RESEARCH ARTICLE OPEN ACCESS

Association of Physical Activity and Parkinson Disease in Women

Long-term Follow-up of the E3N Cohort Study

Berta Portugal, PhD, Fanny Artaud, PhD, Isabelle Degaey, MD, Emmanuel Roze, MD, PhD, Agnès Fournier, PhD, Gianluca Severi, PhD, Marianne Canonico, PhD, Cécile Proust-Lima, PhD, and Alexis Elbaz, MD, PhD

Correspondence
Dr. Elbaz
alexis.elbaz@inserm.fr

Neurology® 2023;101:e386-e398. doi:10.1212/WNL.0000000000207424

Abstract

Background and Objectives

Previous cohort studies reported that a single measure of physical activity (PA) assessed at baseline was associated with lower Parkinson disease (PD) incidence, but a meta-analysis suggested that this association was restricted to men. Because of the long prodromal phase of the disease, reverse causation could not be excluded as a potential explanation. Our objective was to study the association between time-varying PA and PD in women using lagged analyses to address the potential for reverse causation and to compare PA trajectories in patients before diagnosis and matched controls.

Methods

We used data from the Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale (1990–2018), a cohort study of women affiliated with a national health insurance plan for persons working in education. PA was self-reported in 6 questionnaires over the follow-up. As questions changed across questionnaires, we created a time-varying latent PA (LPA) variable using latent process mixed models. PD was ascertained using a multistep validation process based on medical records or a validated algorithm based on drug claims. We set up a nested case-control study to examine differences in LPA trajectories using multivariable linear mixed models with a retrospective timescale. Cox proportional hazards models with age as the timescale and adjusted for confounders were used to estimate the association between time-varying LPA and PD incidence. Our main analysis used a 10-year lag to account for reverse causation; sensitivity analyses used 5-, 15-, and 20-year lags.

Results

Analyses of trajectories (1,196 cases and 23,879 controls) showed that LPA was significantly lower in cases than in controls throughout the follow-up, including 29 years before diagnosis; the difference between cases and controls started to increase ~10 years before diagnosis (p interaction = 0.003). In our main survival analysis, of 95,354 women free of PD in 2000, 1,074 women developed PD over a mean follow-up of 17.2 years. PD incidence decreased with increasing LPA (p trend = 0.001), with 25% lower incidence in those in the highest quartile compared with the lowest (adjusted hazard ratio 0.75, 95% CI 0.63–0.89). Using longer lags yielded similar conclusions.

Discussion

Higher PA level is associated with lower PD incidence in women, not explained by reverse causation. These results are important for planning interventions for PD prevention.

RELATED ARTICLE

Editorial
Physical Activity and the Risk of Parkinson Disease: Moving in the Right Direction
Page 151

MORE ONLINE

CME Course
NPub.org/cmelist

From the Université Paris-Saclay (B.P., F.A., I.D., A.F., G.S., M.C., A.E.), UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif; Neurology Department (E.R.), Hôpital Pitié-Salpêtrière, AP-HP; Sorbonne Université (E.R.); INSERM U1127 (E.R.), CNRS 7225, Brain Institute, Paris, France; Department of Statistics (G.S.), Computer Science, Applications "G. Parenti" (DISIA), University of Florence, Italy; and Université Bordeaux (C.P.-L.), Inserm, Bordeaux Population Health Research Center, UMR1219, France.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

The Article Processing Charge was funded by the authors.

Glossary

AIC = Akaike information criterion; **BMI** = body mass index; **CIF** = cumulative incidence function; **E3N** = Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale; **HR** = hazard ratio; **LPA** = latent PA; **LPMM** = latent process mixed model; **MET** = metabolic equivalent of task; **PA** = physical activity; **PD** = Parkinson disease.

Parkinson disease (PD) is the fastest growing neurologic disorder in terms of prevalence, disability, and deaths, making prevention an urgent public health need in the absence of curative treatments.[1] Physical activity (PA) represents a modifiable health behavior with major benefits for multiple outcomes.[2] A few cohort studies examined the association between PA and PD incidence, with inconsistent results.[3-9] A meta-analysis of 8 studies estimated that participants with the highest PA level, defined in different ways across studies, had 21% lower PD incidence.[10] This association was statistically significant in men and weaker and not significant in women; however, only 4 studies examined this association in women. Determining whether PA plays a role in women is important for developing appropriate interventions.[11]

Reverse causation refers to situations in which the undiagnosed outcome precedes and leads to changes in exposures instead of the other way around.[12] It represents a key issue for the interpretation of the relation between PA and PD.[13] Nonmotor symptoms (e.g., constipation and sleep disorders) and subtle motor signs (e.g., tremor, balance impairment, and rigidity) can be present several years before PD diagnosis.[14,15] Hence, patients with PD may reduce PA during the prodromal phase as a consequence of prodromal symptoms. To address this issue, some previous cohort studies excluded cases over the first 4–10 years of the follow-up.[3,5-8] Although findings were generally consistent with an inverse association, only 1 study reported a significant association.[7] Therefore, larger cohorts with longer follow-ups are needed to assess whether reverse causation contributes to the inverse association between PA and PD. In addition, none of the previous studies used repeated PA measures, whereas analyses of PA trajectories before PD diagnosis would help understand the temporal relation between PA and PD.

Our aim is to examine the association between time-varying PA measures and PD incidence in women from the Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale (E3N) cohort study over 29 years of follow-up, while addressing the potential for reverse causation. We also used a nested case-control design to compare PA trajectories in patients with PD before diagnosis and matched controls.

Methods

E3N Study

We used data from E3N, an ongoing cohort study of 98,995 French women born between 1925 and 1950 and recruited in 1990, who were affiliated with a French national health insurance plan that covers mostly teachers (Mutuelle Générale de l'Education Nationale).[16]

After providing written informed consent (~20% of invited women), participants completed a self-administered questionnaire on lifestyle and medical history at baseline (Q1–1990). Follow-up questionnaires have been sent every 2–3 years. Eleven waves of data collection are available at the present time (latest in 2014, Q11); the average response rate is of ~80% at each questionnaire. Since January 2004, women were also passively followed through health care databases (including drug and medical consultation claims). Causes of death were available.

Standard Protocol Approvals, Registrations, and Patient Consents

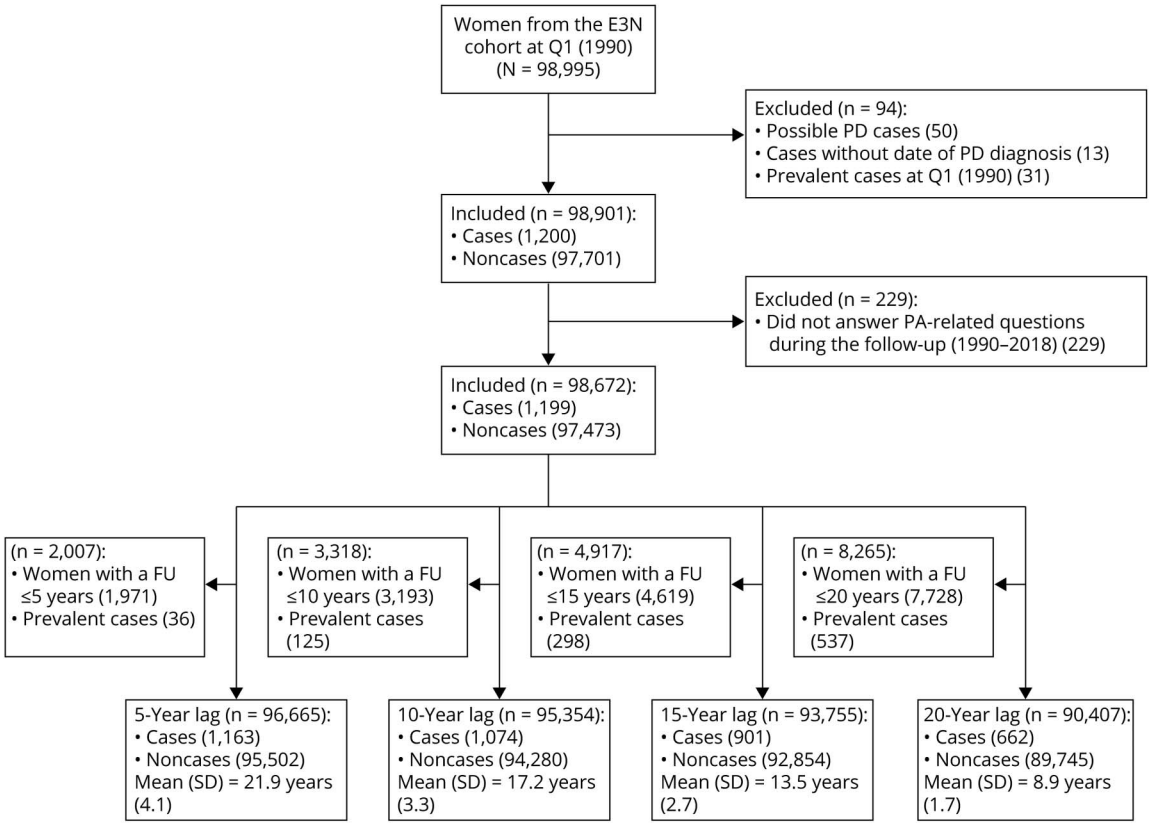
All women signed informed consent in compliance with the rules of the French National Commission for Data Protection and Privacy, from which approval was obtained. The study protocol is registered at ClinicalTrials.gov (NCT03285230).

