Long-Term Trajectories of Body Weight, Diet, and Physical Activity From Midlife Through Late Life and Subsequent Cognitive Decline in Women

Maude Wagner, Francine Grodstein, Cécile Proust-Lima*, and Cécilia Samieri*

* These authors contributed equally to this work.

WEB MATERIAL

WEB APPENDIX 1

Assessment of lifestyle risk factors since midlife

A-MeDi score. Usual dietary intakes consumed during the past year were collected through a 131-item semi-quantitative FFQ (1) in 1984, 1986, 1990, 1994, and 1998. At each time-point, the food item non-responses were imputed using the median of available data at that time-point (2). From 1984 to 1998, 40.8% to 52.3% women were not missing any food item, and only 3.1% to 9.3% were missing more than 4 items at any time-point of assessment (see Web Table 1 for count of missing values). We computed at each time-point an A-MeDi 9-point score that includes 9 components: vegetables (excluding potatoes), legumes, fruits, nuts, whole grains, fish, red or processed meats, alcohol, and the ratio of monounsaturated (MUFA) to saturated fatty acids (SFA). For alcohol (of which moderate consumption is hypothesized beneficial), a consumption between 5 and 15 g per day was assigned the value of 1, 0 otherwise. For red/processed meat, hypothesized detrimental to health, consumption above the median was assigned a value of 0, 1 otherwise. For all seven other components (hypothesized beneficial to health), consumption above the median was assigned the value of 1, 0 otherwise. The possible score ranged therefore between 0 and 9 (perfect adherence). All the repeated A-MeDi scores were defined using the median cut-offs calculated in year 1984 (i.e.; the first dietary assessment) in order to ensure a constant structure of the outcome over time.

Total Physical Activity. Detailed information on leisure-time physical activity was collected in 1986, 1988, 1992, 1994, 1996, 1998, and 2000 through a validated questionnaire (3). As with diet, at each time-point, the item non-responses were imputed using the median of available data at that time-point (2). From 1986 to 2000, 56.7% to 75.5% of women responded to all items, whereas 0.1% to 15.2% only responded to one item at any time-point of assessment (see Web Table 2 for count of missing values). Questionnaire of physical activity evolved during follow-up. From 1992 to 2000, two items were added in the questionnaire (i.e., other vigorous activities [e.g., lawn mowing] and low-intensity exercise [e.g., yoga, stretching, or toning]). Moreover, in 1994 the items "usual pace of walking" and "flights of stairs climbed daily" were not collected. In order to ensure comparable levels of physical activities over the time we (i) did not considered the two items added from 1992 to 2000 and (ii) we imputed in 1994 the values of pace and stairs-climbing collected in 1992 (if missing too, we considered the values of 1996). Women were asked to estimate the average time per week spent during the past year for: running (≥6 miles/hour); jogging (<6 miles/hour); bicycling (including stationary machine); racquet sports; lap-swimming; aerobic dance or use of exercise machines; and walking or hiking outdoors. There were 10 response-categories ranging from 0 minute to at least 11 hours. Participants also indicated their usual outdoor walking pace: easy (<2 miles/hour), normal (2–2.9 miles/hour), brisk (3–3.9 miles/hour), or very brisk (≥4 miles/hour) and the number of flights of stairs climbed daily. For each activity, we assigned a Metabolic Equivalent of Task (MET) according to accepted standards where 1 MET is proportional to the energy expended while sitting quietly (4). MET values were 12 for running; 8 for stair-climbing; 7 for jogging, bicycling, racquet sports, and lap-swimming; 6 for aerobic dance or use of exercise machine, and for walking values varied by reported pace (i.e., from 2.5 METs

[easy], to 4.5 METs [very brisk]). The energy expenditure in MET-hours per week corresponded to MET values multiplied by the time spent performing it.

Statistical analysis

Mean predicted levels of each risk factor (and 95% confidence intervals) given in Figure 3 were computed from the latent process mixed model by using a Monte Carlo approach with 2000 draws. Mean differences between groups at specific times (with 95% confidence intervals) were also computed using a Monte Carlo approach with 2000 draws (5).

We selected the optimal link function to relate the latent process to the observed measures among linear and quadratic I-splines (with 3 to 5 knots placed at the quantiles using the Akaike Information Criterion) (6).

In supplementary analysis, we explored for each factor whether potential differences in trajectories between cases and controls were confounded by the two other factors (in the 13-year time window preceding cognitive assessment that was common to all 3 risk factors of interest). We predicted individual levels of each factor at all the time points using mixed models from the primary analysis; then we adjusted the trajectory of each factor for the time-specific predicted levels of the two other factors, both as a simple effect and in interaction with the case-control status. To ensure comparability of these mutually adjusted trajectories with the trajectories of each factor obtained in the main analysis, we focused on the differences in trajectories between cases and controls, which do not depend on a specific profile of covariates. For each factor, we computed at each time point the case-control difference estimated in both approaches (i.e., analysis not adjusted versus adjusted for the two other factors) and compared them.

REFERENCES

1. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol*. 1985;122(1):51–65.
2. Parr CL, Hjartåker A, Scheel I, Lund E, Laake P, Veierød MB. Comparing methods for handling missing values in food-frequency questionnaires and proposing *k* nearest neighbours imputation: effects on dietary intake in the Norwegian Women and Cancer Study (NOWAC). *Public Health Nutr*. 2008;11(4):361–370.
3. Wolf AM, Hunter DJ, Colditz GA, Manson JE, Stampfer MJ, Corsano KA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol*. 1994;23(5):991–999.
4. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR, Montoye HJ, Sallis JF, et al. Compendium of Physical Activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc*. 1993;25(1):71–80.
5. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. *J Stat Softw*. 2017;(78):1–56.
6. Proust-Lima C, Amieva H, Jacqmin-Gadda H. Analysis of multivariate mixed longitudinal data: a flexible latent process approach. *Br J Math Stat Psychol*. 2013;66(3):470–487.

Web Table 1. Proportion of Responses to the Food Frequency Questionnaire and of Missing Items Used in the Computation of the Alternate-Mediterranean Diet Score at Each Year of Diet Assessment Among the Initial Study Population (*n* = 14,956), Nurses' Health Study, United States, 1984–1998

	Year of Food Frequency Questionnaire				
	1984	**1986**	**1990**	**1994**	**1998**
Response, no. (%)	12,236 (81.8)	11,565 (77.3)	12,655 (84.6)	13,737 (91.8)	12,173 (81.4)
Missing items (%)					
0	40.8	41.6	48.2	43.6	52.3
1	39.0	34.9	28.5	28.8	28.7
2	12.5	13.2	11.9	12.5	9.7
3	4.6	4.9	4.6	5.8	3.8
≥4	3.1	5.4	6.8	9.3	5.5

Note: The total number of food items ranged from 64 to 70 during 1984–1998.

