

Daily steps are a predictor of, but perhaps not a risk factor for Parkinson's disease: findings from the UK Biobank

Received: 11 July 2025

Accepted: 9 November 2025

Cite this article as: Acquah, A., Creagh, A., Hamy, V. *et al.* Daily steps are a predictor of, but perhaps not a risk factor for Parkinson's disease: findings from the UK Biobank. *npj Parkinsons Dis.* (2025). <https://doi.org/10.1038/s41531-025-01214-6>

Aidan Acquah, Andrew Creagh, Valentin Hamy, Alaina Shreves, Charilaos Zisou, Charlie Harper, Stefan van Duijvenboden, Chrystalina Antoniadou, Derrick Bennett, David Clifton & Aiden Doherty

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Daily steps are a predictor of, but perhaps not a risk factor for Parkinson's disease: Findings**from the UK Biobank**

Authors: Aidan Acquah^{1,2,3}, Dr. Andrew Creagh^{2,3}, Dr. Valentin Hamy⁴, Alaina Shreves^{1,2,5}, Charilaos Zisou^{1,2}, Dr. Charlie Harper^{1,2}, Dr. Stefan Van Duijvenboden^{1,2}, Prof. Chrystalina Antoniadou⁶, Prof. Derrick Bennett¹, Prof. David Clifton^{3,7}, Prof. Aiden Doherty^{1,2}

Affiliations:

1. Nuffield Department of Population Health, University of Oxford, Oxford, UK
2. Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK
3. Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK
4. GSK, UK
5. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
6. Nuffield Department of Clinical Neurosciences, Medical Sciences Division, University of Oxford, Oxford, UK
7. Oxford Suzhou Centre for Advanced Research, University of Oxford, Suzhou 215123, Jiangsu, China

Correspondence to:

Aiden Doherty, PhD

Professor of Biomedical Informatics

University of Oxford

Nuffield Department of Population Health

Big Data Institute, Old Road Campus

Oxford, UK OX3 7LF

Email: aiden.doherty@ndph.ox.ac.uk

Abstract

Previous studies link lower physical activity with incident Parkinson's disease (PD) but rely on self-reported data and fail to address reverse causation. This study used accelerometer-derived daily step count, an objective measure of physical activity, to examine its association with incident PD in the UK Biobank, within successive periods of follow-up. PD cases were identified through hospital inpatient and death data, and associated with machine learning-derived step counts using Cox regression models, adjusted for age, sex, demographic and lifestyle factors. For 94,696 valid participants and a median follow-up of 7.9 years, 407 incident PD cases occurred. Every 1,000 steps higher were associated with 8% lower risk of PD (hazard ratio 0.92, 95% confidence interval 0.89-0.94). However, this association was no longer statistically significant when excluding follow-up periods closer to the time of accelerometer wear, suggesting that low activity may be an early marker, but not a risk factor for PD.

ARTICLE IN PRESS

Introduction

Parkinson's disease (PD) is the second most prevalent and most rapidly growing neurodegenerative disorder^{1,2}, with an estimated prevalence of 9.4 million³ cases in 2020, marking a substantial increase from 5.2 million cases in 2004⁴. The prodromal phase of PD has many well-recognised motor-related symptoms, such as subtle motor dysfunction, that are believed to occur over a decade before diagnosis⁵. Recognising these early indicators could provide crucial insights for improved disease understanding and may offer the possibility of identifying modifiable risk factors for developing PD.

In existing literature, low self-reported physical activity has been associated with a higher risk of incident PD^{6,7}. Despite these observations, uncertainties remain, stemming from a reliance on self-report measurements of physical activity, which are crude and less reliable, creating greater uncertainties in observed associations^{8,9}. In addition, PD's progressive and prolonged development poses challenges in distinguishing between incident and prevalent cases. This ambiguity can introduce reverse-causation bias, where the observed associations may be influenced by the change in behaviour of individuals who already have underlying PD, rather than vice-versa. One common approach to mitigate potential reverse causation is to exclude the first few years of follow-up data. In PD studies, there is inconsistency around the number of years removed across studies, ranging between four to ten years, however, reduced activity levels were still significantly associated with incident PD after this exclusion⁶.