Ascertainment of Patients With PD

We provide a detailed description of PD ascertainment in E3N elsewhere.[17] Briefly, potential PD cases were identified through self-reported doctor diagnoses of PD in follow-up questionnaires, antiparkinsonian drug claims until December 2018, and death certificates. Potential patients with PD were contacted to confirm the PD diagnosis when possible by mail. For women who confirmed a diagnosis of PD or parkinsonism and for women who could not be contacted (e.g., deceased or contact refusal), we obtained from treating neurologists or general practitioners (GPs) detailed medical documentation (year of PD onset and diagnosis, cardinal motor signs and other neurologic symptoms, use of neuroleptics, treatment, responsiveness to treatment, and diagnosis). Finally, an expert panel adjudicated PD status (definite, probable, possible, or no PD) based on all the medical documentation available.[17] Only patients with definite and probable PD were retained in the analyses. When medical documentation was not available, we used an algorithm based on antiparkinsonian drug claims and medical visits, which was previously validated against a clinical diagnosis (94% sensitivity and 88% specificity).[17] Among women who were considered to have PD, the diagnosis was based on medical records for 62% (62% self-reported PD) and the algorithm for 38% (56% self-reported PD).[17]

Year of PD diagnosis was set as the year of diagnosis available in medical records or, in the decreasing order of priority,

Figure 1 Flowchart for Inclusion Into the Study Using 5-, 10-, 15-, and 20-Year Lags Between Physical Activity and PD Incidence in Survival Analyses



E3N = Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale; FU = follow-up; PD = Parkinson disease; Q = questionnaire.

self-reported year of diagnosis, year of the first use of antiparkinsonian drugs, and year of the first questionnaire where PD was self-reported.

PD incidence rates in E3N are in agreement with those in women from Western Europe (1992–2018) according to the Global Burden of Disease, which supports the validity of our case ascertainment strategy.[17]

Physical Activity

Six questionnaires included PA-related questions (1990–Q1, 1993–Q3, 1997–Q5, 2002–Q7, 2005–Q8, and 2014–Q11; eTable 1, links.lww.com/WNL/C801). PA related to different recreational and household activities was assessed over the follow-up in different ways and units, and 11 types of activities were included in at least 1 questionnaire. Metabolic equivalent of task (MET) values were attributed to each activity based on a compendium and expert opinion (eTable 1). For each activity, METs were multiplied by their frequency and duration to obtain a PA score (MET-hour/week).

The baseline questionnaire (1990–Q1) included 6 closed-ended questions on recreational PA: usual distance walked daily (<500/500–2,000/>2,000 m), the average number of flights of stairs climbed daily (0/1–4/>5), weekly average time spent in light household activities (0/1–4/5–13/≥14 hours), weekly average time spent in heavy household activities (0/1–4/≥5 hours), weekly average time spent in moderate recreational activities (e.g., light gardening or sports of moderate intensity; 0/1–4/5–13/≥14 hours), and weekly average time spent in vigorous recreational activities (e.g., vigorous sports; 0/1–4/≥5 hours).[18]

In subsequent questionnaires (1993–Q3, 1997–Q5, 2002–Q7, and 2014–Q11), PA-related questions were derived from a modified version of the Baecke Questionnaire.[19,20] The questions assessed the duration (hours/week) that participants spent walking (including walking to work, shopping, and leisure time), cycling (including cycling to work, shopping, and leisure time), and engaging in sports during 2 typical weeks over the past year, 1 in summer and 1 in winter. These questionnaires used open-ended questions, allowing women to provide more detailed information on the frequency and duration of each activity than at Q1. In addition, the number of hours practicing each activity was ascertained in the winter and summer, and durations were averaged over the summer and winter. The questionnaire from 2005–Q8 was similar to Q1 but asked open-ended rather than closed-ended questions.

Table 1 Participants' Characteristics at Baseline (1990–Q1)

Baseline characteristics, n (%)	Total (n = 96,665)	LPA at baseline[a] Quartile 1 (24,166 [25.0])	Quartile 2 (24,166 [25.0])	Quartile 3 (24,166 [25.0])	Quartile 4 (24,167 [25.0])
PA, METs-h/wk, mean (SD)[a,b]	45.3 (30.1)	27.0 (15.7)	37.3 (18.7)	46.0 (21.7)	70.8 (38.7)
Age, y, mean (SD)	49.3 (6.6)	49.9 (6.5)	48.5 (6.4)	48.5 (6.4)	50.5 (7.0)
≤49	56,279 (58.2)	11,035 (53.1)	17,129 (63.3)	16,993 (63.5)	11,122 (50.3)
50–59	31,825 (32.9)	7,804 (37.6)	7,986 (29.6)	7,847 (29.3)	8,188 (37.0)
≥60	8,561 (8.9)	1,941 (9.3)	1,910 (7.1)	1,920 (7.2)	2,790 (12.7)
Education					
<High school diploma	12,815 (13.8)	3,703 (16.0)	2,970 (12.8)	2,807 (12.1)	3,335 (14.4)
≥High school diploma	79,814 (86.2)	19,373 (84.0)	20,210 (87.2)	20,362 (87.9)	19,869 (85.6)
Missing	4,036	1,090	986	997	963
Smoking					
Never	52,014 (53.9)	12,822 (53.1)	12,580 (52.2)	12,850 (53.2)	13,762 (57.0)
Ex	29,911 (31.0)	7,298 (30.2)	7,659 (31.8)	7,707 (32.0)	7,247 (30.0)
Current	14,564 (15.1)	4,012 (16.7)	3,852 (16.0)	3,565 (14.8)	3,135 (13.0)
Missing	176	34	75	44	23
Age at menarche, y, mean (SD)	12.8 (1.4)	12.8 (1.4)	12.8 (1.4)	12.8 (1.4)	12.9 (1.4)
≤11	19,761 (20.9)	5,030 (21.4)	4,963 (21.0)	4,942 (21.0)	4,826 (20.4)
12–13	47,656 (50.5)	11,942 (50.7)	12,118 (51.2)	11,892 (50.4)	11,704 (49.5)
≥14	27,002 (28.6)	6,564 (27.9)	6,573 (27.8)	6,734 (28.6)	7,131 (30.1)
Missing	2,246	630	512	598	506
Menopausal status					
Premenopausal	53,531 (57.5)	12,620 (53.7)	14,166 (61.7)	14,568 (62.9)	12,177 (51.6)
Natural menopause	32,520 (34.9)	8,841 (37.6)	7,119 (31.1)	7,058 (30.5)	9,502 (40.2)
Artificial menopause	6,410 (6.9)	1,810 (7.7)	1,436 (6.3)	1,382 (6.0)	1,782 (7.5)
Unknown type of menopause	694 (0.7)	217 (1.0)	200 (0.9)	129 (0.6)	148 (0.7)
Missing	3,510	678	1,245	1,029	558
Parity					
Nulliparous	11,586 (12.1)	3,705 (15.5)	2,865 (11.9)	2,549 (10.6)	2,467 (10.3)
One child	15,571 (16.2)	4,405 (18.4)	4,067 (17.0)	3,701 (15.4)	3,398 (14.2)
Two children	40,724 (42.5)	9,632 (40.3)	10,617 (44.2)	10,722 (44.7)	9,753 (40.6)
≥3 children	28,011 (29.2)	6,168 (25.8)	6,449 (26.9)	7,030 (29.3)	8,364 (34.9)
Missing	773	256	168	164	185
Place of residence					
Rural	13,089 (14.8)	2,138 (9.7)	2,844 (12.9)	3,525 (16.0)	4,582 (20.7)
Urban	75,327 (85.2)	19,939 (90.3)	19,233 (87.1)	18,570 (84.0)	17,585 (79.3)
Missing	8,249	2,089	2,089	2,071	2,000
BMI, kg/m^2, mean (SD)	22.6 (3.2)	22.9 (3.5)	22.5 (3.1)	22.4 (3.0)	22.6 (3.0)
<18.5	3,995 (4.2)	1,100 (4.7)	1,006 (4.3)	1,011 (4.3)	878 (3.7)

