Supplementary Online Content

Wagner M, Helmer C, Tzourio C, Berr C, Proust-Lima C, Samieri C. Evaluation of the concurrent trajectories of cardiometabolic risk factors in the 14 years before dementia. *JAMA Psychiatry*. Published online July 24, 2018.
doi:10.1001/jamapsychiatry.2018.2004

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Detailed Methodology

Statistical model

Latent process mixed models were used to model the mean trajectory of each cardiometabolic factor according to time (assuming a quadratic trajectory with time and time squared), case-control grouping and confounding factors. The latent process mixed model is a mixed model that handles possibly non-Gaussian longitudinal markers; the marker is transformed to correct the departure from the normal distribution using parameterized link functions.[1] For each cardiometabolic risk factor, the optimal link function was selected according to the Akaike Information Criterion among link functions based on quadratic I-splines with three to five knots and the linear link function.

For SBP and DBP specifically, models included two additional binary indicators (measure taken at T_0 *versus* later; measure based on the average of two measures *versus* one measure) to control for potential "white coat effect" (i.e., the transient increase in blood pressure induced by the stress associated with a primary clinical examination). Such white coat effect on blood pressure measures in our cohort is illustrated in **eFigure 1**.

As generally recommended when there is no intuitive reference group for a covariate, the 3-category study center variable was coded using a contrast in which the reference was set to the average of covariate modalities. Thus, the reference group belonged to an "average" study center in the analyses.

Wald tests of the differences in trajectories between cases and controls

We tested the differences in trajectories between cases and controls using Wald tests. Two types of comparisons were made. First, we compared risk factor trajectories over the entire time period, through a global evaluation of group-by-time interactions (i.e., simultaneous evaluation of interactions parameters between case-control status and both time and time squared). Second, we compared mean predicted risk factor values at different time points. For the latter, we primarily used the nominal 5% significance level. In secondary analyses, we also used a corrected threshold (based on the joint distribution of the statistics over pre-specified time points [two-year backward intervals from dementia diagnosis]) to account for the multiple testing.[2]

Supplementary analyses: consideration of medication use

For blood pressure, plasma lipids and glycemia, we explored whether differences in trajectories between cases and controls were modified by medication use. We considered two supplementary analyses:

1. one for polymedication at baseline. We defined a binary variable for the use of ≥4 medications at baseline *versus* less;
2. one for factor-specific medication use during follow-up. We created at each visit three binary indicators for use of each factor-specific medication (i.e., antihypertensive medication for analyses of blood pressure, lipid-lowering medication for analyses of plasma lipids, antidiabetic medication for analyses of glycemia), and we summarized the repeated indicators into three binary variables reflecting medication use (yes/no) at least once between cohort baseline and the matching visit for each factor-specific medication (distribution of the variables in both groups is given in **eTable 1**).

In the two analyses, we added the binary variable as a supplementary covariate in the models. We also included the second-order interaction of medication use with case-control status, and the two third-order interactions of medication use with both case-control status and each of the two time variables (time and time squared). Finally, we specifically evaluated whether medication use variable modified the differences in trajectories between cases and controls by testing the global significance of the three interaction terms simultaneously.

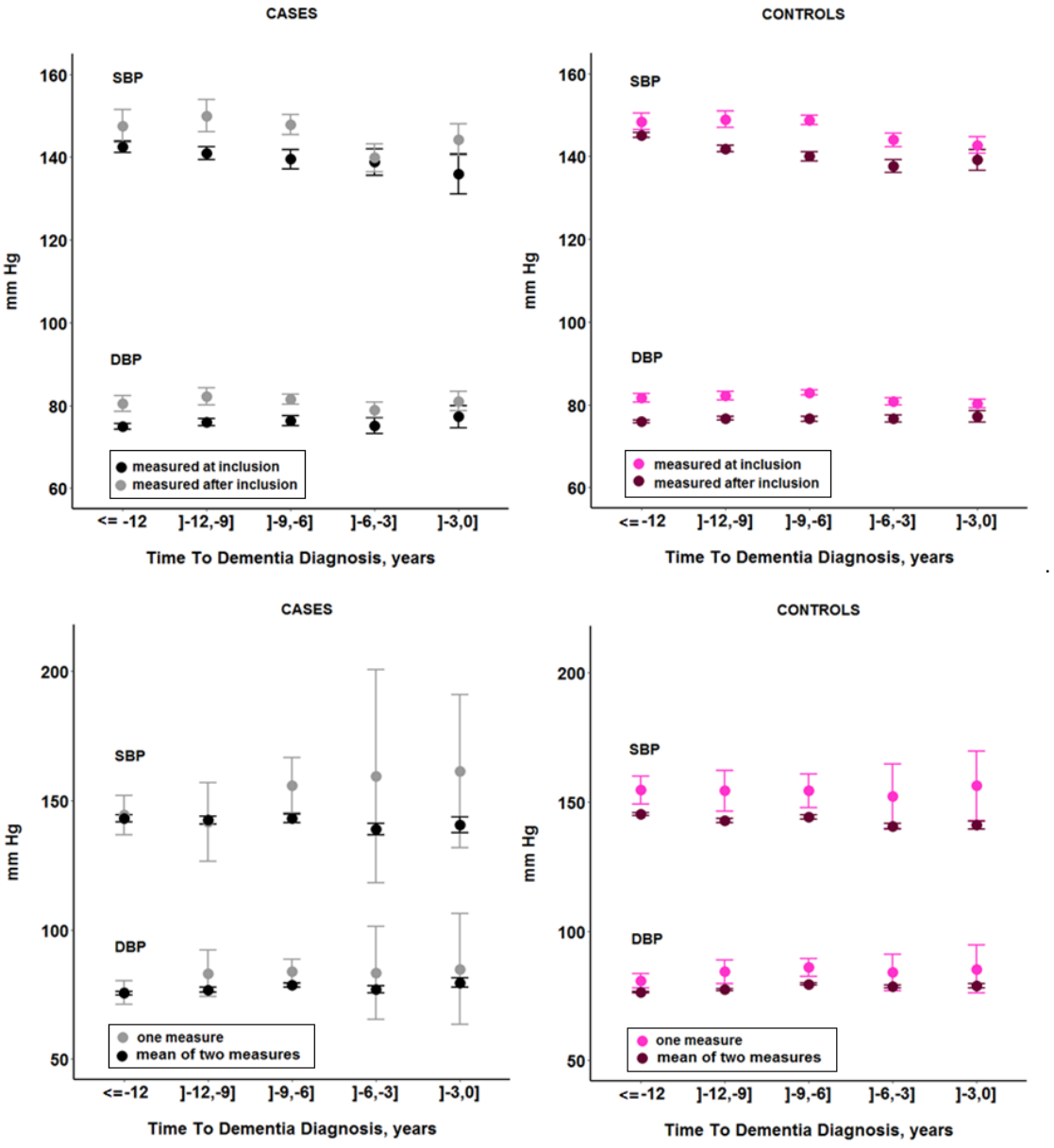
eTable 1. Consumption of at Least One Cardiometabolic Treatment Medication Between Baseline and the Matching Visit of Incident Dementia Cases and Matched Controls.

