who meet the same inclusion criteria at the beginning of the intervention. We choose 2006 as baseline for the matching and further modeling as the largest share of M.O.B.I.L.I.S. participants started the intervention in this year (see Table 1) and besides the SOEP data for weight are also available. The SOEP sample [33] is representative of the German population so that comparability for socio-economic and other background variables should be reasonably high. We performed propensity score matching estimated on the baseline covariates BMI, age, and gender for the two nearest neighbors in the SOEP sample to build our artificial control group.

Unfortunately, the SOEP data only contain information on weight and BMI as risk factors for diabetes, so we lose the information on the additional risk indicators included in the M.O.B.I.L.I.S. dataset in the control group. As data on weight is only available in the SOEP sample every 2 years, we compute the annual BMI development in the missing years between 2007 and 2011 through linear interpolation of the known values in 2006, 2008, and 2012. Thus, we are able to adjust the control group to the 1-year period of the intervention and the average 4-year follow-up length. Figure 2 depicts the interpolated BMI values for the intervention group and the control group.

The average development in the matched control group in comparison to the measured intervention group values can be found in the first column of Table 3. Being matched on the baseline BMI, the control group shows no significant change in BMI in the period until completion of the program. Over the 4 years of the follow-up period, the average BMI in the control group grows at a low rate from 34.458 to 35.088 kg/m^2, and compared to the baseline value of 34.887 kg/m^2, the overall rise in weight in the control group is not significantly different from zero. Thus, the average BMI in the control group is more or less constant over the observed period, while the intervention group shows the pattern of initial weight loss and a rebound in the following 4 years as described before.

The observed reduction in BMI from the start to the end of the intervention is 7.4 %, while the estimated average treatment effect (ATE) of BMI development is a 5.9 % reduction for the intervention group, compared to the control group. In the period between completion and 4-year follow-up, the observed BMI in the intervention group grows by 4.6 %, whereas the ATE in this period only shows a 2.8 % rise in BMI for the intervention group. As the ATE is significantly below the observed rebound effect in the intervention group, we can state that the effects of regaining weight in the intervention group at the follow-up are on average lower when we take the representative control into account. The differences in distribution over

Table 3 BMI means and subgroup development

	BMI[a]	BMI < 30	30 ≤ BMI < 35	35 ≤ BMI
Baseline intervention[b]	34.887 (2.692)	0.0	0.5287	0.4713
Completion intervention	32.317 (3.509)	0.2718	0.5037	0.2244
Completion control	34.458 (2.792)	0.0698	0.4589	0.4713
Follow-up intervention	33.850 (3.862)	0.1621	0.4514	0.3865
Follow-up control	35.088 (3.277)	0.0474	0.4638	0.4888

$N = 401$

[a] Measured in kg/m^2

[b] Control matched on BMI at the baseline before the intervention

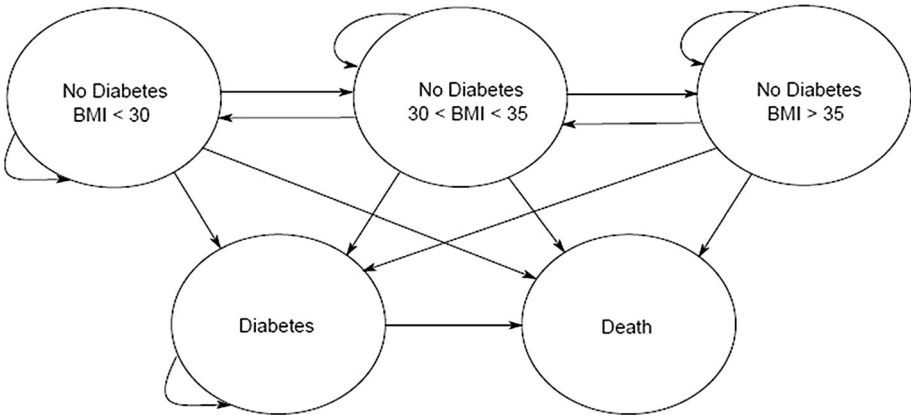

Fig. 3 States of the Markov model

the three obesity subgroups overweight, obese, and severely obese (BMI < 30, 30 ≤ BMI ≤ 35, BMI > 35) in the M.O.B.I.L.I.S. population and the control group are presented in Table 3. All statistical analyses were performed using Stata 12.