Continued

Table 1 Participants' Characteristics at Baseline (1990–Q1) *(continued)*

Baseline characteristics, n (%)	Total (n = 96,665)	LPA at baseline[a] Quartile 1 (24,166 [25.0])	Quartile 2 (24,166 [25.0])	Quartile 3 (24,166 [25.0])	Quartile 4 (24,167 [25.0])
18.5–24.9	73,801 (78.3)	17,537 (74.7)	18,648 (79.0)	18,832 (79.7)	18,784 (79.4)
25–29.9	13,658 (14.5)	3,767 (16.1)	3,247 (13.8)	3,221 (13.6)	3,423 (14.5)
≥30.0	2,860 (3.0)	1,046 (4.5)	693 (2.9)	555 (2.4)	566 (2.4)
Missing	2,351	716	572	547	516
Hypercholesterolemia[c]	36,574 (37.8)	10,763 (38.6)	7,694 (36.4)	7,727 (36.6)	10,390 (39.2)
High blood pressure[c]	28,843 (29.8)	8,966 (32.1)	6,113 (28.9)	6,085 (28.8)	7,679 (29.0)
Diabetes[c]	2,651 (2.7)	1,010 (3.6)	551 (2.6)	497 (2.4)	593 (2.2)
Cardiovascular disease[c]	2,743 (2.8)	1,011 (3.6)	573 (2.7)	527 (2.5)	632 (2.4)

BMI = body mass index; LPA = latent PA; MET = metabolic equivalent of task; PA = physical activity.
Baseline characteristics are shown for participants available for survival analyses based on a 5-year lag.
[a] PA was not available at baseline for 1,169 women (1.2%); therefore, we used the first LPA value available over the follow-up for these women (95% at Q3–1993 or Q5–1997).
[b] Total PA assessed at the baseline questionnaire (1990–Q1).
[c] Assessed at the end of the follow-up.

There were 5 questions in common for Q3, Q5, and Q7; Q5 and Q7 included the same questions; Q8 had 4 matching questions with Q1; Q11 had 4 identical questions with Q5 and Q7 (eTable 1, links.lww.com/WNL/C801).

Covariates

Participants' characteristics were collected through follow-up questionnaires. Birth date, education level (</≥high school), region of residence (French commune categorized into rural/urban according to the French Institut national de la statistique et des études économiques),[21] age at menarche (≤11/12–13/≥14 years), and parity (nulliparous/1 child/2 children/≥3 children) were assessed at baseline (Q1). Menopausal status (premenopausal/natural/artificial/unknown type of menopause) and weight were determined at all waves; height was determined at Q1, Q4 (1994), and Q6–Q11 (2000–2014) and standardized by using the most frequent value. Body mass index (BMI) was computed as weight divided by height squared (kg/m^2) and categorized into 4 categories (underweight: <18.5/normal weight: 18.5–24.9/overweight: 25.0–29.9/obese: ≥30.0 kg/m^2). Information on smoking status (never/ex/current) was collected from Q1 to Q8. Diet, including adherence to the Mediterranean diet, total intake of caffeine (in milligrams), and daily intake of lactose (in grams), was assessed at Q3 (1993) using a validated dietary questionnaire and categorized in quartiles.[22] We used self-reports in E3N questionnaires throughout the follow-up to ascertain comorbidities, including history of hypercholesterolemia, cardiovascular diseases (ischemic heart disease/stroke), and high blood pressure. Type 2 diabetes was ascertained using self-reports in questionnaires before 2004 and antidiabetic drug claims thereafter.

Nested Case-Control Study

To compare PA trajectories in PD cases and matched controls, we set up a nested case-control study. Each incident PD case was individually matched to 20 controls using incidence density sampling.[23] To be selected as controls, women had to be alive and at risk of PD at the date of diagnosis of the matched case (index date, T_0) and have the same exact age (rounded to the nearest integer without decimals) at T_0.

Statistical Analysis

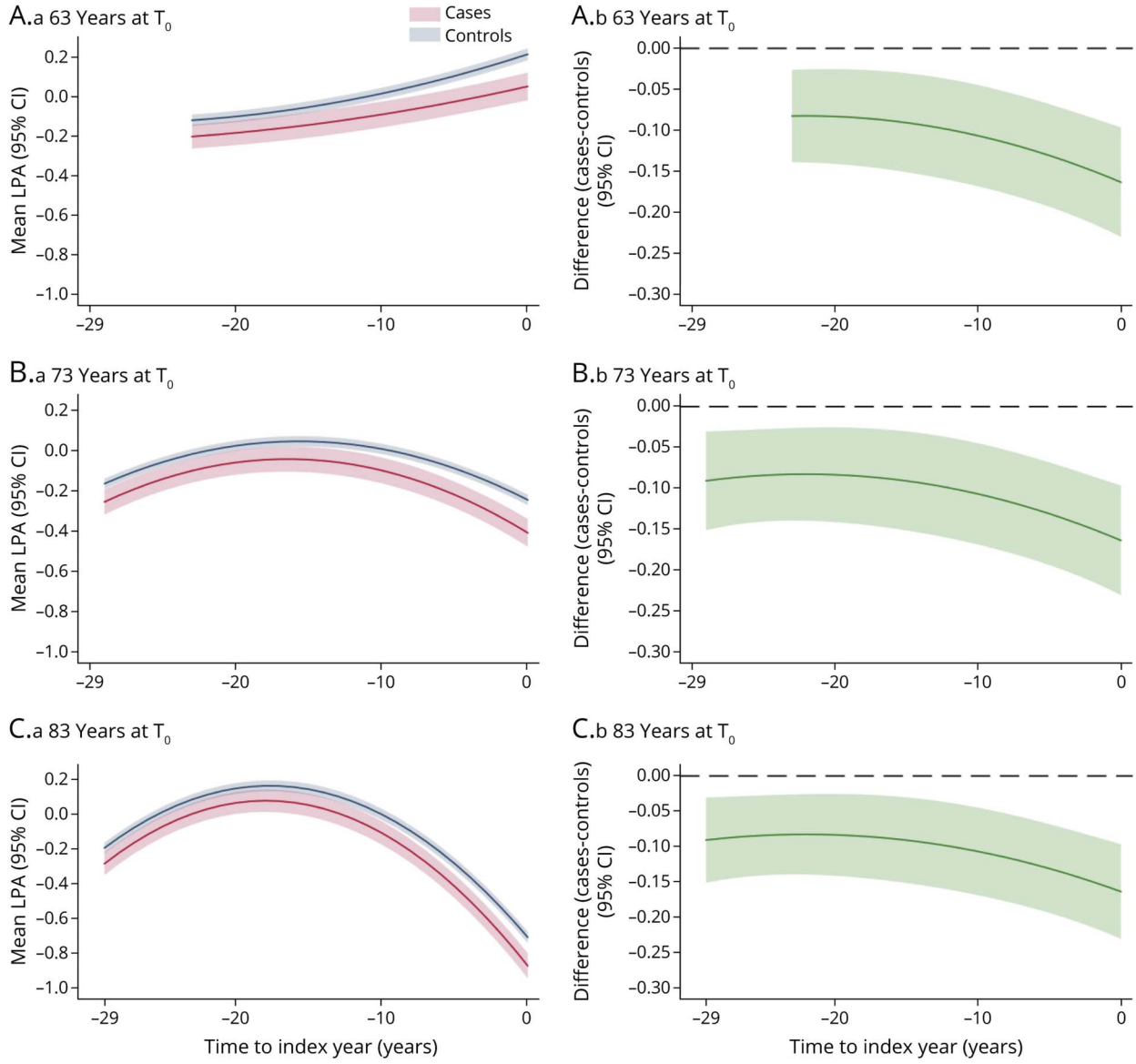
Statistical analyses were conducted using SAS 9.4 (SAS Institute Inc., Cary, NC), the lcmm R package (R Foundation for Statistical Computing, Vienna, Austria),[24] and Stata 15.1 (StataCorp, College Station, TX).