Characteristic	Cases (n = 785)	Controls (n = 3140)
Antihypertensive	574 (73.1)	2240 (71.3)
Lipid lowering	380 (48.4)	1429 (45.5)
Antidiabetic	116 (14.8)	277 (8.8)

Supplementary analyses: robustness to model hypotheses

For triglycerides and glycemia, which had extreme asymmetric distributions, we investigated the ability of the quadratic I-splines link functions to correctly normalize the data by comparing our results with trajectories obtained when log-transforming the data before applying the normalization. We also compared the quadratic trajectories with more flexible functions of time (regression splines with two internal knots).

eFigure 1. Means and 95% Confidence Intervals of the Observed Measures of Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP) in the Retrospective Time Since the Diagnosis of Dementia Visit by Periods of 3 Years for Incident Cases (*N*=785, Left Panels) and Matched Controls (*N*=3140, Right Panels), According to the Measurement Occasion (Initial Visit or Subsequent Visits, Top Panels) and to the Number of Measures at Any Visit (One Measure or Mean of Two Measures, Bottom Panel), the Three-City Study, France, 1999-2014.



eAppendix. R Code for the Analyses of Body Mass Index.

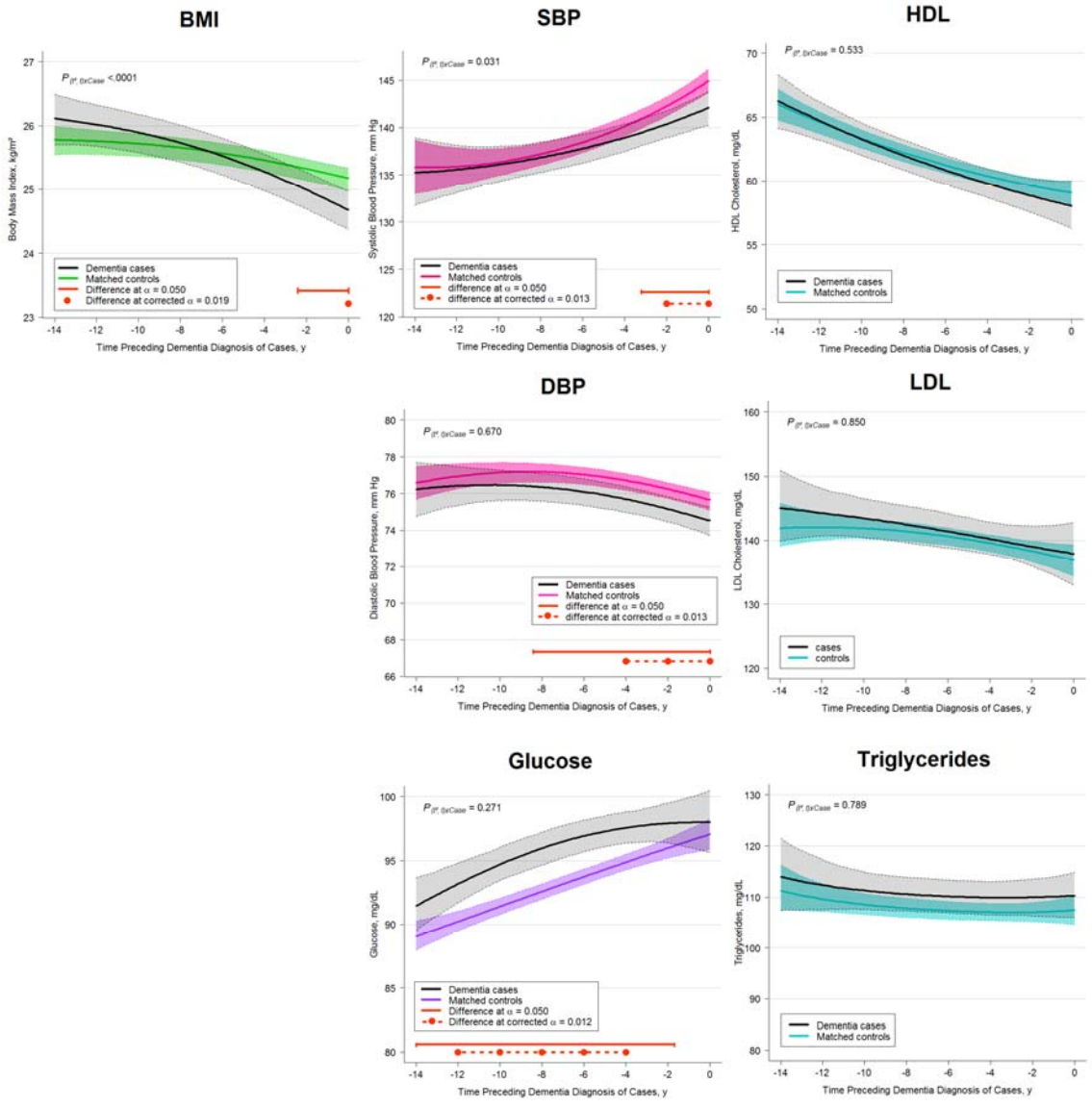
```
## This code applies to a dataset named 3C_bmi, which contains the longitudinal data (one row per individual follow-up visit) and the following variables:

# bmi: continuous body mass index
# time: retrospective time since diagnosis visit of dementia of cases and matching visit of controls
# status: 0=control, 1=case
# gender: 0=women, 1=men
# age0: age at inclusion given in decades and centered around 76 (mean age at study baseline)
# education: 0=below high school, 1=high school diploma, 2=higher than high school diploma
# center1/center2: two continuous variables (center1: 0=Bordeaux, 1=Dijon, -1=Montpellier; center2: 1=Bordeaux, 0=Dijon, -1=Montpellier) parameterized to consider mean cohort in reference
# 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) + center1 + center1*time + center1*I(time^2) + center2 + center2*time + center2*I(time^2) + gender + gender*time + gender*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 = 3C_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 = 3C_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 = 3C_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 = 3C_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 in both groups (i.e., time = I(time^2) = 0)
R> WaldMult(m2, pos=c(2,3))                              # 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))
```