Simulation framework

Evaluating the long-term effects of the M.O.B.I.L.I.S. intervention requires information on the development of obesity-related diseases beyond the data of the 4-year follow-up. Abstracting from other diseases such as myocardial infarction and stroke, our analyses focus on the development of diabetes based on the individual obesity level. Using a Markov cohort simulation we estimate the long-term effects of the M.O.B.I.L.I.S. intervention on the prevalence of type 2 diabetes in the control and the intervention group. The Markov model is a variation of the model used in a previous study to evaluate the Finnish GOAL Intervention [16].

The state-transition Markov model consists of five mutually exclusive (disease) states and discrete 1-year intervals. The model structure depicted in Fig. 3 allows us to follow the starting population over a 20-year time horizon, using the annual forecasts for every state of interest. We limit the time horizon of the Markov model to 20 years because we think that any weight differences observed between the groups at a later time can no longer be traced back to the intervention. This is in line with findings of previous follow-up studies [8]. Members of the intervention and control groups move between the Markov states according to given transition probabilities (see Table 4 for an overview). Both groups enter the model at stage zero according to the observed (estimated) distributions over the model states at the 1-year follow-up. The five Markov states are: (1) no diagnosed diabetes and BMI < 30, (2) no diagnosed diabetes and 30 < BMI < 35, (3) no diagnosed diabetes and BMI > 35 (afterwards we will refer to all three of these states as "non-diabetes"), (4) diagnosed diabetes, and (5) the absorbing Markov state death. The initial average age of the individuals entering the Markov model is 54 years, which is consistent with the average age at the follow-up in both groups. With respect to the gender shares, we simulate the model according to the mixed composition in the follow-up sample and in a separate subgroup analysis for females only. We do not run a separate subgroup for males, as the low number of male participants makes a sound analysis impossible.

The annual transition probabilities between the three non-diabetes states are computed from the analysis of the BMI development between the end of the M.O.B.I.L.I.S.

Table 4 Annual transition probabilities in the Markov model

Transition to from	BMI < 30	30 ≤ BMI < 35	35 ≤ BMI	Diabetes[b]	Death[b]
BMI < 30	0.80298[a]	0.18302[a]	0	0.00834	0.00566
	0.63606	0.34994			
30 ≤ BMI < 35	0.02467[a]	0.83391[a]	0.11256[a]	0.02320	0.00566
	0.01526	0.89981	0.05607		
35 ≤ BMI	0	0.04142[a]	0.90911[a]	0.04381	0.00566
		0.03502	0.91551		
Diabetes[b]	0	0	0	0.99425	0.00575

Upper line intervention group, *lower line* control group

[a] Depending on the assumed further weight development in the intervention group, the values adjust to the control group over time

[b] Values for the age group 55–60. Age-dependent in 5-year steps

intervention and the follow-up,[1] and the respective changes in the control group. We do not further consider the BMI progress within the year of the program for the computation of the transition probabilities, as this would implicitly assume regular repetition of the intervention. However, the singular 1-year intervention effect is captured by the composition of the intervention group and control group with respect to the Markov states at the start of the simulation. To account for a further equalization in weight development of the intervention group to the general time trend reflected by the control group, we assume a linear adjustment of the non-diabetes transition probabilities starting after the 4-year follow-up to the control group values over 10 years. This assumption is in line with previous findings on long-term effects of weight reduction [8, 28].[2] In line with the data for the development in the intervention group and the control group, we assume that annual weight changes are not big enough to jump from the lowest to the highest non-diabetes state (or vice versa) in one step.

All other annual transition probabilities are based on the results of other studies and German epidemiological data: non-diabetes to diabetes [31, 4], non-diabetes to death and diabetes to death (Statistisches Bundesamt 2010; GBE 2006; [1]), are all one way. The transition probabilities to the state of death are adjusted to the gender composition of the samples and they are age-dependent in 5-year steps. Due to missing data, the transition from non-diabetes to death does not differ by BMI category. All of the people who die remain in this state forever and we only regard diagnosed cases of type 2 diabetes mellitus where no cure is feasible. Those probabilities are assumed to be equal for the two groups of individuals, but they vary by age and are adjusted to the gender composition of the groups. All simulations were performed using TreeAge Pro Healthcare (Release 1.0 b1 2001; TreeAge Software Inc.).