Daily step counts serve as a valuable proxy for measuring physical activity levels. One of the primary advantages of using average daily steps over other physical activity metrics is their simplicity, allowing for clear communication and understanding among the general population. This ease of interpretation helps in translating research findings into effective public health messaging¹⁰. Existing research into the association between physical activity and incident PD have typically used metabolic equivalent of task (MET) scores to estimate physical activity⁶. However, the increasing adoption of wearable devices, capable of estimating step counts, has encouraged research on their accuracy and reliability, particularly in populations diagnosed with PD^{11,12}. The translatable nature of daily step counts encourages the expansion of this research into incident PD, allowing for results to offer insight into public health recommendations for potential disease prevention.

We, therefore, aimed to investigate the association between accelerometer-measured daily step counts and incident Parkinson's disease, and how this association changed with latter windows of follow-up.

Results

Participant characteristics

After exclusions, this study consisted of 94,696 UK Biobank physical activity monitoring study¹³ participants (Figure 1).

[Figure 1]

As shown in Table 1, among these participants, 56% were female, 97% identified as White. Most were in the least to second least deprived TDI quintile (<-1.509), compared to the TDI distribution of the general UK population in 2011¹⁴. Those with step counts in the highest quintile (12,369+ daily steps) were younger, and had lower BMI compared to those in the lowest step count group (<6,276 daily steps).

[Table 1]

The overall mean of median daily steps calculated in the study population was 9,446 steps and the standard deviation was 3,822 steps. Participants with prevalent PD (n = 150), who were excluded from the study, exhibited the lowest average step count, as compared to incident and non-PD participants. Among those with incident PD, average median daily step counts were progressively higher with longer follow-up time from accelerometer wear to the first reporting of PD in their health records. Despite this, the average median daily steps in the incident PD population remained lower than that of non-PD participants, as illustrated in Supplementary Figure 2.

Association of daily steps with incident Parkinson's disease

Participants had a median follow-up time of 7.9 years (IQR: 7.4, 8.4), during which 407 incident PD cases occurred, with a median time to diagnosis of 5.2 years (IQR: 3.1, 6.8). We observed a strong and significant inverse linear association between median daily step quintiles and incident PD using data from the entire duration of follow-up (Figure 2). Those who walked over 12,369 steps daily had a 59% lower risk of PD (HR 0.41; group specific 95% CI: 0.31-0.54), compared to those walking less than 6,276 steps daily (HR 1.00; group-specific 95% CI: 0.84-1.19), after multivariable adjustment. There was no evidence of departure from linearity (p=0.65).

[Figure 2]

For the multivariable model, every 1,000 daily steps higher was associated with an 8% lower risk of PD (HR 0.92; 95% CI 0.89-0.94). The sequential adjustment of covariates from the univariable to multivariable model showed minimal changes to the strength of the association and the chi-squared increased by 33% (shown in the supplementary material). The further adjustment of additional covariates minimally altered the observed

associations. However, adjustment for sleep duration did attenuate the strength of association, with a reduction in the chi-squared by 64%, as shown in Supplementary Figure 5. Furthermore, we observed no significant difference across subgroups, including age groups, sex, BMI status and history of depression. After removing 9,734 individuals with prevalent neurological cases, the observed association was similar to that observed in our main study sample (HR 0.92; 95% CI 0.89-0.95).

To address potential reverse causality, we assessed how the association between step counts and PD risk may have changed over follow-up periods of <2, 2-4, 4-6 and 6+ (Figure 3). We observed the strongest association within the first two years of follow-up (HR 0.83; 95% CI 0.70-0.90), during which there were 55 incident cases (Figure 3). However, the association was attenuated for later windows of follow-up, trending towards not statistically significant. This is despite having higher relative numbers of cases in latter periods of follow-up, reducing uncertainty in estimates. For the follow-up period of over 6 years, the HR was no longer statistically significant, with a hazard of 0.96 (95% CI: 0.92 – 1.01).