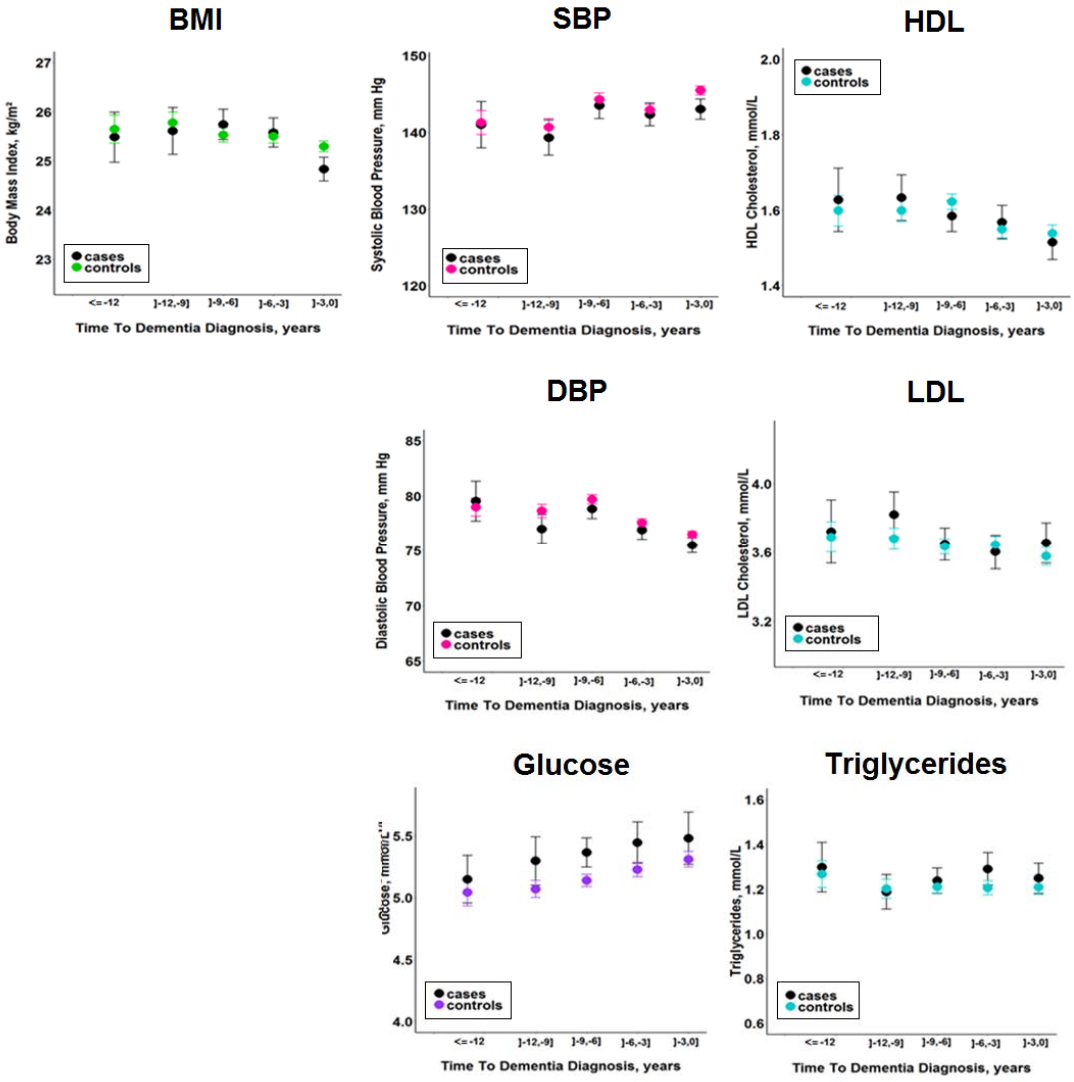
```
            {mat[i,] <- c(1,tim[i],tim[i]*tim[i])
            }
      rho <- matrix(0, nrow=length(tim), ncol=length(tim))
      for (j in 1:length(tim))
            {for (k in 1:length(tim))
                  {rho[j,k] <- (t(mat[j,]) %*% matv %*% mat[k,])/(sqrt(t(mat[j,]) %*% matv
                   %*% 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(9,22,23), tim=c(0,-2,-4,-6,-8,-10,-12,-14))
# The corrected significance threshold for bmi was 0.019
# Time-specific Wald tests were:
# at the matching visit (time=0) (i.e., status= 1)
R> WaldMult(m2, pos=c(9))                                    # P value = 0.001
# 2 years before the matching visit (i.e., status= 1; status*time= -2; status*I(time^2)= 4)
R> WaldMult(m2, pos=c(9,22,23), contrasts=c(1,-2,4))    # P value = 0.026
# 4 years before the matching visit (i.e., status= 1; status*time= -4; status*I(time^2)= 16)
R> WaldMult(m2, pos=c(9,22,23), contrasts=c(1,-4,16))   # P value = 0.237

# 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
(woman, 76 years-old at baseline, educational level lower than high school, and mean cohort).
# d.1. Creation of the profile for which trajectories are to be displayed
R> datnew <- data.frame(time = seq(-14, 0, length=100))
R> datnew$age0 <- 0
R> datnew$gender <- 0
R> datnew$center1 <- 0
R> datnew$center2 <- 0
R> datnew$education <- 0
# d.2. Prediction of the trajectories for controls and cases
R> datnew$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="Time To Dementia
Diagnosis, years", ylab="Body Mass Index, kg/m²", legend=NULL)
R> plot(cases, col=1, lwd=c(4, 1), add=T)
```