Latent PA Process

When measurement tools of a risk factor change over the follow-up, latent process mixed models (LPMMs) allow to include observations from different measurement tools and to define trajectories of a latent process, provided that the tools measure the same quantity. This latent process represents the unmeasured common factor underlying the observations obtained through different tools.[25,26] We aimed at predicting a latent PA (LPA) process for all the participants from the cohort with at least 1 PA assessment (N = 98,766) using LPMMs.

To link repeated individual PA measures (in MET-hour/week) obtained from different questionnaires to their common underlying latent process and correct the departure of each activity from a normal distribution, we first selected the best activity-specific parameterized link function (among quadratic I-splines link functions with 2, 3, 4, and 5 knots)

Figure 2 Trajectories of Latent Physical Activity in Cases With PD and Matched Controls Up to 29 Years Before the Index Date



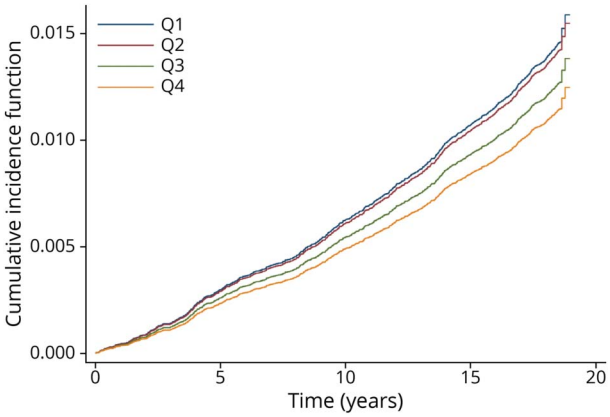
Figures A.a, B.a, and C.a show the trajectories (95% CI) of mean LPA in 1,196 PD cases and 23,879 matched controls based on a linear mixed model with a quadratic function of retrospective time; the model's coefficients are shown in eTable 5 (links.lww.com/WNL/C801). Figures A.b, B.b, and C.b show the differences (95% CI) between the mean trajectories of latent physical activity in PD cases and controls. Differences whose CI do not include 0 (horizontal dashed line) are statistically significant. We used a retrospective timescale, with T_0 (time = 0) representing the year of PD diagnosis in cases and the index date in controls. The model was adjusted for PD status, age at T_0, and 2-way interactions of time with PD status and age at T_0. It was further adjusted for baseline parity, place of residence, age at menarche, and time-varying smoking and menopausal status. Given the significant interaction between age at T_0 and time, trajectories were plotted for 3 different ages at T_0 (63, 73, and 83 years) and the most common profile of E3N participants (never smokers, age at menarche at 12–13 years, natural menopause, 2 children, and living in urban areas). E3N = Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale; LPA = latent physical activity; PD = Parkinson disease.

according to the Akaike information criterion (AIC) in activity-specific LPMMs. We then included the 11 activities (with the shapes of activity-specific link functions previously determined) in the same LPMM for multivariate longitudinal outcomes to model the LPA trajectory according to age. Age was centered at its mean at Q1 (49 years), and change over age was approximated by natural cubic splines with 2 internal knots (lower AIC). The within-participant correlation was captured by independent random intercept and slopes on the 3 functions of age.

We finally obtained patient-specific LPA predictions for each patient at each visit based on the best fitting LPMM, which were used in subsequent trajectories and survival analyses.

As it relies on maximum likelihood estimation, the LPMM allows to handle missing values of PA under the assumption of missingness at random, that is, missing values at a given time point can be predicted by the observed values of PA and other covariates (age).[27]

Figure 3 Cumulative Incidence Function of PD Over the Follow-up (2000–2018) According to LPA Assessed at Q1 (1990)



The cumulative incidence function of PD as a function of the quartiles of LPA assessed at Q1 (1990) is predicted by a Fine-Gray regression model for competing risks, with time since the beginning of the follow-up as the timescale and a 10-year lag (follow-up, 2000–2018). The model was adjusted for age (restricted cubic spline with 3 knots), parity, place of residence, age at menarche, smoking, and menopausal status (all assessed at the beginning of the follow-up). LPA = latent physical activity; PD = Parkinson disease.

Trajectories of LPA in Cases and Controls

LPA trajectories from the index time T_0 to the beginning of the study (retrospective timescale) were examined in cases and controls over 29 years of follow-up using a linear mixed model. The LPA trajectory was modeled as a function of retrospective time and time squared and was adjusted for PD status, age at T_0, and 2-way interactions of time with PD status and age; it was further adjusted for confounders associated with PA and PD, including parity,[28] age at menarche,[28] rural residence,[29] and time-varying smoking[29] and menopausal status.[28] The within-participant correlation was captured by correlated random intercept and slopes on time and time squared.

Survival Analysis

We used Cox proportional hazard models for time-varying variables with age as the timescale to estimate hazard ratios (HRs), 95% CIs, and 2-tailed p values ($\alpha = 0.05$).

To address the potential for reverse causation, we included a lag of increasing duration (5, 10, 15, and 20 years) between time-varying variables (including LPA) and PD incidence. Given results of the analyses of trajectories described above, we used a 10-year lag for our main analyses: LPA was lagged by 10 years, and we started the follow-up in 2000 (i.e., 10 years after the baseline assessment), so that participants who developed PD before 2000 (prevalent cases) were excluded. Women were followed from 2000 until PD diagnosis or the end of follow-up (maximum of the date of the last questionnaire and last drug reimbursement). The same approach was used for other lags.

LPA was included as time-varying quartiles in the models, and linear trends in HRs were tested through ordinal variables defined by the median of each quartile. We also used restricted cubic splines to test for departures from linearity for continuous LPA. Analyses were adjusted for confounders associated with PA and PD, including baseline parity,[28] age at menarche,[28] rural residence,[29] and time-varying smoking[29] and menopausal status.[28] Missing values were coded as specific categories to retain the same number of participants in all analyses.

We used the Fine-Gray subdistribution hazard model to estimate the cumulative incidence function (CIF) of PD over time in the presence of the competing risk of death.[30] For this analysis, we estimated the CIF of PD as a function of the quartiles of LPA at Q1, with time since the beginning of the follow-up as the timescale and a 10-year lag. Analyses are adjusted for the same covariates described above and for age at the beginning of the follow-up (restricted cubic splines with 3 knots).