[Figure 3]

These findings remained consistent after running a sensitivity analysis using a step count algorithm trained on both healthy and PD populations. The estimate median daily step counts between the two algorithms, healthy and PD-trained compared to healthy-only trained, showed very strong correlation, with a Pearson correlation coefficient¹⁵ of 0.97, as seen in Supplementary Figure 6. Furthermore, a significant association was found between PD-trained median daily step count and incident PD using all periods of follow-up, with covariate adjustment, as shown in Supplementary Figures 7 and 8. As seen previously, this HR, however, was no longer statically significant when considering follow-up periods of over 6 years (HR 0.96; 95% CI 0.92-1.00).

Discussion

Higher accelerometer-derived physical activity levels, indicated by median daily step counts, were associated with a lower risk of incident Parkinson's disease (PD) when using data from all periods of follow-up. This association remained statistically significant after multivariable adjustment, investigating potential confounding factors, and among subpopulations of interest. However, this association was no longer statically significant when considering later periods of follow-up, indicating that reverse causation bias is likely.

Our findings are in contrast to some previous literature in the field that supports an association between physical activity levels and incident PD, even after the consideration of reverse causation^{6,16-18}. In a meta-analysis of this

association with a median (range) follow-up of 12 (6.1-22.0) years, the relative risk for incident PD comparing the highest and lowest categories of total physical activity was 0.79 (95% CI 0.68-0.91), and remained significant when the first four to ten years of follow-up was removed from the constituent studies (0.76; 95% CI 0.64-0.91)⁶. In our study, removing years of follow-up reduced the strength of the observed association. Similar attenuations between physical activity and diseases of ageing have been reported in other studies. Our findings are more consistent with results from a study on physical activity and dementia that was conducted in the Whitehall Study and Million Women Study^{19,20}.

These results, however, support existing literature in finding lower physical activity levels to be a strong indicator/feature of prodromal or early PD^{18,21}. One study reported that this reduction in activity may occur up to 4 years before a clinical diagnosis¹⁸. A separate study assessing minutes of daily physical activity²¹ found a reduction of 29% in comparing clinical PD participants to controls, in minutes of daily physical activity. There are several hypotheses for these activity changes, including changes in overall disability, poorer walking performance, and fear of falls/anxiety/depression²¹. It should be noted, however, that this reduction in activity is not fully explained by known factors, with this study encouraging further research into explaining this behaviour²¹.

Our study has several strengths. First, we used objective measures of activity derived from accelerometer data, which offers greater reliability in the measurement of this exposure compared to questionnaire-based studies. Second, the use of step count as a measure of daily physical activity offers the opportunity for better communication as step counts are easily translatable for public health messaging. Next, this work has built on previous studies investigating PD using the UK Biobank physical activity monitoring study^{16,17,22}, however with longer follow-up, while also addressing some of these studies' limitations. Of these limitations, we explored how the association may change over follow-up periods to better describe the potential impact of reverse causation.

There are also some limitations to our study. Firstly, this analysis was purely observational; hence we cannot exclude the possibility of unmeasured or residual confounding in the observed associations, despite considering a comprehensive set of confounders. Secondly, the UK Biobank provides limited health data linkage for participants, which we use to identify participants with PD. We use the approach recommended by UK Biobank to ascertain PD, combining hospital inpatient, death register and self-reported data²³, which may tend to identify more severe cases. These records do not provide information regarding the severity or subtype of PD, which may offer further insights into this association. Furthermore, primary care data could bolster PD cases, but is currently limited to just 45% of participants and only up to 2017¹³. Lastly, this study is limited by the number of cases

observed within period of follow-up. With 407 incident cases of PD from all periods of follow-up, and just 156 diagnosed at least 6 years after participating in the study, we are somewhat limited in power to observe small effects of greater activity. We are also limited by a maximum follow-up time of 9.4 years. The PD prodromal window is typically reported to be at least 10 years prior to a clinical diagnosis⁵.

Future studies in this population can benefit from higher case numbers from a longer follow-up period, while also giving further insights into this association, beyond the prodromal window of PD. Further investigations should also explore the association between other accelerometer-derived metrics and incident PD, such as metrics relating to gait speed, stride length or the distribution of steps throughout the week. Furthermore, we would encourage similar analyses in other study populations, such as the Parkinson's Progression Markers Initiative (PPMI)-Verily study^{24,25}, containing a higher proportion of individuals with prodromal PD, and the All of Us study²⁶, containing greater diversity in race, ethnicity and socio-economic status.