eFigure 2. Trajectories of BMI, SBP, DBP, Glucose, HDL and LDL Cholesterol, and Triglycerides Among Incident Cases of Dementia (*n*=785) and Matched Controls (*n*=3140) in the 14 Years Preceding the Matching Visit (Visit of Dementia Diagnosis) From a Non-Adjusted Model for Confounding Factors, the Three-City Study, France, 1999-2014.

Mean trajectories with 95% pointwise confidence intervals were predicted by a latent process linear mixed model in the retrospective time since the matching visit. The models included: a quadratic function of time (t, t^2); case-control status and its interaction with the quadratic function of time; correlated random effects on the intercept, time and time squared (models for SBP and DBP also included two binary indicators for white coat effect [value measured at T_0 yes/no, and value based on average of two measures versus one measure]). Observations were normalized by I-spline. Trajectories were plotted for the general profile of the study sample. $P_{(t^2,t)xCase}$ corresponds to the *p*-value of the Wald test evaluating the overall difference in change over time across groups (simultaneous evaluation of time and time squared parameters).



eFigure 3. Means and 95% Confidence Intervals of the Observed Repeated Measures of BMI, SBP, DBP, HDL and LDL Cholesterol, Triglycerides and Glucose in the Retrospective Time Since the Diagnosis of Dementia Visit by Periods of 3 Years Among Incident Cases of Dementia (N=785) and Matched Controls (N=3140) in the 14 Years Preceding the Matching Visit (Visit ff Dementia Diagnosis), the Three-City Study, France, 1999-2014.

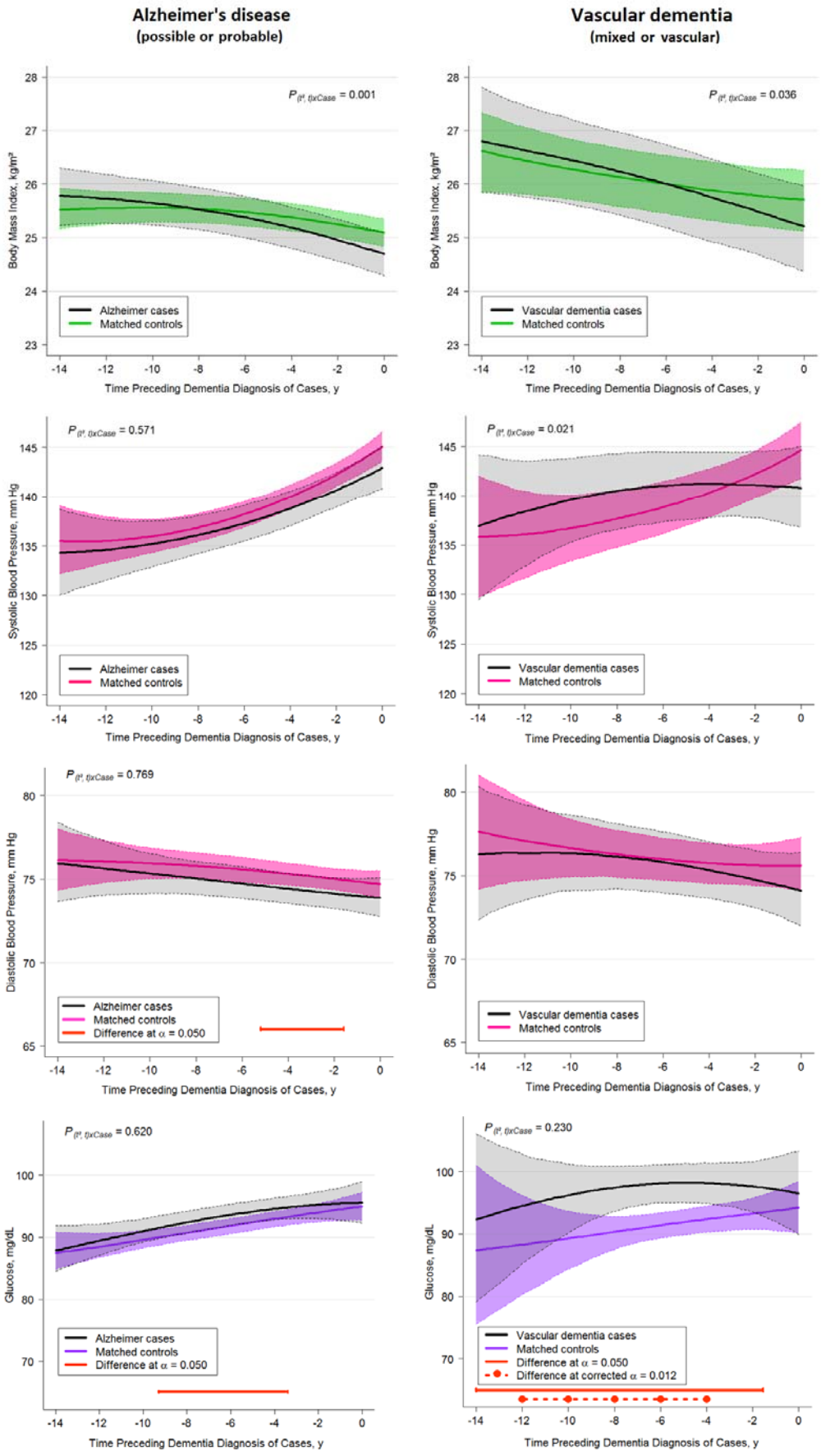


eTable 2. Table Summarizing the Statistical Tests of Difference Between Cases and Controls Trajectories Over the Entire Time Period (Simultaneous Evaluation of Interactions Parameters Between Both Time and Time Squared and Case-Control Status; $P_{(t,t^2)*Case}$) and at Specific Times for Body Mass Index (BMI), Systolic and Diastolic Blood Pressure (SBP/DBP), HDL and LDL Cholesterol, Triglycerides and Glycemia by Using Wald Tests, the Three-City Study, France, 1999-2014