We performed several sets of sensitivity analyses:

1. We constructed an alternative LPA variable by excluding the baseline PA questions (1990–Q1) that were less precise than subsequent measures.
2. To examine whether our findings may have been confounded by diet, we performed analyses based on PA assessed at Q5 (1997) and adjusted for dietary exposures associated with PD risk that were assessed earlier during the follow-up (1993–Q3; Mediterranean diet, caffeine, and dairy intake).[31] Analyses with a lag longer than 10 years were not possible because of an insufficient number of cases.
3. We performed analyses stratified by median age (72.8 years) to examine whether associations were similar for patients with PD who developed PD before or after that age.
4. To examine the influence of the case definition on our findings, we excluded patients with PD predicted by the algorithm and performed analyses restricted to cases validated based on medical documentation (definite or probable; definite only). Because we were concerned that the reduced sample size would lead to insufficient statistical power, we performed additional analyses in which, in addition to validated cases, we also retained cases predicted based on a more specific version of the algorithm (90% sensitivity; 90% specificity).[17] For these analyses, we compared PD incidence in the fourth quartile with that in the other 3 quartiles combined.
5. Higher PA is associated with a lower risk of several comorbidities, including obesity, high blood pressure, dyslipidemia, and cardiovascular disease.[2] Hence, comorbidities may be potential mediators of the association between PA and PD. We ran sensitivity analyses adjusted for time-varying BMI, high blood pressure,

Table 2 Association of Time-Varying Physical Activity With Parkinson Disease Incidence

Latent physical activity	Cases (n)	IR	Age-adjusted HR (95% CI)	*p* Value	Multivariable HR (95% CI)[a]	*p* Value
5-y lag (FU 1995–2018; N = 96,665)	1,163					
Quartile 1	316	0.60	1.00 (ref)	—	1.00 (ref)	—
Quartile 2	304	0.60	1.01 (0.86–1.18)	0.95	0.99 (0.85–1.16)	0.94
Quartile 3	289	0.56	0.96 (0.81–1.12)	0.58	0.94 (0.80–1.11)	0.49
Quartile 4	254	0.45	0.76 (0.64–0.89)	<0.001	0.75 (0.63–0.88)	<0.001
			p-linear trend	<0.001	*p*-linear trend	<0.001
10-y lag (FU 2000–2018; N = 95,354)	1,074					
Quartile 1	286	0.73	1.00 (ref)	—	1.00 (ref)	—
Quartile 2	274	0.69	0.95 (0.80–1.12)	0.57	0.94 (0.80–1.11)	0.50
Quartile 3	268	0.67	0.93 (0.78–1.10)	0.37	0.92 (0.77–1.08)	0.30
Quartile 4	246	0.55	0.76 (0.64–0.90)	0.002	0.75 (0.63–0.89)	0.001
			p-linear trend	0.002	*p*-linear trend	0.001
15-y lag (FU 2005–2018; N = 93,755)	901					
Quartile 1	237	0.80	1.00 (ref)	—	1.00 (ref)	—
Quartile 2	229	0.74	0.93 (0.78–1.12)	0.44	0.92 (0.77–1.10)	0.37
Quartile 3	222	0.72	0.90 (0.75–1.08)	0.27	0.89 (0.74–1.07)	0.20
Quartile 4	213	0.63	0.79 (0.66–0.95)	0.01	0.78 (0.64–0.94)	0.008
			p-linear trend	0.01	*p*-linear trend	0.008
20-y lag (FU 2010–2018; N = 90,407)	662					
Quartile 1	166	0.88	1.00 (ref)	—	1.00 (ref)	—
Quartile 2	178	0.85	0.97 (0.78–1.20)	0.78	0.96 (0.78–1.19)	0.71
Quartile 3	167	0.81	0.92 (0.74–1.14)	0.44	0.90 (0.73–1.12)	0.36
Quartile 4	151	0.73	0.83 (0.67–1.04)	0.11	0.82 (0.65–1.02)	0.08
			p-linear trend	0.09	*p*-linear trend	0.06

FU = follow-up; HR = hazard ratio; IR = age-standardized incidence rate of Parkinson disease per 1,000 person-years.
HRs and 95% CIs calculated using Cox proportional hazards models for time-varying variables with age as the timescale.
[a] Models are adjusted for baseline place of residence (rural/urban), age at menarche (≤11/12–13/≥14 years), parity (nulliparous/1 child/2 children/≥3 children), and time-varying smoking (never/ex/current) and menopausal status (premenopausal/natural menopause/artificial menopause/unknown type of menopause).

hypercholesterolemia, diabetes, and cardiovascular disease to examine their contribution to the association between PA and PD.

Data Availability

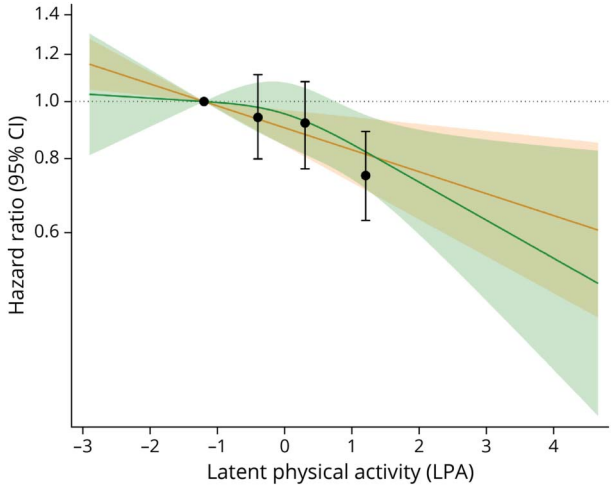
Data on E3N cohort participants are available to bona fide researchers for all types of health-related research, which is in the public interest. Data are made available under managed access owing to governance constraints and the need to protect the privacy of study participants. Raw data requests should be submitted through the E3N website (e3n.fr) or sent to contact@e3n.fr; requests will be reviewed by the E3N Access Committee. Further information is available at e3n.fr/node/78.

Results

Figure 1 shows a flowchart for inclusion of participants into the study. We excluded 50 possible PD cases, 13 cases without a diagnosis date, 31 prevalent cases at Q1, and 229 women who did not answer PA questions at any questionnaire; among the remaining E3N participants, 90% had 3 measures or more of PA during the follow-up, and of these, 80% had 5 or 6 measures available.

The number of participants available for survival analyses decreased with increasing lags. For instance, for our main survival analysis (10-year lag), we further excluded 3,193 women whose follow-up ended before 2000 and 125 patients

Figure 4 HRs of PD in Relation to LPA (10-Year Lag)



HRs and 95% CIs calculated using Cox proportional hazards models for time-varying variables with age as the timescale and adjusted for baseline place of residence (rural/urban), age at menarche (≤11/12–13/≥14 years), parity (nulliparous/1 child/2 children/≥3 children), and time-varying smoking (never/ex/current) and menopausal status (premenopausal/natural menopause/artificial menopause/unknown type of menopause). The dots correspond to HRs for quartiles of LPA compared with the reference quartile (Table 2) together with their 95% CIs (vertical bars); HRs are plotted at the median of each quartile. The orange solid line represents the HR of PD for continuous LPA modeled as a linear variable, and the shaded area corresponds to the 95% CI. The HR of PD decreased linearly with an increasing level of LPA. The green solid line represents the HR of PD for continuous LPA modeled with restricted cubic splines, and the shaded area corresponds to 95% CI; 3 knots provided the best fit (lower AIC values). There was no significant departure from linearity (p = 0.25). AIC = Akaike information criterion; HR = hazard ratio; LPA = latent physical activity; PD = Parkinson disease.

with prevalent PD, leaving 95,354 women followed for 19 years (mean = 17.2, SD = 3.3) of whom 1,074 developed PD. Compared with previous studies, our study included the largest number of patients with PD and had the longest follow-up (eTable 2, links.lww.com/WNL/C801).

Table 1 describes baseline (1990–Q1) participants' characteristics. The mean age was 49.3 years (SD = 6.6), and the mean PA level was 45.3 (SD = 30.1) METs-hour/week. Women with later age at menarche, with ≥3 children, and who lived in rural areas had higher LPA levels than their counterparts. Women with incident PD were older, less frequently smokers, more frequently postmenopausal, and had later age at menarche and more often ≥3 children than those who remained PD free (eTable 3, links.lww.com/WNL/C801).

Trajectories of PA Preceding PD

Among 25,200 women (1,200 cases and 24,000 age-matched controls) included in the nested case-control study, we excluded 45 women (4 PD cases) for whom PA was missing at all visits before T_0, and 80 controls matched to these 4 PD cases, leading to a final sample of 25,075 women (1,196 PD cases and 23,879 controls); 1,156 cases were matched to 20 controls, 39 cases to 19 controls, and 1 case to 18 controls.

eTable 4 (links.lww.com/WNL/C801) describes the characteristics of cases and controls at the index date (T_0). The mean (SD) age of cases and controls was 71.9 years (SD = 7.8). Compared with controls, PD cases were less often smokers and obese, had more frequently menarche ≤11 years or ≥14 years and artificial menopause, and had more children.