Cost analysis

The results of the Markov simulation are the foundation of the subsequent analysis of expected cost savings. The cost analysis (CA) adopts the perspective of the health insurance system and abstracts from effects on human capital, work loss, etc. Also, subjective utility from the health state does not enter the CA. We measure the direct costs of diabetes by monetizing the simulation outcomes for diabetes prevalence in each year and in both groups.

The diabetes costs in year i are defined by the following equation:

$$\text{Costs\_}i\text{\_group} = \left[\text{Costs\_w/ocomplications} + \text{Rate\_}i\text{\_complications} \times \text{Costs\_withcomp}\right] \times \text{Prevalence\_}i\text{\_group}.$$

Discounted overall costs of diabetes for each group are respectively:

$$\text{COSTS}_{\text{group}} = \sum_{i=0}^{20} \frac{1}{(1+r)^i} \text{Costs}_{i\text{\_group}}$$

Based on the results by [15], we assume no complications for the first 5 years after the diagnosis of diabetes. Starting from year 6 on, we assume a linear increase of the average complication rate by 5 % annually until year 10 after the diagnosis of diabetes. Subsequently, we assume that the average complication rate remains constant at 25 %. Regarding the treatment costs of diabetes in Germany, we use direct costs of €850 for the disease without complications (medications, physicians' outpatient services) and €8830

[1] We account for the actual length of the follow-up (between 3 and 6 years) and afterwards linearly adjust the annual transition probabilities to the average follow-up period of 4 years in both groups.

[2] Since after the end of the program members of the intervention group on average gain weight faster than members of the control group due to the rebound effect, this assumption also prevents the implausible outcome that after a number of years the former are on average heavier than the latter.

with complications[3] accounting for further direct costs from impatient care, medical devices, transportation, and long-term care [22, 23]. The intervention costs—for a sickness fund—are €685 (€585 before 2009), resulting in average intervention costs according to the sample composition of €673. The intervention costs per individual are based on the accounting of the provider and those are also the costs the sickness funds have to pay (without the out-of-pocket payment of €100). For our evaluation, we only take into account the costs without the deductible of €100, as we adopt the perspective of the sickness fund. All costs used for the analysis are expressed in 2011 euros and all future costs were compounded at an annual discount rate of 3 % [19].

Sensitivity analyses

To check for the validity of our results, we conduct one-way sensitivity analyses regarding the values of all parameters of the model. Besides, we conduct a probabilistic sensitivity analysis in which we vary all costs and parameters entering the model. The respective parameter values are drawn simultaneously in a Monte Carlo simulation from beta (all probabilities), log-normal (diabetes costs) and gamma distributions (start age, time until complications). Distribution parameters rely on mean values and confidence intervals for known values. Where no confidence interval is available, we assume that a 50 % change in the assumed values corresponds to a distance of two standard deviations from the mean value.

Results

Mortality

The simulated cumulative annual mortalities depicted in Fig. 4 for the mixed-gender groups indicate that the intervention has no significant effect on the longevity of the individuals. The overall time trend of the mortality rates is the same for the intervention group and the control group, with the annual rates differing only at insignificantly low margins. The cumulative mortality after 20 years is 27.00 % in the control group compared to 27.28 % in the intervention group. Similarly, the difference in average annual mortality rates between 1.35 % in the control group and 1.36 % in the intervention group is not significantly different from zero. The simulation results for the female subgroup show that the described mortality effects are independent of the gender composition of the groups. In this case, the cumulative mortality rate after 20 years is 20.26 % (20.05 %) in the intervention (control) group. Summarizing the results, we can state that, at least in our model setup, the M.O.B.I.L.I.S. intervention has no effects on long-term mortality rates.