	P-value of group-by-time interaction	P-value at specific times prior to diagnosis (in years)							
	$P_{(t,t^2)*Case}$	-14	-12	-10	-8	-6	-4	-2	0
BMI	**<.0001**	0.134	0.140	0.257	0.603	0.792	0.237	**0.026**	**0.001**
SBP	**0.049**	0.709	0.789	0.809	0.676	0.385	0.100	**0.007**	**0.001**
DBP	0.575	0.818	0.400	0.102	**0.022**	**0.007**	**0.003**	**0.003**	**0.024**
HDL	0.573	0.709	0.850	0.900	0.628	0.406	0.258	0.242	0.351
LDL	0.914	0.372	0.285	0.294	0.373	0.458	0.581	0.782	0.914
Triglycerides	0.797	0.885	0.532	0.255	0.141	0.093	0.090	0.214	0.518
Glycemia	0.294	**0.045**	**<.001**	**<.0001**	**<.0001**	**<.0001**	**<.0001**	**0.016**	0.408

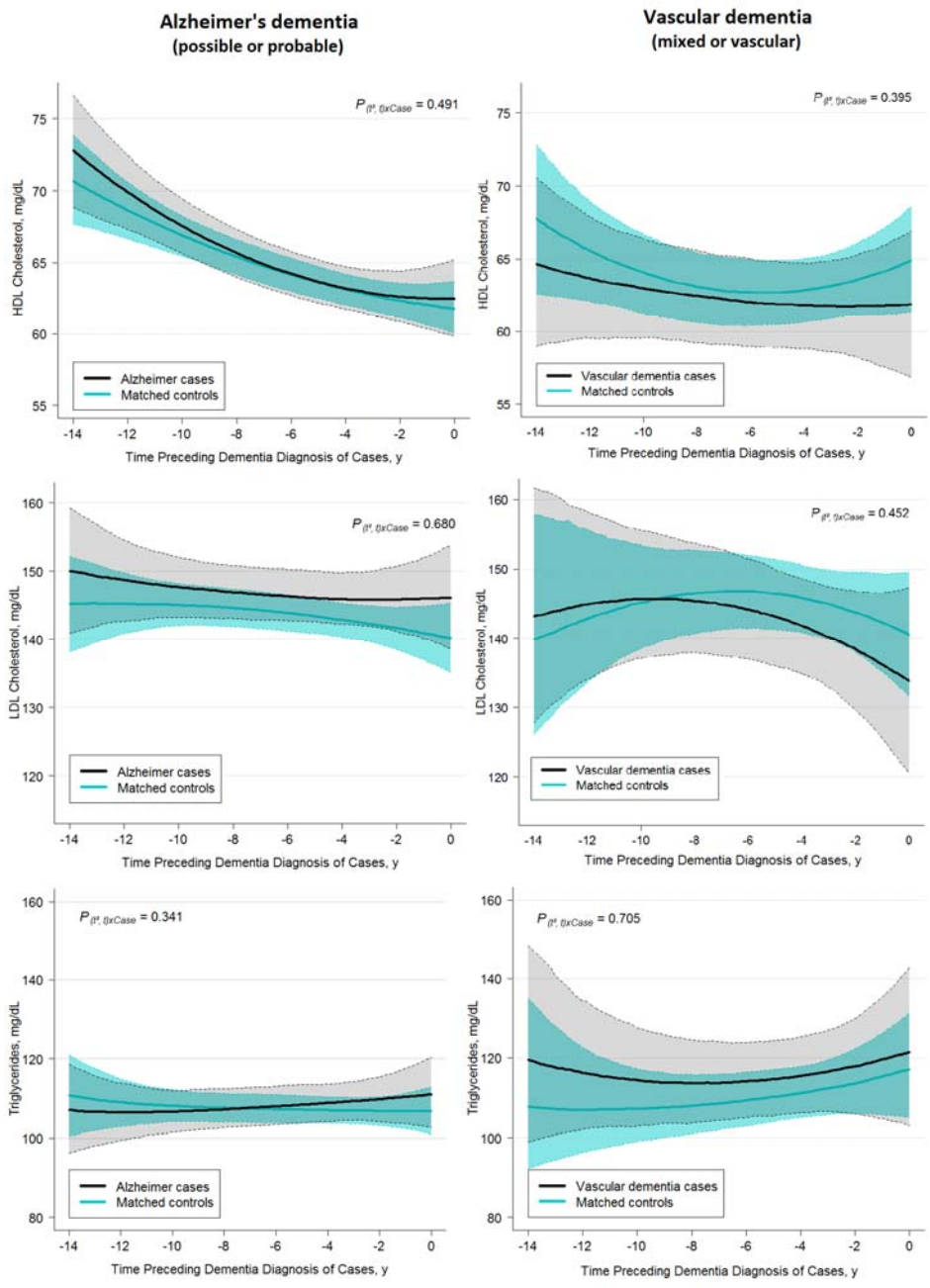
eFigure 4. Trajectories of BMI, SBP, DBP, Glucose, HDL, LDL Cholesterol and Triglycerides Among Alzheimer's Disease Subsample (Left Panel) and Vascular Dementia Subsample (Right Panel) in the 14 Years Preceding the Matching Visit (Dementia Diagnosis of the Case). Left Panel: Incident Cases of Possible or Probable Alzheimer (N=537) and Matched Controls (N=2148); Right Panel: Incident Cases of Mixed or Vascular Dementia (*n*=162) and Matched Controls (*n*=648), the Three-City Study, France, 1999-2014.