Web Table 2. Number of Responses to the Physical Activity Questionnaire and Proportions of Missing Items Used in the Computation of the Total Physical Activity at Each Year of Assessment Among the Initial Study Population (*n* = 14,956), Nurses' Health Study, United States, 1986–2000 [a]

	Year of Physical Activity Questionnaire						
	1986 (10 Items)	**1988 (9 Items)**	**1992 (9 Items)**	**1994 (7 Items[a])**	**1996 (9 Items)**	**1998 (9 Items)**	**2000 (9 Items)**
Response, no. (%)	12,725 (85.1)	14,553 (97.3)	13,878 (92.8)	14,075 (94.1)	13,644 (91.2)	12,605 (84.3)	118 (0.8)
Missing items (%)							
0	74.1	56.7	72.3	71.6	75.5	75.4	70.3
1	5.3	3.5	5.2	4.1	6.4	6.1	5.9
2	0.6	0.7	0.6	0.6	0.5	0.6	0.9
3	0.4	0.6	0.4	0.4	0.3	0.4	1.7
4	0.5	2.7	1.4	1.6	1.1	0.9	4.2
5	1.7	9.7	5.9	6.6	4.4	4.3	14.4
6	6.2	16.2	12.9	15.2	10.6	10.8	2.5
7	10.8	0.8	1.3		1.3	1.4	
8	0.5	9.2	0.1		0.2	0.1	
9	0.1						

[a] Data on the items "usual walking pace" and "flights of stairs climbed daily" were not collected in 1994.

WEB APPENDIX 2

R Code for the Analyses of Body Mass Index

```
## This code applies to a dataset named NHS_bmi, which contains the longitudinal data (one row per nurse follow-up visit) and the following variables:

# bmi: continuous body mass index
# time: retrospective time preceding the ascertainment of cognitive decline (years)
# status: 0=control, 1=case of cognitive decline
# age0: age at inclusion given in decades and centered around 51 (mean age at study baseline)
# education: 0=registered nurse, 1=Bachelor's degree, 2=Master or doctorate
# ID: unique subject ID

R> library("lcmm","mvtnorm")


## a. ESTIMATION OF MODELS USING DIFFERENT LINK FUNCTIONS
# linear
R> m0 <- lcmm(bmi ~ time + I(time^2) + status + status*time + status*I(time^2) + age0 + age0*time + age0*I(time^2) + education + education*time + education*I(time^2), random=~ time + I(time ^2), subject="ID", data = NHS_bmi)
# quadratic I-splines with 3 knots placed at the quantiles of the distribution
R> m1 <- lcmm(bmi ~ time + … + education*I(time^2), link = c("3-quant-splines"), random=~ time + I(time ^2), subject="ID", data = NHS_bmi)
# with 4 knots placed at the quantiles
R> m2 <- lcmm(bmi ~ time + … + education*I(time^2), link = c("4-quant-splines"), random=~ time + I(time ^2), subject="ID", data = NHS_bmi)
# with 5 knots placed at the quantiles
R> m3 <- lcmm(bmi ~ time + … + education*I(time^2), link = c("5-quant-splines"), random=~ time + I(time ^2), subject="ID", data = NHS_bmi)

R> summary(m0)
R> summary(m1)
R> summary(m2)
R> summary(m3)
# selection of the model with the best Akaike Information Criterion: m2

## b. ASSESSMENT OF THE MODEL GOODNESS-OF-FIT
R> plot(m2)
# All subjects
R> plot(m2, which="fit", var.time="time", ylab="BMI, kg/m²")
R> plot(m2, which="fit", var.time="time", ylab="BMI, kg/m²", marg=F)
# Among groups
R> plot(m2, which="fit", var.time="time", subset=which(status==1), ylab="BMI, kg/m²", main="Cases")
R> plot(m2, which="fit", var.time="time", subset=which(status==0), ylab="BMI, kg/m²", main="Controls", marg=F)

## c. WALD-TESTS
# pos: vector containing the position in m2$best of the parameters to test
```