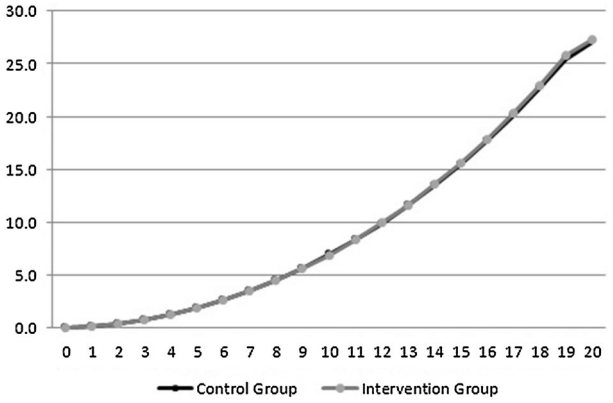

Fig. 4 Cumulative mortality rates

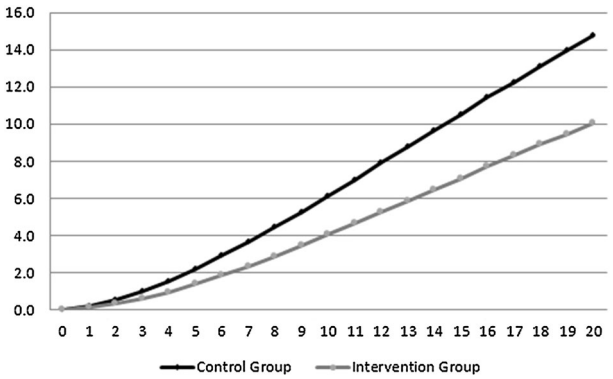

Fig. 5 Annual diabetes prevalence rates

Diabetes prevalence

The annual prevalence rates of diabetes over the 20 years of the simulated time horizon are depicted for the intervention group and the control group in Fig. 5. The prevalence level in the intervention group is markedly below that in the control group over all the 20 years, with an average annual rate of 4.59 % in the intervention group compared to 6.87 % in the control group. The simulated diabetes prevalence at the end of the simulation is 31.9 % lower in the intervention group than in the control group, with an absolute difference of 4.73 percentage points. The estimated diabetes prevalence in the M.O.B.I.L.I.S. population would be considerably lower if the success of the intervention persisted over the whole period. Looking at the literature on the long-term effects of other interventions [28, 32], this scenario seems unrealistic, but we included the simulation as a benchmark in the sensitivity analysis.

[3] Including macrovascular disease, nephropathy, neuropathy, and retinopathy.

Table 5 Overview of net cost savings

Avg. complications	Mixed group (€)	All female (€)
NO	−227	−269
After 5 years	411	349
After 10 years	**327**	**299**
After 15 years	269	228

The assumed scenario is in *bold*. 25 % Complications on average

The prevalence of diabetes for the female subgroup shows the same pattern as described above for the mixed group, though the differences in prevalence rates are a bit less pronounced. On average, the annual prevalence rate is 3.67 % (5.10 %) in the intervention (control) group. At the end of the simulated time horizon, the diabetes prevalence is 8.04 % in the intervention group compared to 10.97 % in the control group. This prevalence being 26.7 % lower in the intervention group compared to 31.9 % in total indicates that the male share of the mixed sample has a positive influence on the effectiveness of the intervention as a whole.

Cost analysis

As discussed in the Methods section, the extent of the net cost savings depends on the average time between the onset of type 2 diabetes and the onset of complications (see Table 5 for an overview). In the extreme case of no complications at all, the M.O.B.I.L.I.S. intervention does not pay off from the perspective of a sickness fund. In that situation, with annual type 2 diabetes costs of €850, the difference in the discounted (avoided) diabetes costs between control and intervention group amounts to €446. With average intervention costs of €673 for a sickness fund, this leads to net costs per participant of €227 for the mixed group and €269 for the female subgroup.[4]

Following the rationale of the computation described in the Methods section—with an average time until the onset of complications of 10 years—the M.O.B.I.L.I.S. intervention does save costs. Based on the assumption that the costs of type 2 diabetes rise linearly from €850 (cost without complications) in the first 5 years after diagnosis to €3057 (25 % complications on average) in year 10 and afterwards, the differences in the costs of diabetes are as follows: The discounted diabetes costs cumulated over the whole period of 20 years are €1000 lower for individuals from the intervention group than for those from the control group. For the mixed-gender group, this results in net cost savings of €327 from the perspective of a sickness fund. The positive result also holds for the female subgroup, with net cost savings of €299.

[4] As Table 4 contains cost savings, these differences of 227 and 269 € appear as negative numbers.