Figure 2 shows LPA trajectories in cases and controls for the most common profile of covariates and 3 ages at T_0; eTable 5 (links.lww.com/WNL/C801) presents estimates from the corresponding model.

After an initial increase of LPA in cases and controls, it decreased with a steeper decline in cases than in controls in the 10 years before T_0, because of a significant interaction between PD status and time (p interaction = 0.003). Cases had a significantly lower PA level at T_0 than controls (difference −0.164, 95% CI −0.230 to −0.097); this difference decreased over the 10 years before T_0, but it remained significant throughout the follow-up, and LPA was significantly lower in cases than in controls 29 years before T_0; hence, the difference in LPA between cases and controls was larger at the index date than at the beginning of the study. Based on these findings, we used a 10-year lag for our main survival analysis.

In addition, the LPA level decreased with age and increased with the number of children, and it was lower in smokers than nonsmokers, in urban compared with rural regions, and in postmenopausal compared with premenopausal women.

PA and PD Incidence

Figure 3 shows the CIF of PD over the follow-up, while taking the competing risk of death into account; the CIF was reduced by 22% (95% CI 7%–34%, p = 0.007) in women in the highest quartile of LPA compared with those in the lowest quartile.

Table 2 shows the association between time-varying LPA and PD incidence. In our main analysis (10-year lag), the hazard of PD decreased with increasing PA (p trend = 0.001), with 25% lower incidence in those in the highest quartile compared with the lowest (HR 0.75, 95% CI 0.63–0.89). The hazard of PD decreased linearly with an increasing PA level; analyses based on splines showed no departure from linearity (p = 0.25; Figure 4). Using longer lags yielded similar associations (Table 2); the inverse association was borderline significant for the 20-year lag based on a smaller number of PD cases (p trend = 0.06).

Sensitivity analyses adjusted for Mediterranean diet and caffeine and dairy intake (Table 3) or excluding baseline PA assessments (eTable 6, links.lww.com/WNL/C801) yielded similar results. Analyses stratified by median age showed similar associations in both age groups (eTable 7). In analyses using alternative PD definitions, associations were similar to those from our main analyses (eTable 8). In analysis adjusted for comorbidities that represent potential mediators (BMI,

Table 3 Association of PA (Assessed at Q5 in 1997) With PD Incidence: Analyses Adjusted for Diet (Assessed at Q3 in 1993)

PA (METs-hour/week) in 1997 (Q5)	Cases (n)	IR	Age-adjusted HR (95% CI)	*p* Value	Model 1 HR (95% CI)[a]	Model 1 *p* Value	Model 2 HR (95% CI)[b]	Model 2 *p* Value
5-y lag (FU 2002–2018; N = 81,777)	881							
Quartile 1	235	0.79	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Quartile 2	207	0.68	0.85 (0.70–1.02)	0.09	0.85 (0.70–1.02)	0.08	0.84 (0.70–1.01)	0.07
Quartile 3	225	0.69	0.86 (0.72–1.03)	0.10	0.85 (0.70–1.02)	0.07	0.84 (0.70–1.01)	0.06
Quartile 4	214	0.61	0.77 (0.64–0.92)	0.005	0.75 (0.62–0.90)	0.002	0.74 (0.62–0.90)	0.002
			p-linear trend	0.01	*p*-linear trend	0.005	*p*-linear trend	0.004
10-y lag (FU 2007–2018; N = 79,836)	685							
Quartile 1	176	0.87	1.00 (ref)	—	1.00 (ref)	—	1.00 (ref)	—
Quartile 2	160	0.77	0.87 (0.71–1.08)	0.22	0.87 (0.70–1.08)	0.20	0.86 (0.70–1.07)	0.18
Quartile 3	182	0.83	0.94 (0.77–1.16)	0.57	0.93 (0.75–1.14)	0.47	0.92 (0.74–1.13)	0.41
Quartile 4	167	0.72	0.82 (0.66–1.01)	0.07	0.80 (0.65–0.99)	0.04	0.79 (0.64–0.98)	0.04
			p-linear trend	0.11	*p*-linear trend	0.07	*p*-linear trend	0.06

FU = follow-up; HR = hazard ratio; IR = age-standardized incidence rate of PD per 1,000 person-years; MET = metabolic equivalent of task; PA = physical activity; PD = Parkinson disease.
HRs and 95% CIs calculated using Cox proportional hazards models with age as the timescale.
A total of 86,824 women participated to the Q5 wave of data collection, of whom 83,574 had information on PA; we further excluded 46 participants who were possible PD cases, 7 participants who had no date of PD diagnosis, and 83 prevalent PD cases at Q5 (n = 136). Of the remaining 83,438 women, 94 developed PD, and 1,567 were censored free of PD within the first 5 years of follow-up, leaving 81,777 (881 PD) women for the analyses based on a 5-year lag; 290 developed PD, and 3,312 were censored free of PD within the first 10 years of follow-up, leaving 79,836 (685 PD) women for the analyses based on a 10-year lag. Analyses with lags longer than 10 years were not possible because of an insufficient number of cases.
[a] Multivariable model 1 is adjusted for baseline place of residence (rural/urban), age at menarche (≤11/12–13/≥14 years), parity (nulliparous/1 child/2 children/≥3 children), smoking (never/ex/current), and menopausal status (premenopausal/natural menopause/artificial menopause/unknown type of menopause).
[b] Multivariable model 2 is further adjusted for caffeine intake (milligrams, quartiles), adherence to Mediterranean diet (quartiles), and dairy intake (grams, quartiles) assessed in 1993 (Q3).

high blood pressure, hypercholesterolemia, diabetes, and cardiovascular diseases), the association between PA and PD remained unchanged after adjusting for these covariates that explained a small proportion of the association between PA and PD (5.5%) (eTable 9).

Discussion

In this cohort study of ~100,000 French women with 6 PA measures over 29 years of follow-up, increasing PA was associated with reduced PD incidence, while taking into account the potential for reverse causation. Analyses of PA trajectories showed that PD cases had lower PA levels than controls 29 years before the index date, and that case-control differences in the PA level increased ~10 years before diagnosis in agreement with the hypothesis that prodromal PD leads to a reduction in PA.

Previous cohort studies on the relation between PA and PD yielded inconsistent findings (eTable 2, links.lww.com/WNL/C801). One meta-analysis (8 studies, 2,192 patients with PD) with a median follow-up period of 12 years (range 6.1–22.0) showed that PA was associated with lower PD incidence in analyses of men and women combined.[10] This association was statistically significant in men and weaker and not significant in women; however, only 4 studies examined women alone (604 patients with PD)[3,5-7] and did not adjust for characteristics of reproductive life associated with PD. In addition, no previous study used repeated PA measures to take into account changes in PA over the follow-up. Five studies performed sensitivity analysis using a lag between PA assessment and PD incidence (4-year lag, n = 3; 8-year lag, n = 1; and 10-year lag, n = 1), and only 1 with a 4-year lag showed a significant inverse association, overall and in sex-stratified analyses (198 women with PD).[7] One study showed that moderate-to-vigorous PA at ages 35–39 years was associated with lower PD risk in men and women,[7] whereas another found a significant inverse association between the higher PA level in early adulthood and PD in men but not in women.[3]

The main difference between previous studies and ours is that we identified a considerably larger number of women with incident PD over a longer follow-up, allowing us to perform analyses with longer lags while retaining a sufficient number of patients with PD. PA trajectories in controls were consistent with decreasing PA levels in the elderly.[32] This decline was steeper in cases, with a case-control difference that started to increase ~10 years before PD diagnosis, emphasizing the

importance of performing lagged analyses to estimate associations between PA and PD not biased by reverse causation. Analyses with lags ≥10 years confirmed an inverse association between PA and PD incidence that was significant for a 15-year lag and borderline significant for a 20-year lag, likely because of the smaller number of cases. Therefore, our findings suggest that reverse causation is unlikely to explain the inverse association between PA and PD.