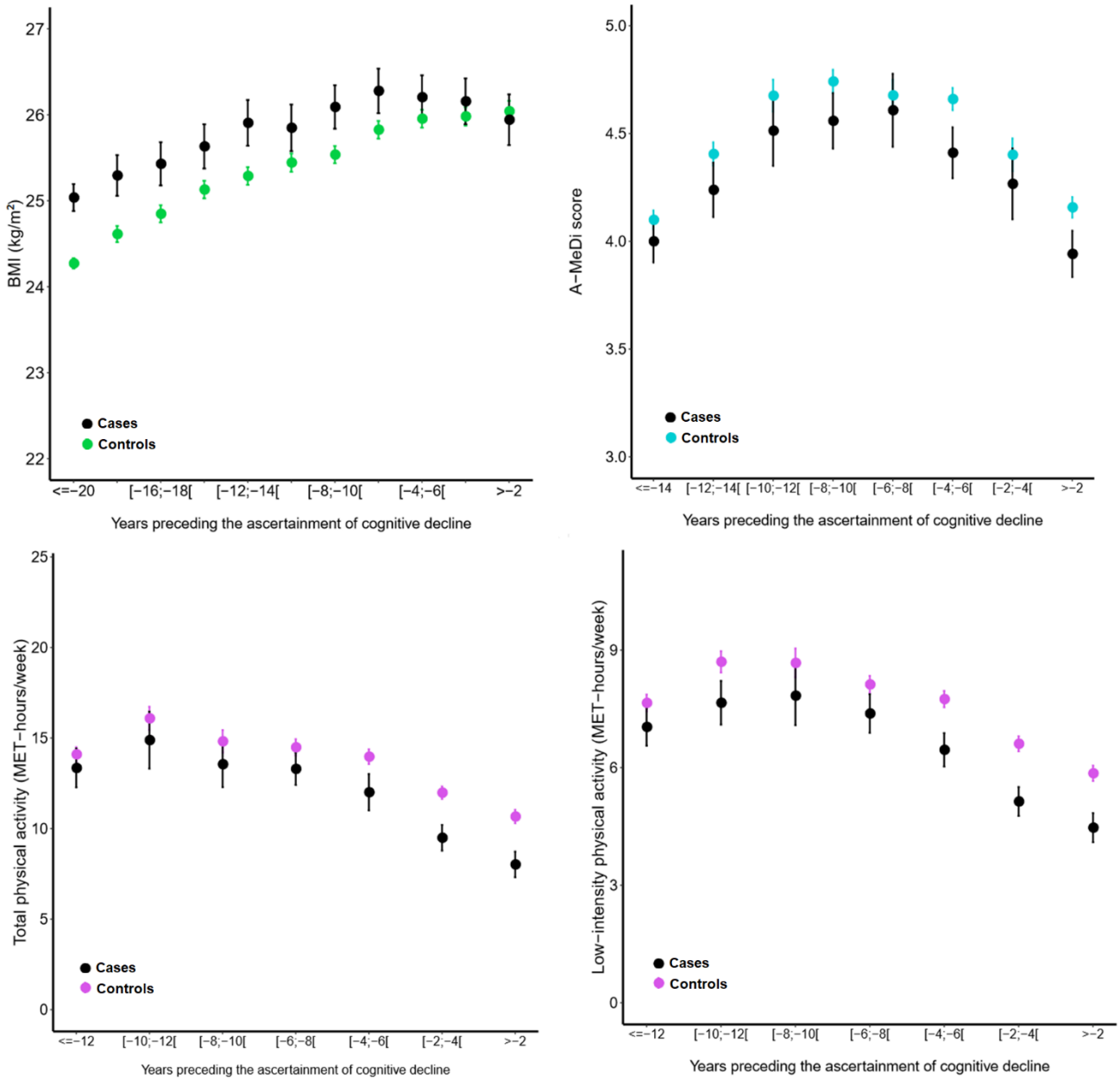
```
# contrasts: numeric vector of same length that pos (vector of 1 by default); the quantity to test
is the dot product of pos and contrasts.

# c.1. Global tests
# overall difference in the evolution among groups (i.e., status*time = status*I(time^2) = 0)
R> WaldMult(m2, pos=c(4,5))  # P value <.0001
# c.2. Tests for differences among groups at specific times
# In the following, function corrected_P gives the corrected significance level for multiple
testing in the comparison of trajectories among groups at different periods of time computed from
the joint distribution of all the statistics. model corresponds to the estimated model, pos
corresponds to the position of the parameters to test in model$best and tim corresponds to the
predefined sequence of times.

R> corrected_P <- function(model, pos, tim)
      {V <- VarCov(model)
      matv <- V[pos, pos]
      mat <- matrix(0, nrow=length(tim), ncol=length(pos))
      for (i in 1:length(tim))
            {mat[i,] <- c(1,tim[i],tim[i]*tim[i])
            }
      rho <- matrix(0, nrow=length(tim), ncol=length(tim))
      for (j in 1:length(tim))
            {for (k in 1:length(tim))
                  {rho[j,k] <- (t(mat[j,]) %*% matv %*% mat[k,])/(sqrt(t(mat[j,]) %*% matv
                   %*% mat[j,])*sqrt(t(mat[k,]) %*% matv %*% mat[k,])
                  }
            }
      threshold <- pnorm(-qmvnorm(p=0.95, tail='both.tails', corr=rho)$quantile)*2
      return(threshold)}

R> corrected_P(m2, pos=c(3,4,5), tim=c(-24,-22,-20,-18,-16,-14,-12,-10,-8,-6,-4,-2,0))
# The corrected significance threshold for bmi was 0.03
# Time-specific Wald tests were:
# 23 years before cognition (i.e., status= 1; status*time= -23y.; status*I(time^2)= (-23y)²)
R> WaldMult(m2, pos=c(3,4,5), contrasts=c(1,-23,529))    # P value <.0001
# 21 years before cognition (i.e., status= 1; status*time= -21y; status*I(time^2)= (-21y)²)
R> WaldMult(m2, pos=c(3,4,5), contrasts=c(1,-21,441))    # P value <.0001
# d.  PLOT OF PREDICTED TRAJECTORIES
# The mean trajectories (with 95% pointwise confidence intervals obtained by a Monte Carlo method
with 2000 draws) were displayed for the most common profile of the study sample (registered nurses,
aged 51 years-old in 1976).
# d.1. Creation of the profile for which trajectories are to be displayed
R> datnew <- data.frame(time = seq(-23, 0, length=100))
R> datnew$age0 <- 0
R> datnew$education <- 0
# d.2. Prediction of the trajectories for controls and cases of cognitive decline
R> datnew$status <- 0
R> controls <- predictY(m2, newdata=datnew, var.time="time", draws=T)
R> datnew$status <- 1
R> cases <- predictY(m2, newdata=datnew, var.time="time", draws=T)
# d.3. Plot of the trajectories
R> plot(controls, ylim=c(22,27), lwd=c(4,1), bty="l", las=1, col=2, xlab="Years preceding the
ascertainment of cognitive decline", ylab="BMI, kg/m²", legend=NULL)
R> plot(cases, col=1, lwd=c(4, 1), add=T)
```

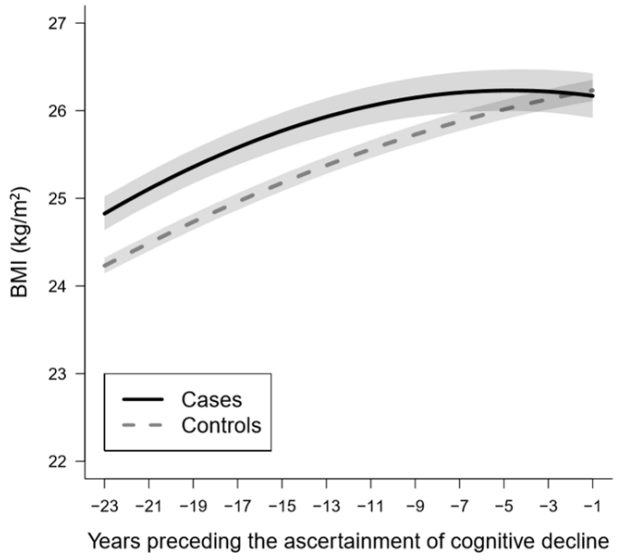
Web Figure 1. Mean and 95% CIs of the Observed Measures of Body Mass Index (BMI), Alternate Mediterranean Diet (A-MeDi) Score and Level of Physical Activity in the Years Preceding the Ascertainment of Cognitive Decline Among Cases of Cognitive Decline (*n* = 1,496) and Controls (*n* = 7,478), Nurses' Health Study, United States, 1976–2000.