Results from this prospective study add to the literature on physical activity and incident PD, suggesting that higher daily step counts were associated with lower risk of incident PD, but only for short follow-up periods. Therefore, the observed association may be explained by reverse causation bias. This work supports the continued labelling of low physical activity as a marker for Parkinson's disease, rather than as a risk factor leading to Parkinson's disease.

Methods

Study Cohort

The UK Biobank is a prospective cohort study of 502,536 adults in the United Kingdom who were recruited between 2006 and 2010. The consented participants completed a questionnaire, interview, and provided physical measurements, as well as blood, urine, and saliva samples^{27,28}. A subgroup of participants consented to wear an Axivity AX3 wrist-worn accelerometer for up to 7 days between 2013-2015¹³. All participants provided written informed consent and the study was approved by the National Information Governance Board for Health and Social Care and the National Health Service North West Multicentre Research Ethics Committee (06/MRE08/65).

Accelerometer-derived physical activity processing

Accelerometer data was processed to derive step counts using the OxWearables "stepcount" package (version 2.1.5), a hybrid self-supervised machine learning model that was trained on ground truth free-living data and counts peaks in detected walking windows of behaviour²⁹. Steps were predicted for each participant's full period of collected data, with missing periods due to non-wear imputed by averaging the step counts in the corresponding

times across valid days²⁹. We then calculated the median of the imputed daily step counts, across the period of monitoring, as the primary exposure in this study.

Ascertainment of Parkinson's disease

The diagnosis of PD was obtained from a combination of the UK Biobank's linked hospital data; inpatient hospital admissions data, and death data. Diagnoses were coded according to the International Classification of Diseases (version 10) coding system, with outcomes recorded as the first instance of a G20 diagnostic code²³. This follows the approach recommended by the UK Biobank²³, and used by other studies in studying PD in the physical activity monitoring cohort^{16,17,22}. Incident PD however was defined at their first diagnosis of PD. Participants were censored at the first date of PD diagnosis, date of death, or at the end of their follow-up period (31 October 2022 for England, 31 August 2022 for Scotland, and 31 May 2022 for Wales), whichever came first³⁰.

Statistical Analysis

We processed the accelerometer data from 103,660 participants. We excluded participants with device calibration or data reading errors (>1% of values outside +/-8g range), and those whose data could not be processed. In addition, participants were excluded due to inadequate wear time, either less than 72 hours of detected wear, or non-wear at a particular time of day for all days of monitoring. We did not implement any day level exclusions, to remove any data due to insufficient wear on a given day. Further exclusions were made for unreasonably high average acceleration (>100 mg). This exclusion criteria aligns with numerous existing studies on this dataset^{22,29,31,32}. We also excluded all 150 prevalent PD and 22 other Parkinsonism cases, diagnosed before accelerometer wear, and participants with missing healthcare linkages or covariate data. The final analysis included 94,696 participants (Figure 1).

A Cox proportional hazard regression model³³ was used to estimate adjusted hazard ratios (HRs) with 95% confidence intervals (CIs) for the association between daily step counts and incident PD. The association was first evaluated using quintiles of median daily step count. CIs when using quintiles of daily steps as the primary exposure were estimated using floating absolute risk, to represent group-specific variances in hazard, therefore referred to as group-specific CIs³⁴.

We then evaluated how the risk of PD was associated with each 1,000 steps per day higher. We sequentially applied the full set of adjustments to the multivariable model. The impact of these sequential adjustments were assessed using chi-squared test. Specifically, for the per-1000 median daily steps with one degree of freedom, we compared the chi-squared values before and after adjustment. A likelihood ratio test comparing quintiles of daily

steps as categorical and ordinal was performed to test that the association between daily steps and incident PD did not significantly depart from linearity³⁵. This allows for the reporting of hazard ratios per 1,000 steps rather than quintiles of step count. At each step of this sequential adjustment, confounding was assessed by observing the chi-squared from a likelihood ratio test of the association.