Mean trajectories with 95% pointwise confidence intervals were predicted by a latent process linear mixed model in the retrospective time since the matching visit. The models included: a quadratic function of time (t, t^2); case-control status, matching variables (i.e., gender, age, education, and cohort center), and their interactions with the quadratic function of time; correlated random effects on the intercept, time and time squared. Observations were normalized by I-splines. Trajectories were plotted for the most common profile of the study sample: a woman from an average study center, aged 76 years at inclusion and with educational level<high school. Note that the choice of the profile only impacts the level of the trajectories; it does not affect the differences between cases and controls or the tests significance. $P_{(t^2,t)xCase}$ corresponds to the *p*-value of the Wald test evaluating the overall difference in change over time across groups (i.e., simultaneous evaluation of time and time squared parameters).



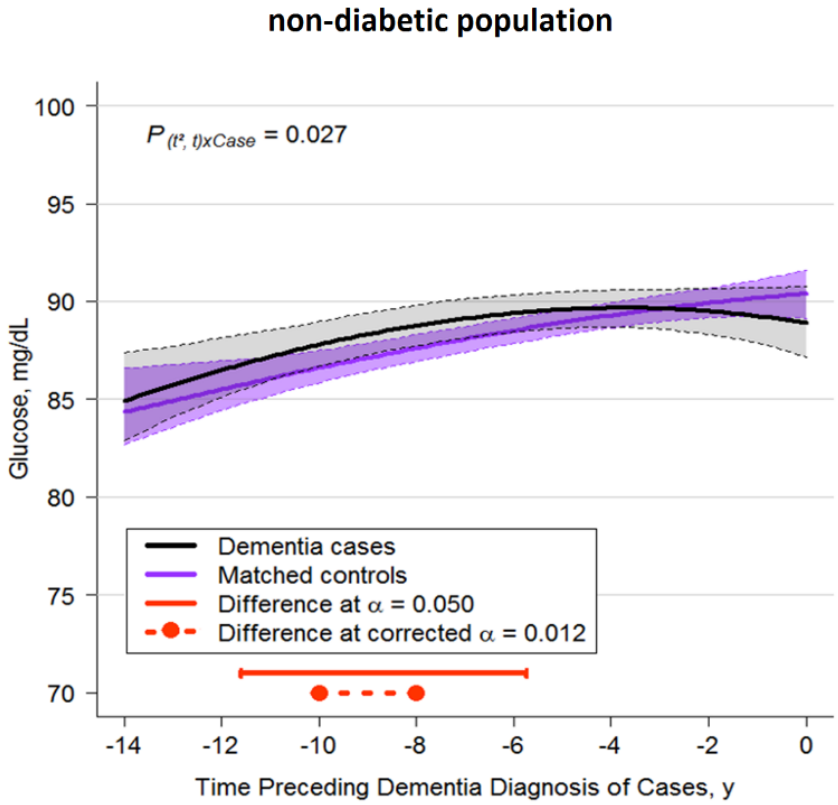
served. 12

(continue)



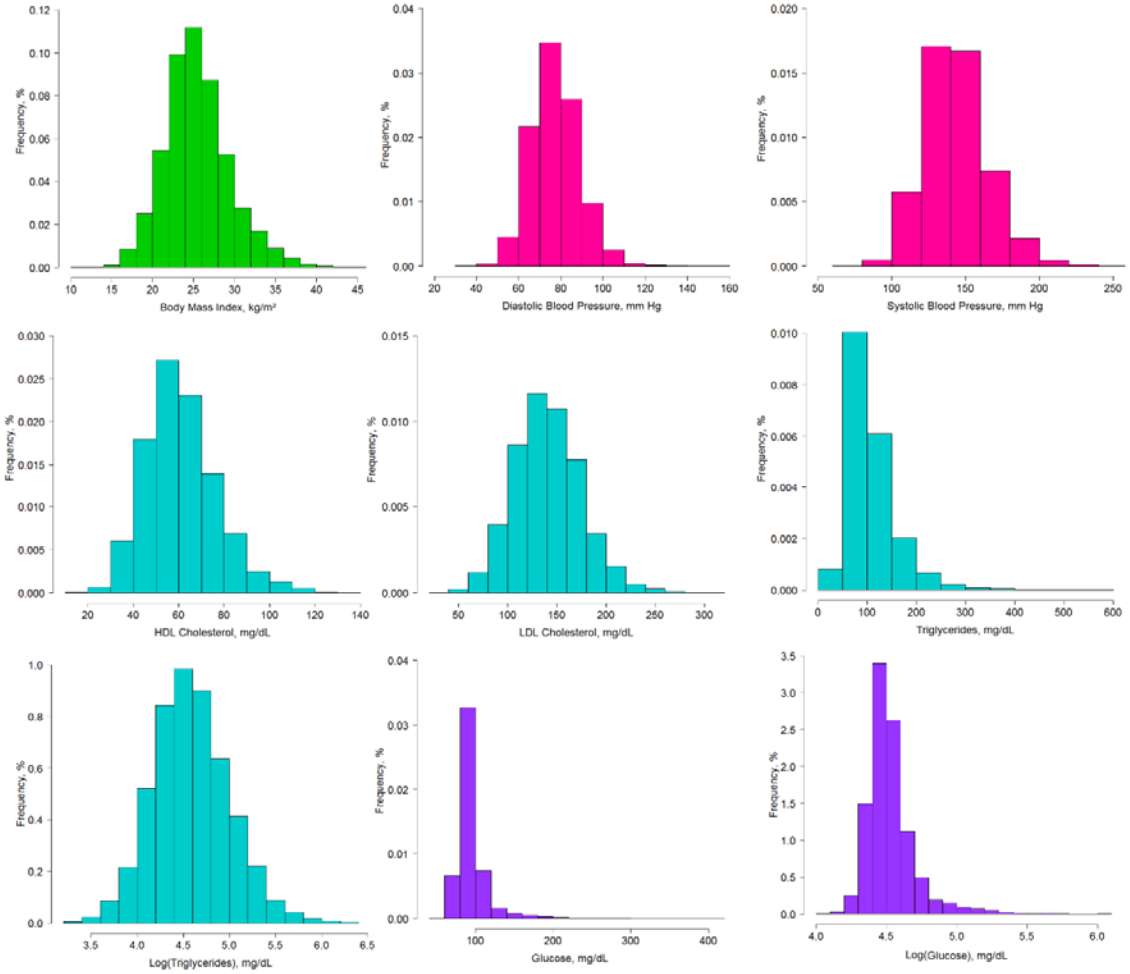
eFigure 5. Trajectories of Glucose Among Incident Cases of Dementia (*n*=650) and Matched Controls (*n*=2794) in Participants Without Diabetes in the 14 Years Preceding the Matching Visit (Dementia Diagnosis of the Case).