Converging evidence from studies in patients with PD, including observational studies[33] and randomized controlled trials,[34,35] suggests that PA improves PD motor and nonmotor symptoms. In patients with PD, aerobic exercise stabilizes disease progression in the corticostriatal sensorimotor network and enhances cognitive performance.[36] Our results extend these findings and suggest that PA may help prevent or delay PD onset, possibly by slowing PD pathologic processes, in agreement with 1 study that showed a reduced prevalence of PD prodromal symptoms in individuals more physically active in midlife.[37] These findings have triggered interest in elucidating the mechanisms that explain beneficial effects of PA for PD. Exercise induces the recovery of motor function and neuroprotection of dopaminergic neurons in animal models of PD, regulates dopaminergic and glutamatergic transmission, mobilizes neurotrophic factors (brain-derived neurotrophic factor/glial cell–derived neurotrophic factor), modulates neuroinflammatory mechanisms, attenuates mitochondrial dysfunction and oxidative stress, and enhances brain plasticity.[38-41] In humans, PA has been associated with brain structural and functional changes until late adulthood.[42] In postmenopausal women, higher fitness levels were associated with higher antioxidant enzyme activity and lower levels of oxidative stress.[43]

We examined whether a set of potential mediators (BMI, high blood pressure, hypercholesterolemia, diabetes, and cardiovascular diseases) explained part of the relation between PA and PD. Our findings are not in favor of this hypothesis because the association between PA and PD was little attenuated after adjusting for these covariates. Hence, the potential mediators examined do not seem to be in the causal pathway between PA and PD. These findings suggest that the mechanisms involved in the relation between PA and PD are independent of these variables and are in favor of a direct protective effect of PA, for instance through motor reserve.[44,45]

The main strengths of our study are its large size and long follow-up, which allowed us to perform lagged analyses to address the potential for reverse causation. We used repeated PA measures rather than a single measure and a method specifically designed to allow longitudinal analyses when measurement tools change over the follow-up.[25] Our approach to ascertain patients with PD yielded incidence rates comparable with those in women from Western Europe, in favor of its validity.[17] Finally, few studies specifically examined the relation between PA and PD in women, possibly because PD is more frequent in men than women, whereas our study focused on this understudied population,[46] and analyses were adjusted for the characteristics of reproductive life.

The main limitation of our study is that we used self-reported PA rather than objective measures (e.g., accelerometer) that are considered more valid, although they do not capture all types of PA (underestimating, for example, cycling and carrying a load). They are however difficult and costly to implement on a large scale and only capture PA over a few days. Measurement error is inevitable for self-reported PA but is reduced through the estimation of the latent process by the LPMM[26] and is likely to be nondifferential and leads to underestimated associations. Recent studies showed that both self-reported and objectively measured PA were associated with outcomes such as all-cause and cardiovascular mortality or self-rated health.[47] Second, there were an insufficient number of questions for vigorous and moderate PA to generate separate latent variables. Third, our analyses based on time-varying PA were not adjusted for diet because it was not recorded at baseline; we performed sensitivity analyses adjusted for dietary characteristics associated with PD, showing that diet was not a strong confounder of the association between PA and PD. Fourth, E3N participants are mostly educated and health-conscious teachers who are not representative of the general population. However, it is generally considered that representativeness is not essential for estimating associations, and associations in occupational cohorts are not necessarily different compared with those estimated in the general population.[48,49] Fifth, we did not obtain medical records for all potential patients with PD and used an algorithm based on drug claims with high sensitivity and specificity for those participants to determine PD status. We confirmed the robustness of our main findings using alternative and stricter PD definitions. Last, all the persons who accepted to participate into the study were assigned female at birth; in addition, they accepted to participate in a study aimed at investigating the risk factors associated with cancer and other major noncommunicable diseases in women. Our findings cannot be extended to persons who were assigned intersex at birth and to men assigned female at birth.

In conclusion, our findings reinforce the evidence in favor of the health benefits of PA and provide stronger evidence than previous studies in favor of an inverse association between PA and PD in women not explained by reverse causation. These results are important for planning interventions for PD prevention[50] and warrant further studies to understand which type and level of PA are beneficial.

Acknowledgment

The authors acknowledge all women enrolled in the E3N cohort for their continued participation. They are also grateful to all members of the E3N study group.

Study Funding

This work was realized with data of the E3N cohort (Inserm) and supported by the Mutuelle Générale de l'Education

Nationale (MGEN), Gustave Roussy Institute, and French League Against Cancer for the constitution and maintenance of the cohort. This work has benefited from state aid managed by the National Research Agency (ANR) under the program "Investment in the future" bearing the reference ANR-10-COHO-0006 and under the program "Young researcher" bearing the reference ANR-18-CE36-0006-01, as well as a subsidy from the Ministry of Higher Education, Research and Innovation for public service charges bearing the reference no. 2102918823, 2103236497, and 2103586016 and from IRESP (Institut de recherche en santé publique). B. Portugal is funded for this work by a doctoral grant from the French Ministry of Research.

Disclosure

B. Portugal is funded by a doctoral grant from the French Ministry of Research. F. Artaud, I. Degaey, A. Fournier, and C. Proust-Lima report no competing interests. E. Roze received a honorarium for speech from Orkyn, Aguettant, and Elivie and for participating in an advisory board from Allergan and Merz Pharma; additionally, he received research support from Merz Pharma, Orkyn, Aguettant, Elivie, Ipsen, Allergan, Everpharma, Fondation Desmarest, AMADYS, ADCY5.org, Agence Nationale de la Recherche, Societé Française de Médecine Esthétique, and Dystonia Medical Research Foundation. G. Severi has obtained research and infrastructure grants related to the E3N cohort (see Study Funding section) as well as cancer research grants unrelated to this study. M. Canonico has obtained research grants from Agence Nationale de la Recherche (ANR). A. Elbaz has obtained research grants from the Michael J. Fox Foundation, Plan Ecophyto (French Ministry of Agriculture), and France Parkinson. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology* September 5, 2022. Accepted in final form April 3, 2023. Submitted and externally peer reviewed. The handling editor was Associate Editor Peter Hedera, MD, PhD.

Appendix Authors

Name	Location	Contribution
Berta Portugal, PhD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; and analysis or interpretation of data
Fanny Artaud, PhD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Isabelle Degaey, MD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France	Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data
Emmanuel Roze, MD, PhD	AP-HP, Hôpital Pitié-Salpêtrière, Neurology Department; Sorbonne Université; INSERM U1127, CNRS 7225, Brain Institute, Paris, France	Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data
Agnès Fournier, PhD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France	Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data
Gianluca Severi, PhD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France; Department of Statistics, Computer Science, Applications "G. Parenti" (DISIA), University of Florence, Italy	Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data
Marianne Canonico, PhD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France	Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data
Cécile Proust-Lima, PhD	Université Bordeaux, Inserm, Bordeaux Population Health Research Center, UMR1219, France	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; and analysis or interpretation of data
Alexis Elbaz, MD, PhD	Université Paris-Saclay, UVSQ, Univ. Paris-Sud, Gustave Roussy, Inserm, U1018, Team "Exposome, Heredity, Cancer, and Health," CESP, Villejuif, France	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data

References

1. Dorsey ER, Elbaz A, Nichols E, et al. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2018;17(11):939-953.
2. Xiao J. *Physical Exercise for Human Health*. Springer; 2020.
3. Chen H, Zhang SM, Schwarzschild MA, Hernan MA, Ascherio A. Physical activity and the risk of Parkinson disease. *Neurology*. 2005;64(4):664-669.
4. Logroscino G, Sesso HD, Paffenbarger RS Jr, Lee IM. Physical activity and risk of Parkinson's disease: a prospective cohort study. *J Neurol Neurosurg Psychiatry*. 2006;77(12):1318-1322.
5. Thacker EL, Chen H, Patel AV, et al. Recreational physical activity and risk of Parkinson's disease. *Mov Disord*. 2008;23(1):69-74.
6. Yang F, Trolle Lagerros Y, Bellocco R, et al. Physical activity and risk of Parkinson's disease in the Swedish National March Cohort. *Brain*. 2015;138(2):269-275.
7. Xu Q, Park Y, Huang X, et al. Physical activities and future risk of Parkinson disease. *Neurology*. 2010;75(4):341-348.
8. Saaksjarvi K, Knekt P, Mannisto S, et al. Reduced risk of Parkinson's disease associated with lower body mass index and heavy leisure-time physical activity. *Eur J Epidemiol*. 2014;29(4):285-292.