Sensitivity analyses

The analysis in the previous section shows that the results of the CA are affected by the assumed time span before diabetes with complications sets in. However, only in the extreme case without any complications are the results affected qualitatively. Altering the timespan from 10 years to 5 or 15 years shows that the results are robust for moderate changes in this model assumption and it provides a confidence interval for the magnitude of the cost savings. In the mixed group, if 5 years pass before the onset of complications, the net cost savings is €411, while the figure is reduced to €269 if 15 years pass before complications arise.

Besides the assumed model structure, the results of the CA may also be sensitive to the choice of model parameters. To check for the robustness of the results, we therefore performed one-way sensitivity analyses for each external parameter used in the model, the results of which are reported in Table 6. For the various transition probabilities in the model, we conducted the sensitivity analyses in an interval from zero to twice the used value, while we chose a 10-year interval for all variables measuring a time span. Table 6 shows that the qualitative result of positive net cost savings is robust to changes in start age, length of time horizon, and time until the onset of complications as well as to variations of the mortality rates. The same is true for realistic values of the discount rate. Not surprisingly, if the diabetes probability of people with BMI between 30 and 35 (or over 35) were zero, a program like this would not be worth its costs. However, the break-even values of the parameters reported in the last column show that these probabilities could fall quite considerably before the program loses its cost-saving nature.

The probabilistic sensitivity analysis leads to a 95 % confidence interval for the net cost savings ranging from €161.60 to €410.06. The interval is based on the 2.5 and the 97.5 percentile values of the net cost savings for the mixed-gender group. Thus, we can summarize that the overall result of positive net cost savings is quite robust to the choice of model parameters.

Discussion

We are well aware of the limitations and shortcomings of the current study that have to be considered. The first point to look at is the lack of a randomized control group for the evaluation. Although we are missing socioeconomic background variables for the intervention group to compare them directly with our matched control group, differences with respect to socioeconomic status should be considered a minor problem, since the intervention group was recruited

Table 6 Overview sensitivity analyses

Variable	Net savings (low value)	Net savings (used value)	Net savings (high value)	Break-even value
Discount rate	465 (0)	327 (0.03)	106 (0.06)	0.079
Diabetes mortality	469 (0)	327 (0.00575)	68 (0.012)	0.019
Diabetes probability BMI > 35	−122 (0)	327 (0.04381)	1040 (0.09)	0.022
Mortality rate BMI > 35	381 (0)	327 (0.00566)	193 (0.012)	0.035
Diabetes probability 30 < BMI < 35	−59 (0)	327 (0.02320)	1456 (0.05)	0.0059
Diabetes probability BMI < 30	479 (0)	327 (0.00834)	232 (0.017)	0.073
Start age	401 (49)	327 (54)	284 (59)	
Number of stages (time horizon)	62 (15)	327 (20)	392 (25)	10.6
Time until average complications	411 (5)	327 (10)	269 (15)	

The lower and the upper bound of the one-way sensitivity analyses, as well as the used values are in *parentheses*. The last column reports the critical parameter value for the break-even point of the net savings. All net savings are expressed in 2011 euros

from all over Germany and the control group is a matched representative subsample of the German population.

There are two main problems with the artificial control group. First, the BMI and weight values in the SOEP are based on self-reported data. We are aware of possible underreporting bias compared to the values in the intervention group that were collected by physicians. However, this underreporting is likely to occur at all measuring points, the difference of the values between two points of time (end of intervention and 4-year follow-up) should not be biased, and it is exactly this development over time that we compare between the two groups. The second problem is, of course, that in the control group we cannot account for a possible selection bias with respect to participation in the intervention group. This might exaggerate the effects for the intervention group as participants have the intention to lose weight.

The only way to avoid this selection bias would be to run a randomized controlled trial with equally motivated persons of whom only one part is admitted to the intervention whereas the others are used as controls. In the medical literature, RCTs are controversial on ethical grounds (see e.g., [13]) because they require consciously withholding a presumably beneficial treatment from patients who could benefit from it. When the long-term effects are to be studied, these persons would even have to be banned from the intervention for a number of years (in our case, 4 years). This would not only be ethically problematic but virtually infeasible since similar programs as M.O.B.I.L.I.S. are offered in many places throughout Germany and thus members of the control group could not be prevented from participating in any of these, which most of them probably would do. Thus, it would be unclear with what the M.O.B.I.L.I.S. program would be compared in that case.