Mean trajectories with 95% pointwise confidence intervals were predicted by a latent process linear mixed model in the retrospective time since the matching visit. The models included: a quadratic function of time (t, t^2); case-control status, matching variables (i.e., gender, age, education, and cohort center), and their interactions with the quadratic function of time; and correlated random effects on the intercept, time and time squared. Glucose observations were normalized by I-splines with 3 internal knots. Trajectories were plotted for the most common profile of the study sample: a woman from an average study center, aged 76 years at inclusion and with educational level<high school. Note that the choice of the profile only impacts the level of the trajectories and it does not affect the differences between cases and controls. Number of observations for cases and controls is given at the bottom per 2-year interval. $P_{(t^2,t)xCase}$ corresponds to the *p*-value of the Wald test evaluating the overall difference in change over time across groups (i.e., simultaneous evaluation of time and time squared parameters).



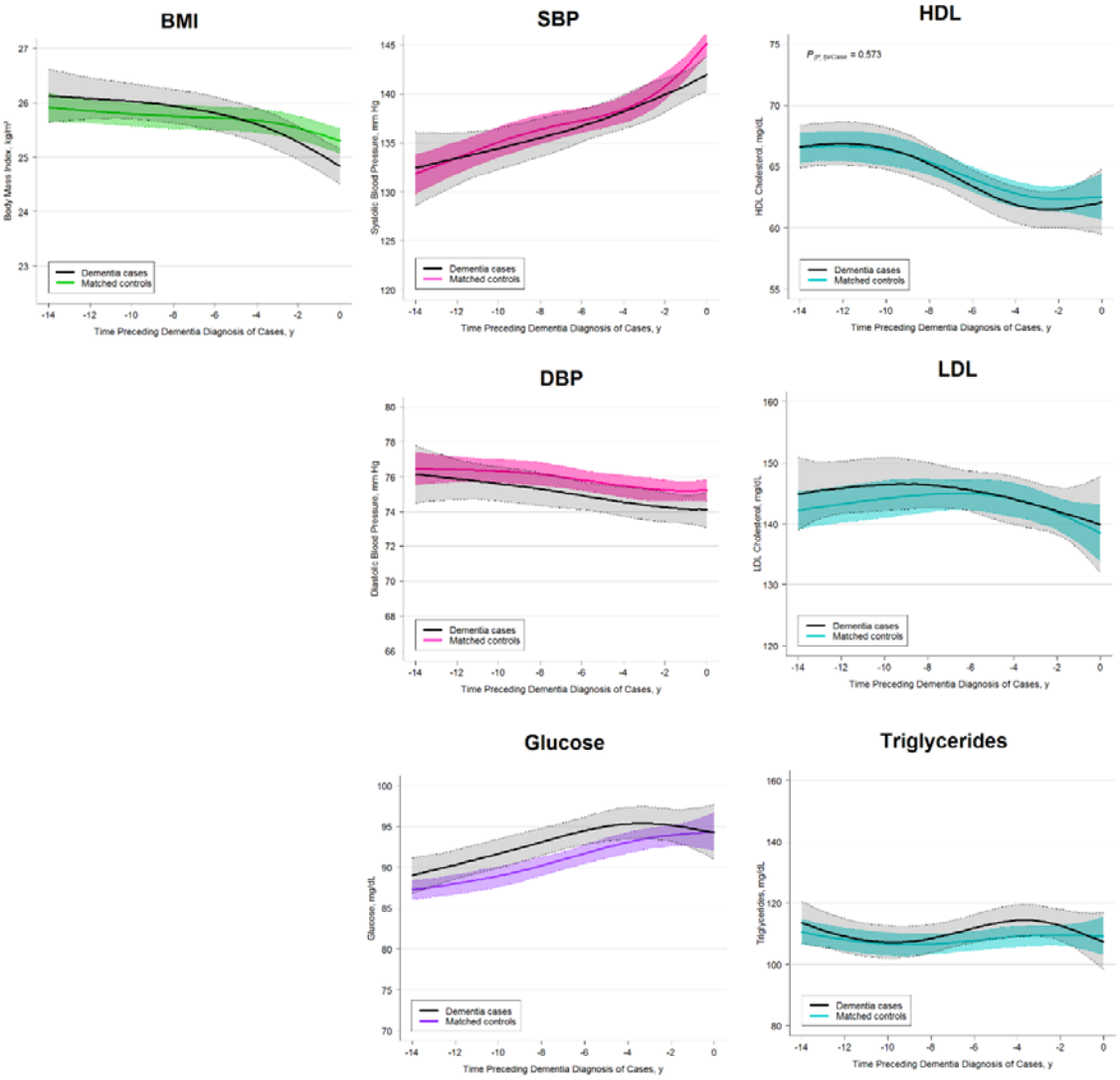
	≤-12	]-12;-10]	]-10;-8]	]-8;-6]	]-6;-4]	]-4;-2]	]-2;0]
Cases	90	75	145	255	107	195	151
Controls	338	306	660	1098	482	934	716

eFigure 6. Distributions of Repeated Measures (Pooled Over Study Visits) of Body Mass Index, Systolic and Diastolic Blood Pressure, HDL and LDL Cholesterol, Triglycerides, Log(Triglycerides), Glucose, and Log(Glucose) in the Nested Cases-Control Sample (*n*=3925), the Three-City Study, France, 1999-2014.