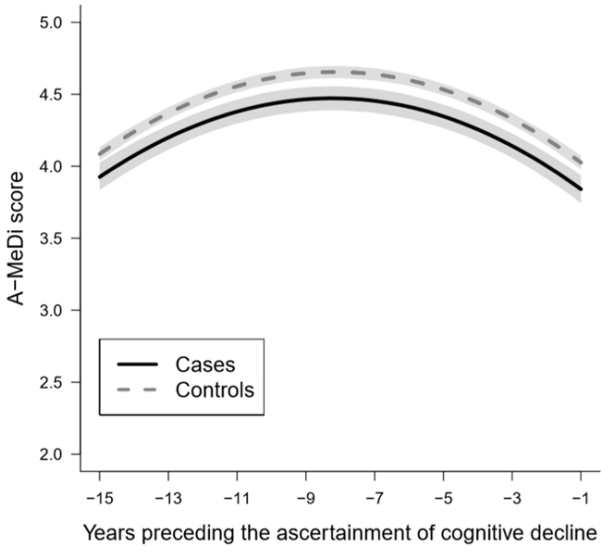
Web Figure 2. Trajectories of Body Mass Index (BMI), Alternate Mediterranean Diet (A-MeDi) Score, and Total and Low-Intensity Physical Activity, in the 23, 15, and 13 Years Preceding the Ascertainment of Cognitive Decline Among Cases of Cognitive Decline (*n* = 1,496) and Controls (*n* = 7,478) in the Nurses' Health Study, United States, 1976–2000.

Mean predicted trajectories (solid lines for cases, dotted lines for controls) with 95% pointwise confidence intervals (95% CI) (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time^2) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976). Outcomes were normalized by I-splines and the corrected threshold for multiple testing were $\alpha = 0.03$ for BMI, $\alpha = 0.02$ for A-MeDi and low-intensity physical activity, and $\alpha = 0.01$ for total physical activity.

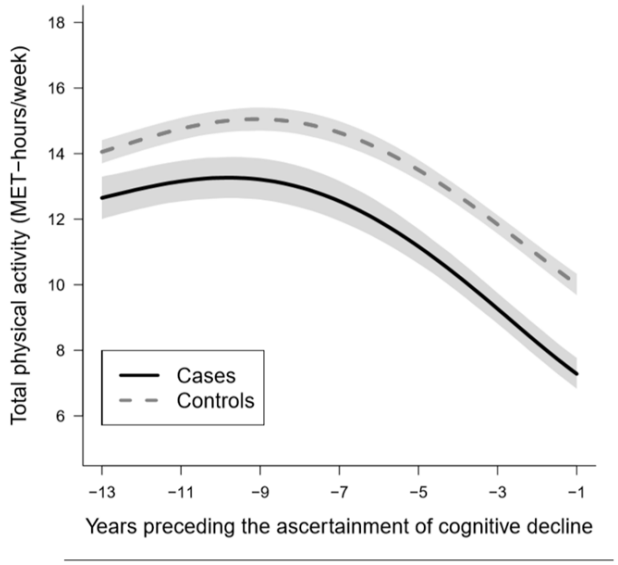
[Graph appears on next page.]



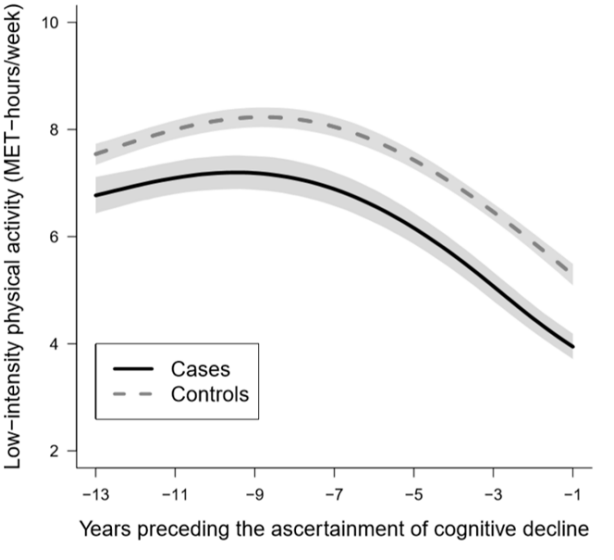
Mean difference compared to controls	0.59	0.62	0.63	0.62	0.59	0.55	0.49	0.42	0.32	0.22	0.09	- 0.07
CI 95% – lower limit	0.39	0.42	0.42	0.41	0.38	0.33	0.27	0.19	0.09	-0.03	-0.17	-0.33
CI 95% – upper limit	0.80	0.82	0.83	0.83	0.81	0.78	0.72	0.65	0.56	0.45	0.33	0.19



	-0.16	-0.17	-0.18	-0.18	-0.19	-0.19	-0.19	-0.19
	-0.26	-0.26	-0.27	-0.27	-0.28	-0.28	-0.28	-0.29
	-0.06	-0.08	-0.09	-0.10	-0.10	-0.10	-0.11	-0.09