As a final defense of our procedure, we would argue that the motivation of participants for losing weight does not alter the implications of the analysis for the evaluation of public prevention programs. As long as prevention programs offered by sickness funds are voluntary, we will always observe similar self-selection of participants.

Besides this selection effect at the baseline, we have to consider the differences between responders and non-responders at the 4-year follow-up, as mentioned in the data section and depicted in Fig. 1. The follow-up responders had on average a larger weight loss by the end of the program, which might in general exaggerate the long-term effectiveness of the whole intervention in the present analysis. However, looking at BMI development between the completion of the program and the follow-up in detail, it appears that this self-selection of responders does not obviously bias our findings in one direction. The effect of an initial BMI decline on the BMI development in the 4-year follow-up period is significantly negative, with an estimated coefficient of −0.429 (SE: 0.054) in a simple OLS regression model used as a robustness check. Thus, the observed rebound effect is stronger for those who lost more weight during the program. This larger regain in BMI has in turn negative consequences for the long-term effectiveness of the intervention. Considering these two points, it is not clear that the larger share of initially successful responders in the follow-up overstates the effects of the intervention.

Regarding the structure of the Markov model, two points can be criticized, the pure foundation on BMI changes and the focus on diabetes as the only cost relevant outcome. We think the former issue is justified by the fact that our control group dataset does not allow for a richer model, as the BMI is the only relevant risk factor we can observe. The focus on the cost of diabetes might understate the effectiveness of the intervention in the CA as we omit possible cost savings due to other obesity-associated diseases. Consequently, we regard our positive cost-saving results as a lower benchmark for the total cost effects of the

M.O.B.I.L.I.S. intervention and think that considering further diseases should add further cost savings. The effect on the simulated mortality should on the other hand be minimal, as the overall mortality rates for obese persons used in the model as transition probabilities reflect the other diseases as well.

Conducting the cost analysis from the perspective of the health care system might neglect some additional benefits for the society at large. Nevertheless, we think the health care system perspective is the right one for this analysis as we evaluate a program advertised and reimbursed by sickness funds. Besides looking at the simulation results, we think there is another point for our perspective. As projected mortality rates in the intervention group and control group do not differ significantly, taking those effects into account would not change the CA results we found from the perspective of the health care system. A further issue is that our analysis neglects sick leave costs of diabetes. This is a valid point, although in our model framework these costs—for a population with a starting age of 54 and a time horizon of 20 years—could only occur in the first years. Considering the development of diabetes prevalence rates in the first 10 years of our model (Fig. 5), we cannot exclude those additional indirect costs. However, with constantly higher rates in the control group, the results will not change qualitatively.

The difference in the magnitude of effects between the mixed group and the female subgroup basically reflects the lower diabetes prevalence over the whole simulation horizon of women. On the other hand, this effect is decreased by the higher female life expectancy. In combination with effects due to the composition of the mixed group, we cannot make any predictions on the effectiveness of this intervention for a subgroup of male participants.

Conclusions

In this paper we have shown that from the perspective of the health care system, real-world diabetes prevention programs for obese people such as the M.O.B.I.L.I.S. intervention may pay for themselves in the long run. We consider the net cost savings found in our study as a lower bound for the total cost effects of the intervention, as we only evaluate the costs of diabetes. Taking avoided costs of further obesity-related diseases into account would presumably raise the net effects of the intervention. On the other hand, the estimated cost savings themselves may be slightly biased upward due to the (mild) self-selection of participants. It seems, however, reasonable to assume that this upward bias is smaller than the downward bias mentioned before. From a policy point of view, the results indicate that allowing sickness funds to invest in prevention might help to reduce health care expenditures, if the funds are able to detect effective programs and address the appropriate target groups.

Acknowledgments This research was funded by the German Ministry of Education and Research (BMBF) under the grant no. 0315671 and by BARMER GEK. The authors are grateful to M.O.B.I.L.I.S. e.V. for collecting and providing the data. Valuable comments by Rolf Holle, Normann Lorenz, and Jürgen Maurer as well as participants of the iHEA World Congress 2013 in Sydney, the Augsburg Workshop on the Economics of Health Behaviors and Prevention 2013 and of the Annual Meeting of the German Health Economics Association (DGGÖ) 2014 in Munich and two anonymous referees are gratefully acknowledged.