Multivariable models used age as the time scale, as age was a strong potential confounder in this analysis^{36,37}. These models were adjusted for the following covariates: season of device wear (spring, summer, autumn, winter), sex (female, male), ethnicity (white, non-white), Townsend Deprivation Index (TDI, stratified by quintiles in UK population¹⁴), geographic region of assessment centre (London, Midlands, Yorkshire, Northeast, Northwest, Southeast, Southwest England, Scotland, Wales), educational attainment (degree, diploma, A/AS levels, professional qualification, GCSE/O levels, no degree/qualification), employment status (employed, not employed), smoking status (never, previous, current) and alcohol consumption (never, < 3 times per week, ≥ 3 times per week) and coffee intake (non-drinker, ≤ 2 cups a day, > 2 cups a day). We adjusted for season of device wear as a technical covariate, which explains part of the variation in physical activity levels³⁸. The remaining covariates were chosen as known risk factors for PD, particularly those from the PREDICT-PD risk model³⁶, with possible association to physical activity, and reliable and accurate recording in UK Biobank records, while not hypothesised to be on the causal pathway^{36,39}. All covariate data were provided by participants at the UK Biobank baseline assessment visit, with details provided in Supplementary Table 2. The proportional hazards assumption was tested using Schoenfeld residuals⁴⁰, and no violations were observed.

We conducted several orthogonal analyses using the models assessing the risk for each 1,000 daily steps higher. First, we expanded our multivariable-adjusted models by additionally adjusting for body mass index (BMI) at baseline, depression, type 2 diabetes, constipation, bladder dysfunction derived through a combination of self-report and health records, and sleep duration derived from accelerometer data. These covariates were chosen as known health-related PD risk factors with an association with physical activity, that were hypothesised possibly be on the causal pathway^{36,41,42}. We then explored whether the observed associations differed within subpopulations of interest, including by age groups, sex, BMI (<25, 25–30, 30+ kg/m²)⁴³ and history of depression. Additionally, we performed a sensitivity analysis examining the associations after excluding participants with any history of neurological disorder. To assess potential reverse causation and evaluate how associations changed over time, we estimated HRs within successive follow-up intervals (0–2, 2–4, 4–6, and ≥ 6 years) using Lexis expansion^{44,45}. For each participant, total follow-up time was divided into these intervals, generating one record for each period in which the individual contributed person-time. Each record represented time at risk within that

interval, with follow-up left-truncated at the start and right-censored at the end of the interval. HRs were then estimated separately for each interval to assess temporal changes in the associations, following a similar approach to that described by Floud et al. (2020)²⁰.

We also performed a final sensitivity analysis, to ensure the model could reliably detect steps in individuals at-risk of incident PD. To do so, we modified the stepcount algorithm deployed on the UK Biobank population, so that instead of only being trained on the healthy OxWalk population, it was additionally trained on the Michael J. Fox Foundation Levodopa Response (MJFF-LR) dataset⁴⁶, consisting of wrist-worn accelerometer data collected from 28 clinical PD individuals. The full association analysis pipeline was then re-run, using this PD-trained stepcount algorithm, to observe how this may affect our findings.

Statistical analyses were performed using R, version 4.3.2, using Cox regression models from the survival package, version 3.8-3. Results were reported according to Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Supplementary Table 3)⁴⁷.

Declarations

Data availability

This study uses UK Biobank data, available at: <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>.

Code availability

The underlying code for this study is available in Github and can be accessed via this link https://github.com/OxWearables/steps_PD.

Acknowledgements

We thank Prof. Sarah Floud and Tonima Trisa for their guidance in the handling of reverse causation in this association.

Funding

This research has been conducted using the UK Biobank Resource under Application Number 59070. This research is supported by the EPSRC Centre for Doctoral Training in Health Data Science [EP/S02428X/1], and by GSK plc as part of an iCase studentship [EP/V519741/1]. AD and SvD are supported by the Wellcome Trust [223100/Z/21/Z]. AD and DB are supported by Novo Nordisk. AD, CH, and DB are supported by SwissRe. CZ and AD are supported by the Oxford BHF Centre of Research Excellence [RE/18/3/34214]. DC is supported by

the Pandemic Sciences Institute at the University of Oxford; the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC); an NIHR Research Professorship; a Royal Academy of Engineering Research Chair; the Wellcome Trust funded VITAL project [204904/Z/16/Z]; the EPSRC [EP/W031744/1]; and the InnoHK Hong Kong Centre for Cerebro-cardiovascular Engineering (COCHE). AS is supported by the National Institutes of Health's Intramural Research Program and by the National Institutes of Health's Oxford Cambridge Scholars Program. CA is supported by the National Institute for Health Research Oxford Biomedical Research Centre and has received research grant support from UCB Pharma.