9. Llamas-Velasco S, Contador I, Mendez-Guerrero A, et al. Physical activity and risk of Parkinson's disease and parkinsonism in a prospective population-based study (NEDICES). *Prev Med Rep*. 2021;23:101485.
10. Fang X, Han D, Cheng Q, et al. Association of levels of physical activity with risk of Parkinson disease: a systematic review and meta-analysis. *JAMA Netw Open*. 2018; 1(5):e182421.
11. LaHue SC, Comella CL, Tanner CM. The best medicine? The influence of physical activity and inactivity on Parkinson's disease. *Mov Disord*. 2016;31(10):1444-1454.
12. Westreich D. *Epidemiology by Design: A Causal Approach to the Health Sciences*. Oxford University Press; 2020.
13. Elbaz A. Prodromal symptoms of Parkinson's disease: implications for epidemiological studies of disease etiology. *Rev Neurol (Paris)*. 2016;172(8-9):503-511.
14. Savica R, Rocca WA, Ahlskog JE. When does Parkinson disease start? *Arch Neurol*. 2010;67(7):798-801.
15. Schrag A, Horsfall L, Walters K, Noyce AJ, Petersen I. Prediagnostic presentations of Parkinson's disease in primary care: a case-control study. *Lancet Neurol*. 2015;14(1): 57-64.
16. Clavel-Chapelon F, Group ENS. Cohort profile: the French E3N cohort study. *Int J Epidemiol*. 2015;44(3):801-809.
17. Canonico M, Artaud F, Degaey I, et al. Incidence of Parkinson's disease in French women from the E3N cohort study over 27 years of follow-up. *Eur J Epidemiol*. 2022; 37(5):513-523.
18. Tehard B, Friedenreich CM, Oppert JM, Clavel-Chapelon F. Effect of physical activity on women at increased risk of breast cancer: results from the E3N cohort study. *Cancer Epidemiol Biomarkers Prev*. 2006;15(1):57-64.
19. Cust AE, Smith BJ, Chau J, et al. Validity and repeatability of the EPIC physical activity questionnaire: a validation study using accelerometers as an objective measure. *Int J Behav Nutr Phys Act*. 2008;5(1):33.
20. The InterAct Consortium. Validity of a short questionnaire to assess physical activity in 10 European countries. *Eur J Epidemiol*. 2012;27(1):15-25.
21. Insee. Base des unités urbaines [online]. Accessed May 11, 2019. insee.fr/fr/information/2115018.
22. Van Liere MJ, Lucas F, Clavel F, Slimani N, Villeminot S. Relative validity and reproducibility of a French dietary history questionnaire. *Int J Epidemiol*. 1997;26(-suppl 1):S128.
23. Desai RJ, Glynn RJ, Wang S, Gagne JJ. Performance of disease risk score matching in nested case-control studies: a simulation study. *Am J Epidemiol*. 2016;183(10):949-957.
24. 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(2).
25. Wagner M, Dartigues JF, Samieri C, Proust-Lima C. Modeling risk-factor trajectories when measurement tools change sequentially during follow-up in cohort studies: application to dietary habits in prodromal dementia. *Am J Epidemiol*. 2018;187(4): 845-854.
26. 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.
27. Commenges D, Jacqmin-Gadda H. Chapter 4. Models for longitudinal data. In: Commenges D, Jacqmin-Gadda H, eds. *Dynamical Statistical Models*. Chapman & Hall; 2016:63-82.
28. Pesce G, Artaud F, Roze E, et al. Reproductive characteristics, use of exogenous hormones and Parkinson disease in women from the E3N study. *Brain*. 2022: awac440. doi:10.1093/brain/awac440
29. Noyce AJ, Bestwick JP, Silveira-Moriyama L, et al. Meta-analysis of early nonmotor features and risk factors for Parkinson disease. *Ann Neurol*. 2012;72(6):893-901.
30. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol*. 2009;170(2):244-256.
31. Bianchi VE, Rizzi L, Somaa F. The role of nutrition on Parkinson's disease: a systematic review. *Nutr Neurosci*. 2022:1-24. doi:10.1080/1028415X.2022.2073107
32. Sun F, Norman IJ, While AE. Physical activity in older people: a systematic review. *BMC Public Health*. 2013;13(1):449.
33. Tsukita K, Sakamaki-Tsukita H, Takahashi R. Long-term effect of regular physical activity and exercise habits in patients with early Parkinson disease. *Neurology*. 2022; 98(8):e859-e871.
34. van der Kolk NM, de Vries NM, Kessels RPC, et al. Effectiveness of home-based and remotely supervised aerobic exercise in Parkinson's disease: a double-blind, randomised controlled trial. *Lancet Neurol*. 2019;18(11):998-1008.
35. Schenkman M, Moore CG, Kohrt WM, et al. Effect of high-intensity treadmill exercise on motor symptoms in patients with de novo Parkinson disease: a phase 2 randomized clinical trial. *JAMA Neurol*. 2018;75(2):219-226.
36. Johansson ME, Cameron IGM, Van der Kolk NM, et al. Aerobic exercise alters brain function and structure in Parkinson's disease: a randomized controlled trial. *Ann Neurol*. 2022;91(2):203-216.
37. Hughes KC, Gao X, Molsberry S, Valeri L, Schwarzschild MA, Ascherio A. Physical activity and prodromal features of Parkinson disease. *Neurology*. 2019;93(23): e2157-e2169.
38. Palasz E, Niewiadomski W, Gasiorowska A, Wysocka A, Stepniewska A, Niewiadomska G. Exercise-induced neuroprotection and recovery of motor function in animal models of Parkinson's disease. *Front Neurol*. 2019;10:1143.
39. Xu X, Fu Z, Le W. Exercise and Parkinson's disease. *Int Rev Neurobiol*. 2019;147: 45-74.
40. Gleeson M, Bishop NC, Stensel DJ, Lindley MR, Mastana SS, Nimmo MA. The anti-inflammatory effects of exercise: mechanisms and implications for the prevention and treatment of disease. *Nat Rev Immunol*. 2011;11(9):607-615.
41. Hou L, Chen W, Liu X, Qiao D, Zhou FM. Exercise-induced neuroprotection of the nigrostriatal dopamine system in Parkinson's disease. *Front Aging Neurosci*. 2017;9: 358.
42. Erickson KI, Gildengers AG, Butters MA. Physical activity and brain plasticity in late adulthood. *Dialogues Clin Neurosci*. 2013;15(1):99-108.
43. Pialoux V, Brown AD, Leigh R, Friedenreich CM, Poulin MJ. Effect of cardiorespiratory fitness on vascular regulation and oxidative stress in postmenopausal women. *Hypertension*. 2009;54(5):1014-1020.
44. Bastos PAD, Barbosa R. Motor reserve: how to build neuronal resilience against ageing and neurodegeneration? *Rev Neurol (Paris)*. 2022;178(8):845-854.
45. Sunwoo MK, Lee JE, Hong JY, et al. Premorbid exercise engagement and motor reserve in Parkinson's disease. *Parkinsonism Relat Disord*. 2017;34:49-53.
46. Subramanian I, Mathur S, Oosterbaan A, Flanagan R, Keener AM, Moro E. Unmet needs of women living with Parkinson's disease: gaps and controversies. *Mov Disord*. 2022;37(3):444-455.
47. Evenson KR, Wen F, Herring AH. Associations of accelerometry-assessed and self-reported physical activity and sedentary behavior with all-cause and cardiovascular mortality among US adults. *Am J Epidemiol*. 2016;184(9):621-632.
48. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012-1014.
49. Batty GD, Shipley M, Tabak A, et al. Generalizability of occupational cohort study findings. *Epidemiology*. 2014;25(6):932-933.
50. Janssen Daalen JM, Schootemeijer S, Richard E, Darweesh SKL, Bloem BR. Lifestyle interventions for the prevention of Parkinson disease: a recipe for action. *Neurology*. 2022;99(7 suppl 1):42-51.