References

1. Bender, R., Zeeb, H., Schwartz, M., Joeckel, K.-H., Berger, M.: Causes of death in obesity: relevant increase in cardiovascular but not in all-cancer mortality. J Clin Epidemiol **59**, 1064–1071 (2006)
2. Berg, A., Berg Jr, A., Frey, I., König, D., Predel, H.G.: Ergebnisse zu M.O.B.I.L.I.S.—einem Bewegungsorientierten Schulungsprogramm für adipöse Erwachsene. Deutsches Ärzteblatt **105**(11), 197–203 (2008)
3. Bertram, M.Y., Lim, S.S., Barendregt, J.J., Vos, T.: Assessing the cost-effectiveness of drug and lifestyle intervention following opportunistic screening for pre-diabetes in primary care. Diabetologia **53**(5), 875–881 (2010)
4. Bonora, E., Kiechl, S., Willeit, J., Oberhollenzer, F., Egger, G., Meigs, J.B., Bonadonna, R.C., Muggeo, M.: Population-based incidence rates and risk factors for type 2 diabetes in white individuals—the Bruneck study. Diabetes **53**, 1782–1789 (2004)
5. Brunello, G., Michaud, P.-C., Sanz-de-Galdeano, A.: The rise of obesity in Europe: an economic perspective. Econ Policy **24**, 551–596 (2009)
6. Davidson, M.B., Schriger, D.L.: Effect of age and race/ethnicity on HbA1c levels in people without known diabetes mellitus: implications for the diagnosis of diabetes. Diabetes Res and Clin Pract **87**, 415–421 (2010)
7. Diabetes Prevention Program (DPP) Research Group: Within-trial cost-effectiveness of lifestyle intervention or Metformin for the primary prevention of type 2 diabetes. Diabetes Care **26**, 2518–2523 (2003)
8. Diabetes Prevention Program (DPP) Research Group: 10-year follow-up of diabetes incidence and weight loss in the DPPOS. Lancet **374**, 1677–1686 (2009)
9. Finkelstein, E.A., Ruhm, C.J., Kosa, K.M.: Economic causes and consequences of obesity. Ann Rev Public Health **26**, 239–257 (2005)
10. Finkelstein, E.A., Fiebelkorn, I.C., Wang, G.: National medical spending attributable to overweight and obesity: how much and who's paying? Health Affairs (web exclusive): W3-219-226 (2003)
11. Frey, I., Dapp, N., König, D., Deibert, P., Predel, H.G., Berg, A.: Weight management trough M.O.B.I.L.I.S., an exercise-based weight loss program: 2-year results. Deutsche Zeitschrift für Sportmedizin **61**, 19–22 (2010)
12. Frey, I., Berg, A., Grathwohl, D., Keul, J.: Freiburger Fragebogen zur körperlichen Aktivität–Entwicklung, Prüfung und Anwendung. Sozial- und Präventivmedizin **44**, 55–64 (1999)
13. Gifford, F.: The conflict between randomized clinical trials and the therapeutic obligation. J. Med. Philos. **11**(4), 347–366 (1986)