This research was supported [in part] by the Intramural Research Program of the National Institutes of Health (NIH). The contributions of the NIH author(s) are considered Works of the United States Government. The findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Authors' contributions

AA, AC, DC, and AD contributed to the conception and design of the project. AA wrote the main manuscript text and prepared the figures, with critical input and guidance from CH, AS, and CZ. AA, AC, VH, AS, CZ, CH, SvD, CA, DB, DC, and AD contributed to manuscript revisions, reviewed the final version, and approved the submitted manuscript.

Competing interests

AS, CZ, CH, SvD, CA and DB declare no competing interests. AA, AC and DC were supported by GSK during this research. VH is an employee at GSK. AD has received research grants from SwissRe and Novo Nordisk. AA, AC and AD have released the stepcount package software under an academic use licence that could result in commercial entities paying license fees for their use. The funders/sponsors had no role in the design of this study or the analysis and interpretation of the results.

Rights and Permissions

For the purpose of open access, the authors have applied a CC-BY public copyright licence to any author accepted manuscript version arising from this submission.

References

1. Mhyre TR, Boyd JT, Hamill RW, Maguire-Zeiss KA. Parkinson's Disease. *Subcell Biochem.* 2012;65:389-455. doi:10.1007/978-94-007-5416-4_16
2. 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. doi:10.1016/S1474-4422(18)30295-3
3. Maserejian N, Vinikoor-Imler L, Dilley A. Estimation of the 2020 Global Population of Parkinson's Disease (PD). *MDS Abstr.* Published online 2020.
4. Ahlrichs C, Lawo M. Parkinson's Disease Motor Symptoms in Machine Learning: A Review. *ArXiv Prepr ArXiv13123825.* Published online 2013.
5. Postuma RB, Berg D. Advances in Markers of Prodromal Parkinson Disease. *Nat Rev Neurol.* 2016;12(11):622-634. doi:10.1038/nrneurol.2016.152
6. 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. doi:10.1001/jamanetworkopen.2018.2421
7. 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. doi:10.1002/mds.26728
8. Sallis JF, Saelens BE. Assessment of Physical Activity by Self-Report: Status, Limitations, and Future Directions. *Res Q Exerc Sport.* 2000;71(sup2):1-14. doi:10.1080/02701367.2000.11082780
9. R J Shephard. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med.* 2003;37(3):197. doi:10.1136/bjism.37.3.197
10. Bassett DR, Toth LP, LaMunion SR, Crouter SE. Step Counting: A Review of Measurement Considerations and Health-Related Applications. *Sports Med Auckl Nz.* 2017;47(7):1303-1315. doi:10.1007/s40279-016-0663-1
11. Shokouhi N, Khodakarami H, Fernando C, Osborn S, Horne M. Accuracy of Step Count Estimations in Parkinson's Disease Can Be Predicted Using Ambulatory Monitoring. *Front Aging Neurosci.* 2022;14. doi:10.3389/fnagi.2022.904895
12. Bianchini E, Galli S, Alborghetti M, et al. Four Days Are Enough to Provide a Reliable Daily Step Count in Mild to Moderate Parkinson's Disease through a Commercial Smartwatch. *Sensors.* 2023;23(21):8971. doi:10.3390/s23218971
13. Doherty A, Jackson D, Hammerla N, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PLOS ONE.* 2017;12(2):e0169649. doi:10.1371/journal.pone.0169649
14. Yousaf S, Bonsall A. UK Townsend Deprivation Scores from 2011 census data. Presented at: September 7, 2017. https://s3-eu-west-1.amazonaws.com/statistics.digitalresources.jisc.ac.uk/dkan/files/Townsend_Deprivation_Scores/UK%20Townsend%20Deprivation%20Scores%20from%202011%20census%20data.pdf
15. 11. Correlation and regression | The BMJ. The BMJ | The BMJ: leading general medical journal. Research. Education. Comment. October 28, 2020. Accessed March 24, 2025.