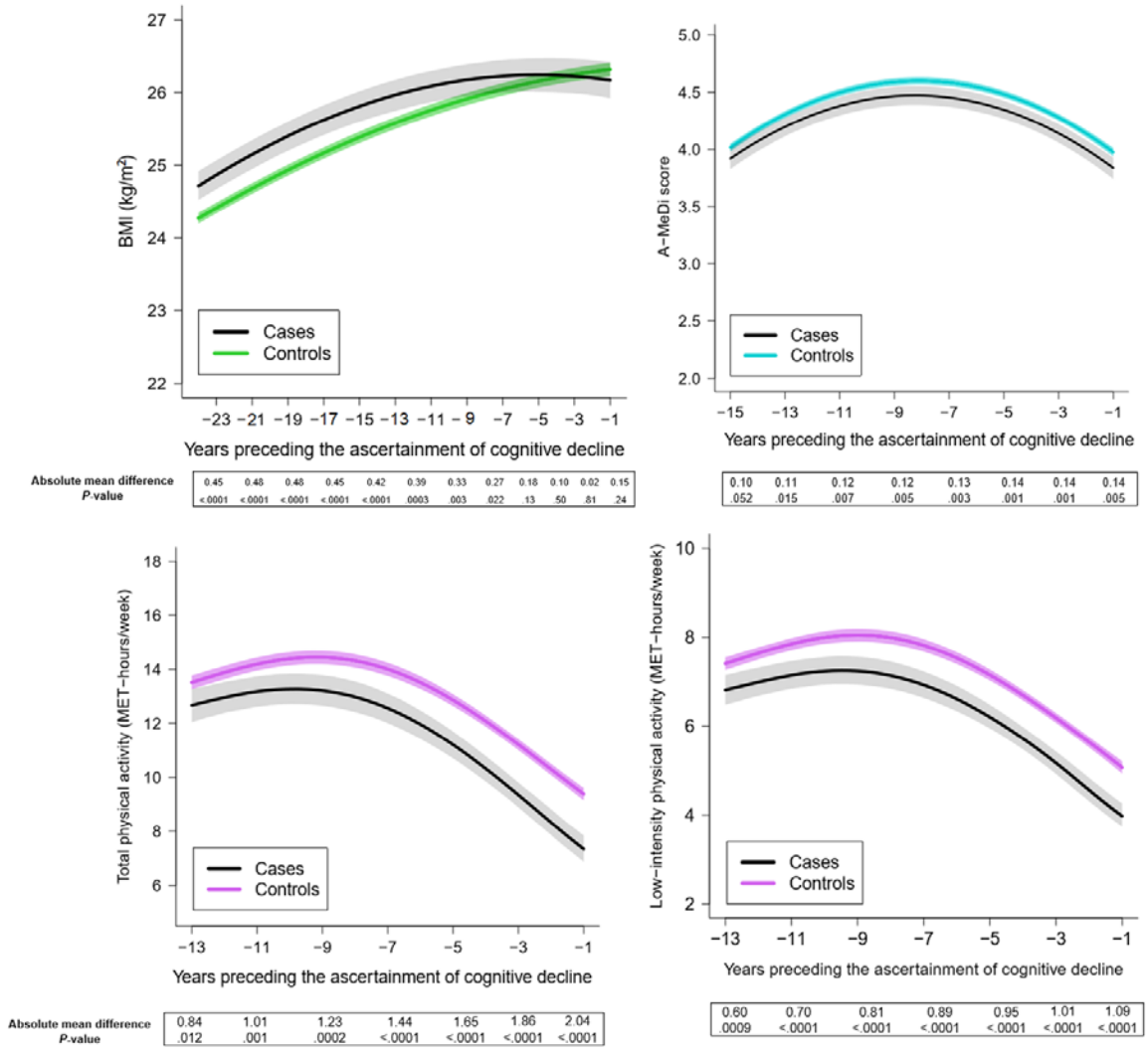
Mean difference compared to controls	-1.41	-1.59	-1.84	-2.08	-2.33	-2.56	-2.71
CI 95% – lower limit	-2.07	-2.21	-2.49	-2.69	-2.87	-3.06	-3.21
CI 95% – upper limit	-0.71	-0.97	-1.22	-1.46	-1.76	-2.03	-2.21



	-0.76	-0.89	-1.03	-1.15	-1.26	-1.38	-1.34
	-1.13	-1.22	-1.38	-1.49	-1.58	-1.67	-1.61
	-0.41	-0.55	-0.67	-0.80	-0.95	-1.10	-1.06

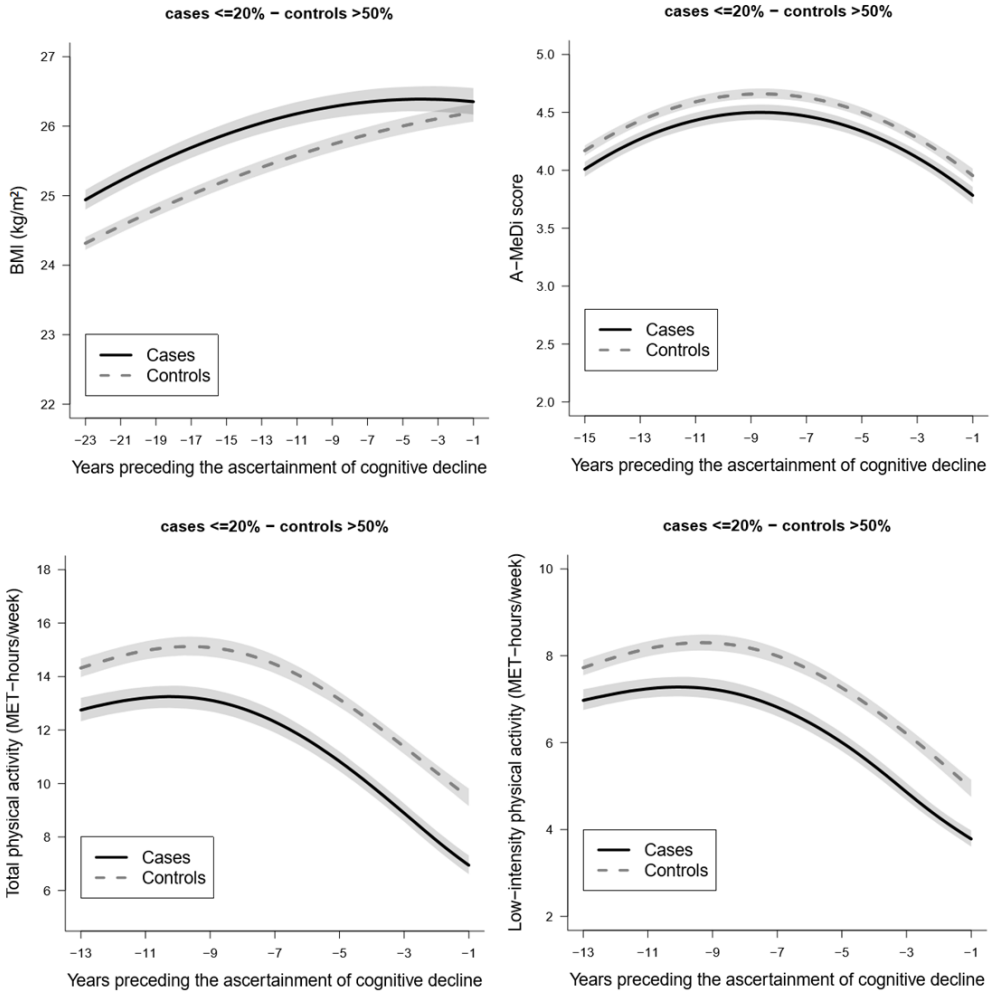
Web Figure 3. Trajectories of Body Mass Index (BMI), Alternate-Mediterranean Diet (A-MeDi) and Level of Physical Activity in the Years Preceding the Ascertainment of Cognitive Decline Among Cases of Cognitive Decline, (≤First Decile of Slope, *n* = 1,496) and Controls (>First Decile of Slope, *n* = 13,460), Nurses' Health Study, United States, 1976–2000.