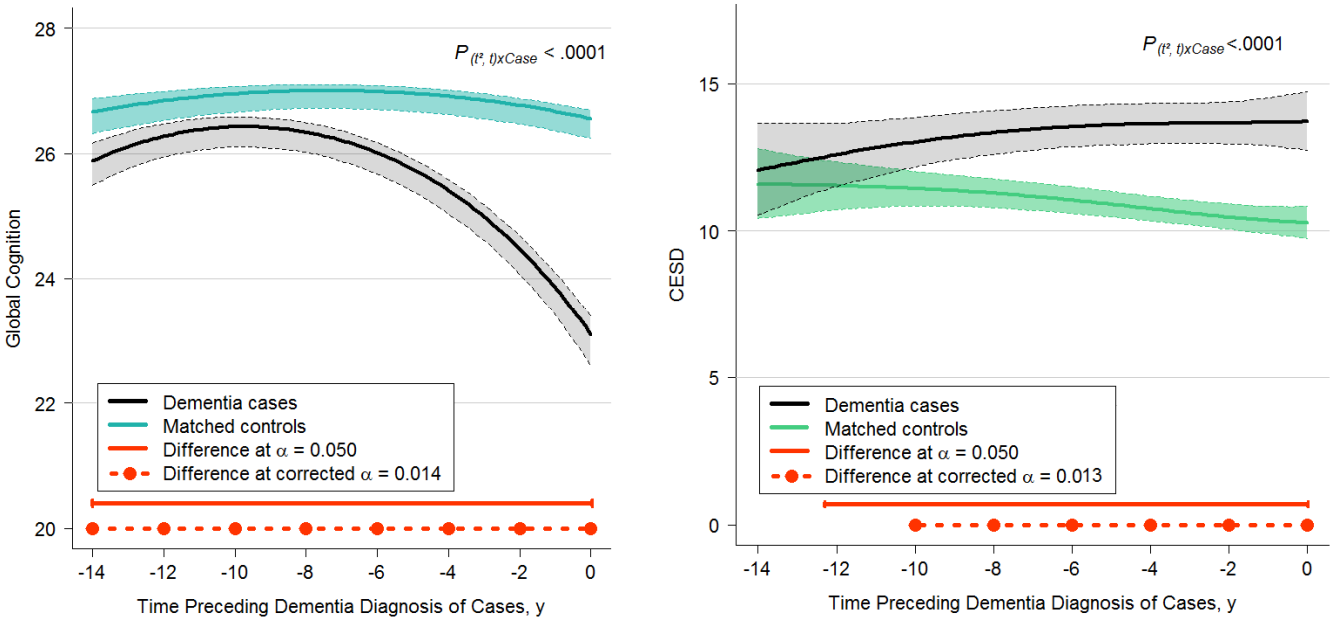
eFigure 7. Trajectories of BMI, SBP, DBP, Glucose, HDL, LDL Cholesterols and Triglycerides Among Incident Cases of D(*n*=785) and Matched Controls (*n*=3140) in the 14 Years Prior to Dementia Diagnosis of Cases When Approximating the Shape of Trajectory by Natural Splines, the Three-City Study, France, 1999-2014.

Mean trajectories with 95% pointwise confidence intervals were predicted by a latent process linear mixed model in the retrospective time since the matching visit. Trajectories were approximated by natural splines (two internal knots) and the models included case-control status, matching variables (i.e., gender, age, education, and cohort center), and their interactions with the function of time as well as correlated random effects on the intercept and function of time. Observations were normalized by I-splines. Trajectories were plotted for the most common profile of the study sample: a woman from an average study center, aged 76 years at inclusion and with educational level<high school. Note that the choice of the profile only impacts the level of the trajectories; it does not affect the differences between cases and controls or the tests significance (not showed here).



eFigure 8. Trajectories of Global Cognition (Left Panel) and Depressive Symptomatology (Right Panel) Among Incident Cases of Dementia (*n*=785) and Matched Controls (*n*=3140) in the 14 Years Preceding the Matching Visit (Dementia Diagnosis of the case).

Mean trajectories with 95% pointwise confidence intervals were predicted by a multivariate latent process linear mixed model in the retrospective time since the matching visit. The models included: a quadratic function of time (t, t^2); case-control status, matching variables (i.e., gender, age, education, and cohort center), and their interactions with the quadratic function of time; and correlated random effects on the intercept, time and time squared. Global cognitive ability (left panel) was defined as the common underlying factor of global cognitive functioning measured by the Mini-Mental State Examination (MMSE), verbal fluency measured by the Isaac's Set Test truncated at 15s and immediate visual memory measured the Benton Visual Retention Test. Each test was transformed by I-splines with 3 internal knots to correct the departure from normality. The scale of global cognitive ability was chosen to be the same as MMSE. Depressive symptomatology (right panel) was measured by the Center for Epidemiological Studies Depression (CESD) scale; CESD was transformed by I-splines with 4 internal knots to correct the departure from normality. Trajectories were plotted for the most common profile of the study sample: a woman from an average study center, aged 76 years at inclusion and with educational level<high school. Note that the choice of the profile only impacts the level of the trajectories and it does not affect the differences between cases and controls or the tests significance. $P_{(t^2,t)xCase}$ corresponds to the *p*-value of the Wald test evaluating the overall difference in change over time across groups (i.e., simultaneous evaluation of time and time squared parameters).



eReferences

1. Proust-Lima C, Dartigues JF, Jacqmin-Gadda H. Misuse of the linear mixed model when evaluating risk factors of cognitive decline. *Am J Epidemiol.* 2011;174(9):1077‐1088.
2. Liquet B, Commenges D. Correction of the P-value after multiple coding of an explanatory variable in logistic regression. *Stat. Med.* 2001;20(19), 2815–2826