14. Göhner, W., Schlatterer, M., Seelig, H., Frey, I., Berg jr, A., Fuchs, R.: Two-year follow-up of an interdisciplinary cognitive-behavioral intervention program for obese adults. J. Psychol. **146**, 371–391 (2012)
15. Guh, D.P., Zhang, W., Bansback, N., Amarsi, Z., Birmingham, C.L., Anis, A.H.: The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC Public Health **9**, 88 (2009)
16. Häußler, J., Hankonen, N., Absetz, P.: Economic evaluation of the GOAL lifestyle intervention to prevent type 2 diabetes. Working Paper Series of the Department of Economics, University of Konstanz, 2015–2 (2015)
17. Icks, A., Rathmann, W., Haastert, B., Gandjour, A., Holle, R., John, J., Giani, G., KORA Study Group: Clinical and cost-effectiveness of primary prevention of type 2 diabetes in a "real-world" routine healthcare setting: model based on the KORA Survey 2000. Diabetic Med **24**, 473–480 (2007)
18. International Diabetes Federation: IDF Diabetes Atlas. Retrieved September 19, 2009. http://www.diabetesatlas.org/ (2009). Accessed 6 Dec 2013
19. IQWiG (2009): Entwurf einer Methodik für die Bewertung von Verhältnissen zwischen Nutzen und Kosten im System der deutschen gesetzlichen Krankenversicherung. Version 2.0. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), Köln
20. Klein, A., Chernyak, N., Brinks, R., Genz, J., Icks, A.: Kosteneffektivität der Primärprävention des Typ-2-Diabetes. Prävention und Gesundheitsförderung **6**, 102–110 (2012)
21. Konnopka, A., Bödemann, M., König, H.-H.: Health burden and costs of obesity and overweight in Germany. European J Health Econ **12**, 345–352 (2011)
22. Köster, I., Huppertz, E., Hauner, H., Schubert, I.: Direct costs of diabetes mellitus in Germany–CoDiM 2000-2007. Exp. Clin. Endocrinol. Diabetes **119**, 377–385 (2011)
23. Köster, I., von Ferber, L., Ihle, P., Schubert, I., Hauner, H.: The cost burden of diabetes mellitus: the evidence from Germany–the CoDiM Study. Diabetologia **49**, 1498–1504 (2006)
24. Lagerstrøm, D., Berg jr, A., Haas, U., Hamm, M., Göhner, W., Fuchs, R., Predel, H.G., Berg, A.: Das M.O.B.I.L.I.S.-Schulungsprogramm. Bewegungstherapie und Lebensstilintervention bei Adipositas und Diabetes. Diabetes aktuell **11**(1), 5–11 (2013)
25. Li, R., Zhang, P., Barker, L.E., Chowdhury, F.M., Zhang, X.: Cost-effectiveness of interventions to prevent and control diabetes mellitus: a systematic review. Diabetes Care **33**, 1872–1894 (2010)
26. Lindgren, P., Lindström, J., Tuomilehto, J., Uusitupa, M., Peltonen, M., Jönsson, B., de Faire, U., Hellenius, M.L., DPS Study Group: Lifestyle intervention to prevent diabetes in men and women with impaired glucose tolerance is cost-effective. Int. J. Technol. Assess. Health Care **23**, 177–183 (2007)
27. Mensink, G.B.M., Schienkiewitz, A., Haftenberger, M., Lampert, T., Ziese, T., Scheidt-Nave, C.: Übergewicht und Adipositas in Deutschland. Bundesgesundheitsblatt **56**, 786–794 (2013)
28. Norris, S.L., Zhang, X., Avenell, A., Gregg, E., Schmid, C.H., Lau, J.: Long-term non-pharmacological weight loss interventions for adults with prediabetes. Cochrane Database Syst. Rev. **18**, CD005270 (2005)
29. Ogden, C.L., Carroll, M.D., Kit, B.K., Flegal, K.M.: Prevalence of obesity in the United States, 2009–2010. NCHS data brief, no 82. Hyattsville, MD: National Center for Health Statistics (2012)
30. Paulweber, B., Valensi, P., Lindström, J., Lalic, N.M., Greaves, C.J., McKee, M., et al.: A European evidence-based guideline for the prevention of type 2 diabetes. Horm. Metab. Res. **42**(S 01), 3–36 (2010)
31. Rathmann, W., Strassburger, K., Heier, M., Holle, R., Thorand, B., Giani, G., Meisinger, C.: Incidence of Type 2 diabetes in the elderly German population and the effect of clinical and lifestyle risk factors: kORA S4/F4 cohort study. Diabetic Med **26**, 1212–1219 (2009)
32. Tuomilehto, J., Lindstrom, J., Eriksson, J.G., Valle, T.T., Hamalainen, H., Ilanne-Parikka, P., Keinanen-Kiukaanniemi, S., Laakso, M., Louheranta, A., Rastas, M., Salminen, V., Uusitupa, M.: Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N. Engl. J. Med. **344**, 1343–1350 (2001)
33. Wagner, G.G., Frick, J.R., Schupp, J.: The German Socio-Economic Panel Study (SOEP)–scope, evolution and enhancements. Schmollers Jahrbuch **127**(1), 139–169 (2007)
34. WHO: The global strategy on diet, physical activity and health. http://www.who.int/dietphysicalactivity/strategy/eb11344/en/index.html (2004). Accessed 18 Sept 2013