Mean predicted trajectories (solid lines) with 95% pointwise CIs (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time^2) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976).



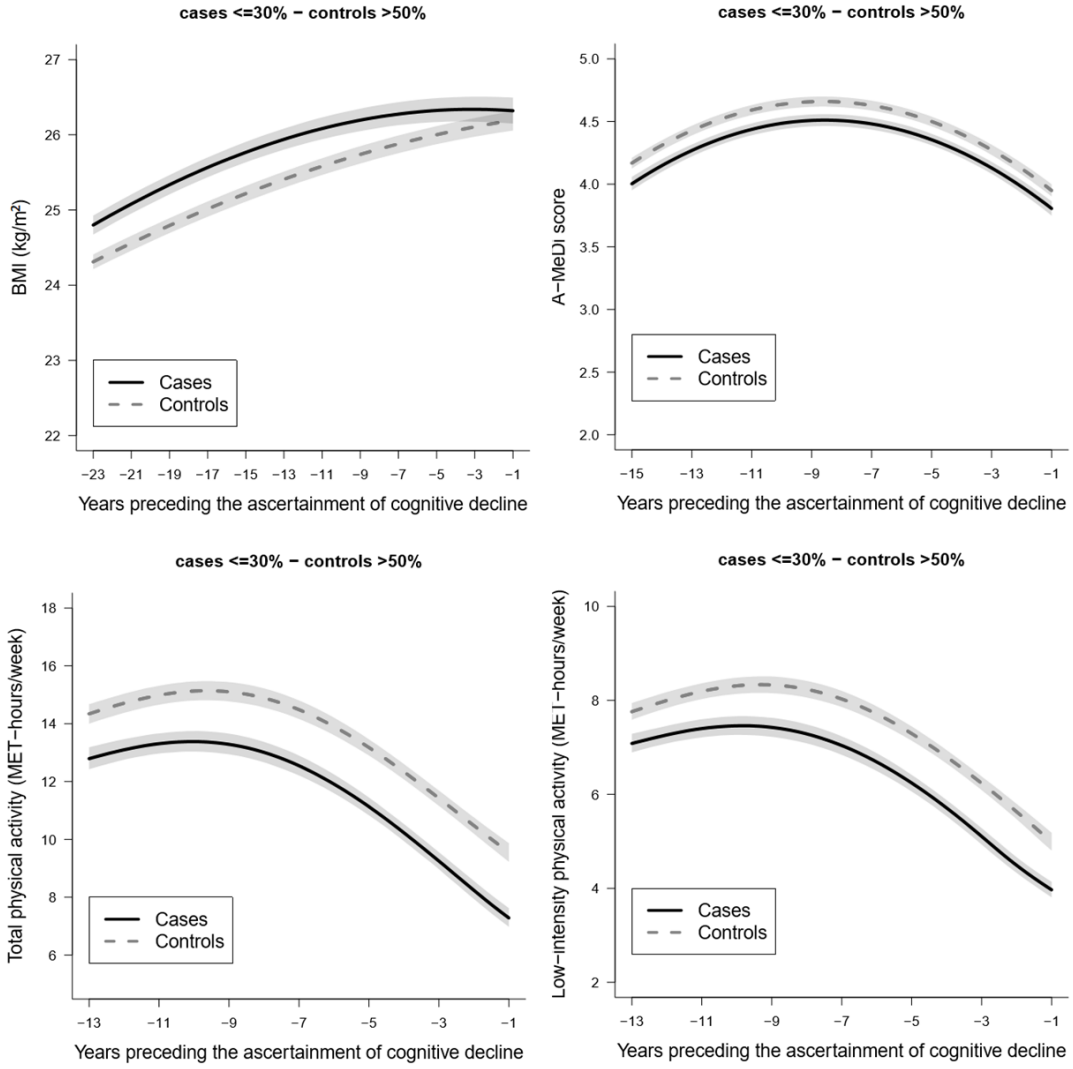
Web Figure 4. Trajectories of Body Mass Index (BMI), Alternate Mediterranean Diet (A-MeDi) Score, and Total and Low-Intensity Physical Activity, in the 23, 15, and 13 Years Preceding the Ascertainment of Cognitive Decline Among Cases of Cognitive Decline (≤Second Decile of Slope, n = 2,991) and Controls (> Median, n = 7,478) in the Nurses' Health Study, United States, 1976–2000.

Mean predicted trajectories (solid lines) with 95% pointwise CIs (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time^2) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976).



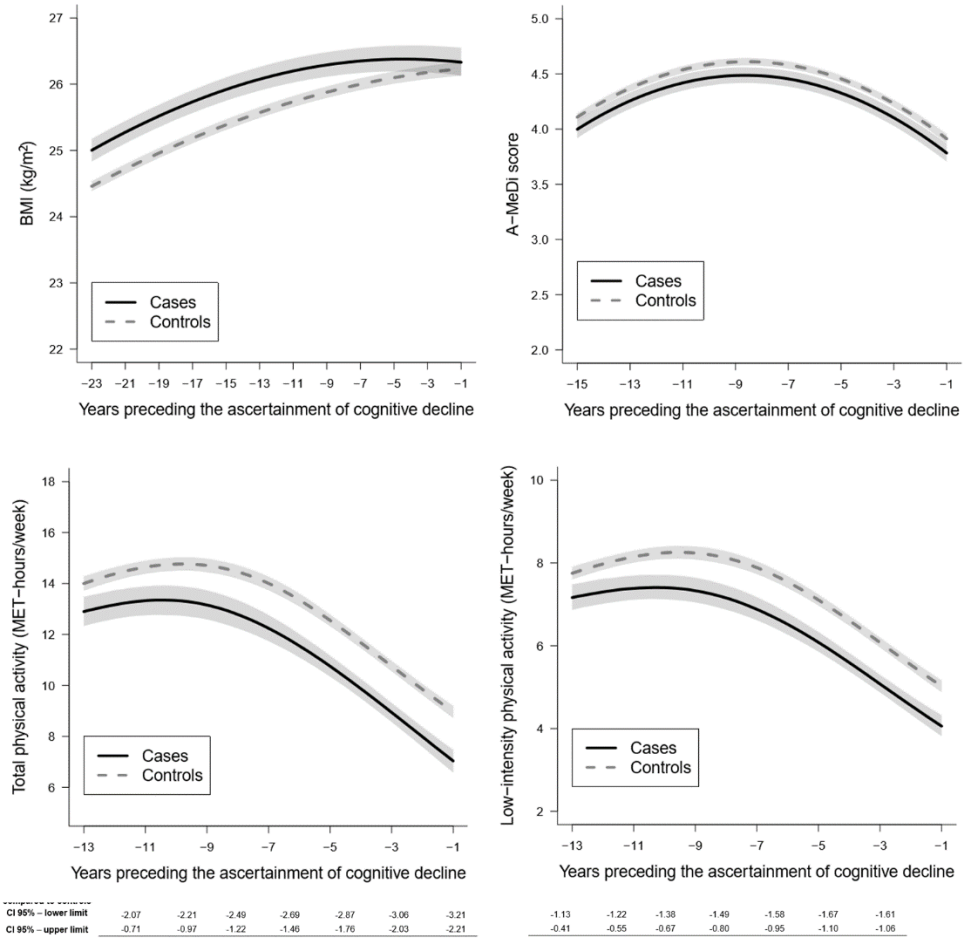
Web Figure 5. Trajectories of Body Mass Index (BMI), Alternate Mediterranean Diet (A-MeDi) Score, and Total and Low-Intensity Physical Activity, in the 23, 15, and 13 Years Preceding the Ascertainment of Cognitive Decline Among Cases of Cognitive Decline (≤Third Decile of Slope, *n* = 4,485) and Controls (> Median, *n* = 7,478) in the Nurses' Health Study, United States, 1976–2000.

Mean predicted trajectories (solid lines) with 95% pointwise CIs (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time^2) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976).



Web Figure 6. Trajectories of Body Mass Index (BMI), Alternate Mediterranean Diet (A-MeDi) Score, and Total and Low-Intensity Physical Activity, in the 23, 15, and 13 Years Preceding the Ascertainment of Cognitive Decline Among Cases of Cognitive Decline Defined From a Quadratic Cognitive Trajectory (n = 1,923) and Controls (n = 13,033) in the Nurses' Health Study, United States, 1976–2000.

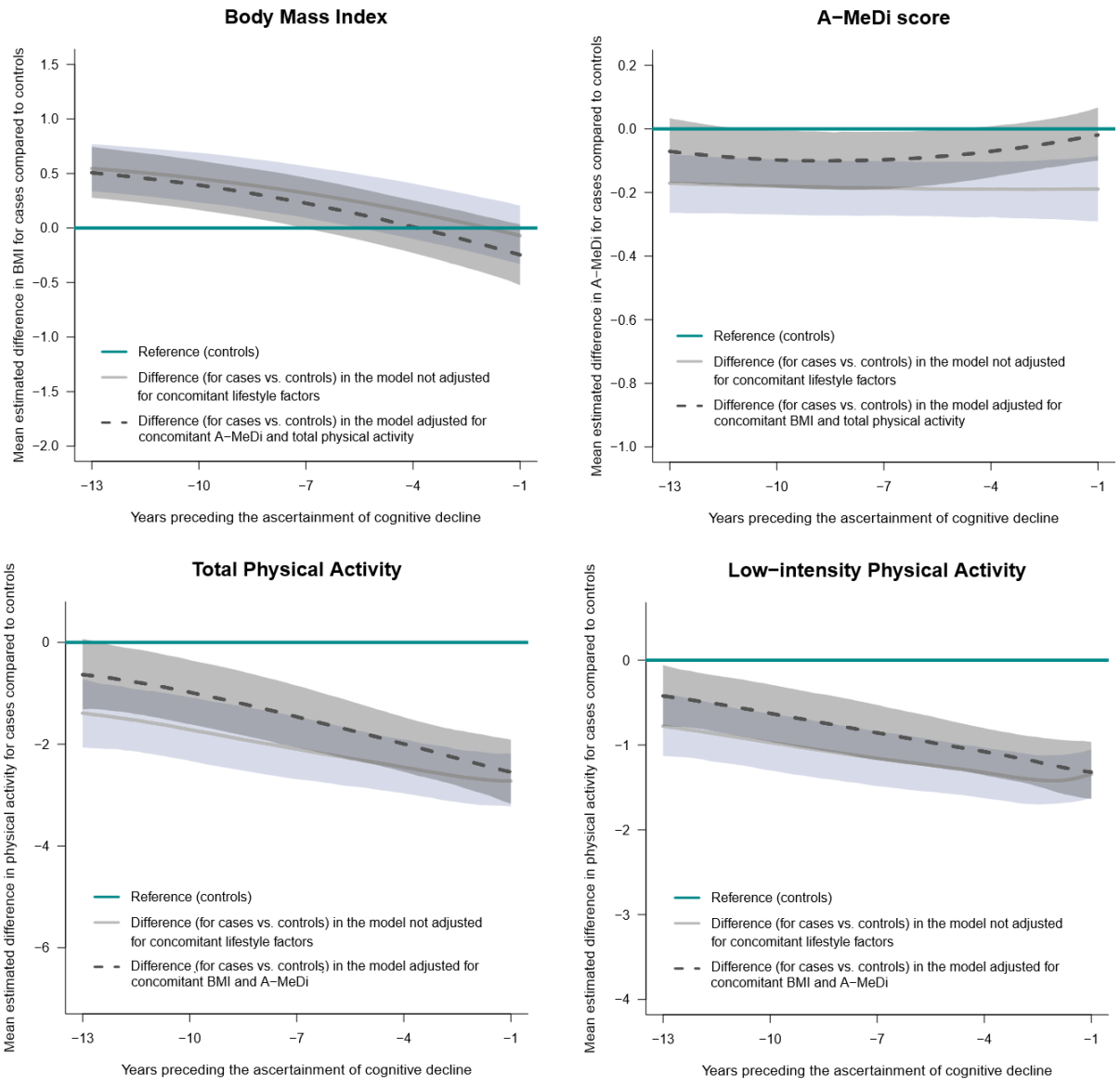
We defined cases and controls by considering a quadratic model for the cognitive change and including in the group of cases all the women who had either a linear slope or a quadratic slope of cognitive decline among the 10% worst. Controls correspond to all non-cases for simplicity. Mean predicted trajectories (solid lines) with 95% pointwise CIs (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time^2) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976).



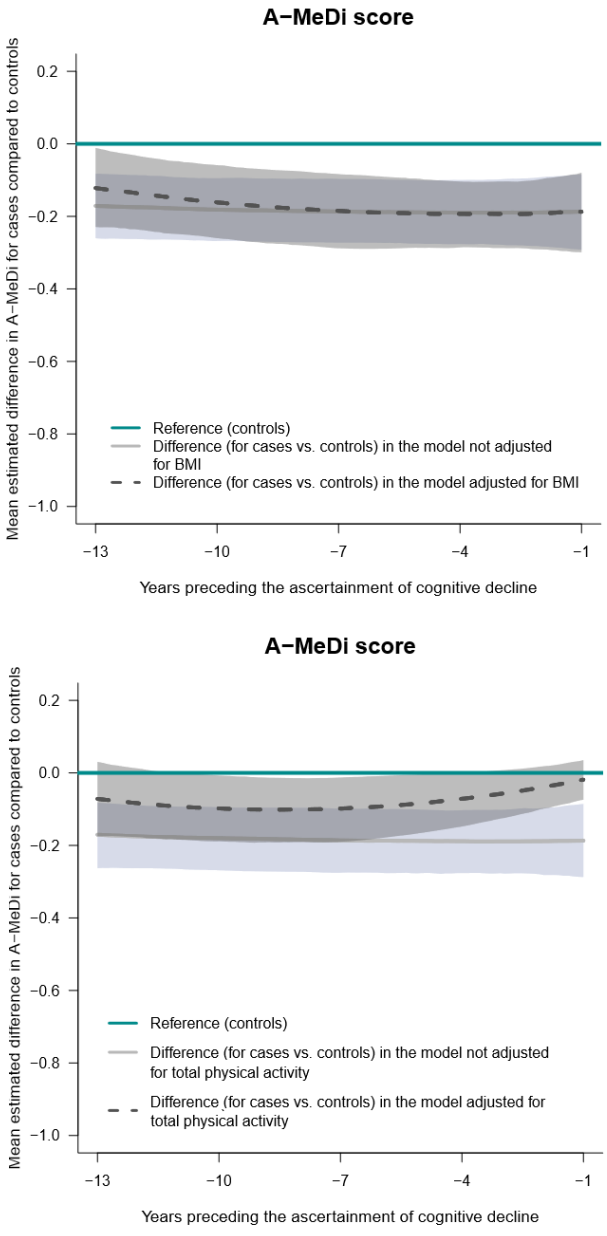
CI 95% – lower limit	-2.07	-2.21	-2.49	-2.69	-2.87	-3.06	-3.21
CI 95% – upper limit	-0.71	-0.97	-1.22	-1.46	-1.76	-2.03	-2.21

	-1.13	-1.22	-1.38	-1.49	-1.58	-1.67	-1.61
	-0.41	-0.55	-0.67	-0.80	-0.95	-1.10	-1.06

Web Figure 7. Estimated Differences in Trajectories of Body Mass Index (BMI), Alternate Mediterranean Diet (A-Medi) Score and Level of Physical Activity Between Cases (*n* = 1,496) And Controls (*n* = 7,478) in the 13-Year Time Window Preceding the Ascertainment of Cognitive Decline From Main Models or Main Models Further Adjusted for the Two Concomitant Other Factors (With Time-Specific Predicted Levels), Nurses' Health Study, United States, 1986–2000.

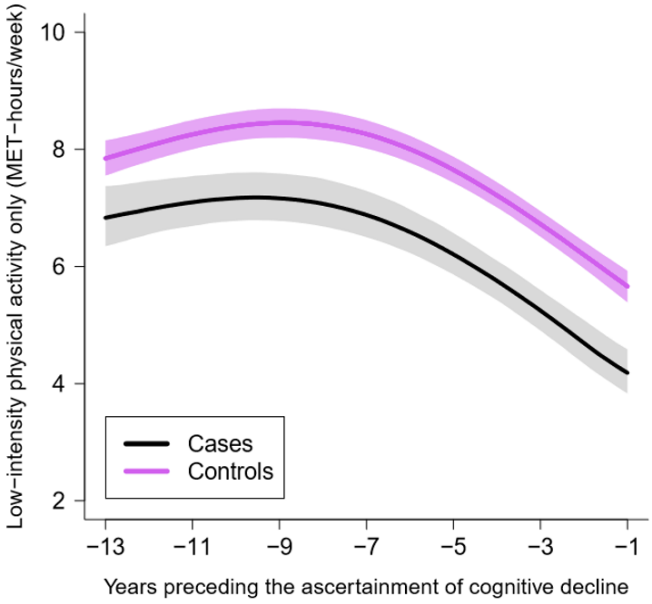


Web Figure 8. Estimated Differences in Trajectories of Alternate-Mediterranean Diet (A-MeDi) Score Between Cases (*n* = 1,496) and Controls (*n* = 7,478) in the 13-Year Time Window Preceding the Ascertainment of Cognitive Decline From Main Models or Main Models Further Adjusted for the two Concomitant Other Factors (With Time-Specific Predicted Levels), Nurses' Health Study, United States, 1986–2000.



Web Figure 9. Trajectories of Low-Intensity Physical Activity in the 13 Years Preceding the Ascertainment of Cognitive Decline Among Cases (*n* = 1,274) and Controls (*n* = 6,137) who Reported no Physical Activities Higher Than 6 METs per Week, Nurses' Health Study, United States, 1986–2000.

Mean predicted trajectories (solid lines) with 95% pointwise CIs (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time^2) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976).



Absolute mean difference	1.01	1.14	1.27	1.37	1.43	1.48	1.48
***P*-value**	<.001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Web Figure 10. Trajectories of Body Mass Index (BMI), Alternate-Mediterranean Diet (A-MeDi) and Level of Physical Activity in the Years Preceding the Ascertainment of Cognitive Decline Among Cases (*n* = 1,496) and Controls (*n* = 7,478) When Modelling the Trajectories With Natural Splines, Nurses' Health Study, United States, 1976–2000.

Mean predicted trajectories (solid lines) with 95% pointwise CIs (indicated with shading) were predicted by a latent process linear mixed model according to the years preceding cognitive decline ascertainment. The model included a quadratic function of time (time, time²) and was adjusted for case-control status, age and educational level and their interaction with the function of time; correlated random effects were considered on both the intercept and the time function. Trajectories were plotted for the most common profile of covariates in the study sample (registered nurse, aged 51 years in 1976).