<https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression>

16. Zhao A, Cui E, Leroux A, Lindquist MA, Crainiceanu CM. Evaluating the prediction performance of objective physical activity measures for incident Parkinson's disease in the UK Biobank. *J Neurol*. Published online August 24, 2023. doi:10.1007/s00415-023-11939-0
17. Liu M, Gan X, Ye Z, et al. Association of accelerometer-measured physical activity intensity, sedentary time, and exercise time with incident Parkinson's disease. *Npj Digit Med*. 2023;6(1):224. doi:10.1038/s41746-023-00969-7
18. Xu Q, Park Y, Huang X, et al. Physical activities and future risk of Parkinson disease. *Neurology*. 2010;75(4):341-348. doi:10.1212/WNL.0b013e3181ea1597
19. Sabia S, Dugravot A, Dartigues JF, et al. Physical activity, cognitive decline, and risk of dementia: 28 year follow-up of Whitehall II cohort study. *BMJ*. Published online June 22, 2017;j2709. doi:10.1136/bmj.j2709
20. Floud S, Simpson RF, Balkwill A, et al. Body mass index, diet, physical inactivity, and the incidence of dementia in 1 million UK women. *Neurology*. 2020;94(2):e123-e132. doi:10.1212/WNL.0000000000008779
21. van Nimwegen M, Speelman AD, Hofman-van Rossum EJM, et al. Physical inactivity in Parkinson's disease. *J Neurol*. 2011;258(12):2214-2221. doi:10.1007/s00415-011-6097-7
22. Schalkamp AK, Peall KJ, Harrison NA, Sandor C. Wearable movement-tracking data identify Parkinson's disease years before clinical diagnosis. *Nat Med*. 2023;29(8):2048-2056. doi:10.1038/s41591-023-02440-2
23. Bush K, Rannikmae K, Wilkinson T, Schnier C, Sudlow C. Definitions of Parkinson's Disease and the Major Causes of Parkinsonism, UK Biobank Phase 1 Outcomes Adjudication. Published online March 2018. https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/alg_outcome_pdp.pdf
24. Marek K, Chowdhury S, Siderowf A, et al. The Parkinson's progression markers initiative (PPMI) - establishing a PD biomarker cohort. *Ann Clin Transl Neurol*. 2018;5(12):1460-1477. doi:10.1002/acn3.644
25. Dorsey ER, Marks Jr. WJ. Verily and Its Approach to Digital Biomarkers. *Digit Biomark*. 2017;1(1):96-99. doi:10.1159/000476051
26. Bailey CP, Dodd KW, McClain JJ, Seo I, Wheeler W, Wolff-Hughes DL. Fitbit Physical Activity and Sleep Data in the All of Us Research Program: Data Exploration and Processing Considerations for Research. *Med Sci Sports Exerc*. Published online July 3, 2025:10.1249/MSS.0000000000003804. doi:10.1249/MSS.0000000000003804
27. UK Biobank. UK Biobank Enable your research. June 24, 2021. Accessed February 13, 2024. <https://www.ukbiobank.ac.uk/enable-your-research>
28. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Med*. 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779

29. Small SR, Chan S, Walmsley R, et al. *Development and Validation of a Machine Learning Wrist-Worn Step Detection Algorithm with Deployment in the UK Biobank*. *Public and Global Health*; 2023. doi:10.1101/2023.02.20.23285750
30. UK Biobank. UK Biobank. May 21, 2024. Accessed June 26, 2024. <https://www.ukbiobank.ac.uk>
31. Walmsley R, Chan S, Smith-Byrne K, et al. Reallocation of Time between Device-Measured Movement Behaviours and Risk of Incident Cardiovascular Disease. *Br J Sports Med*. Published online September 2021:bjsports-2021-104050. doi:10.1136/bjsports-2021-104050
32. Doherty A, Smith-Byrne K, Ferreira T, et al. GWAS Identifies 14 Loci for Device-Measured Physical Activity and Sleep Duration. *Nat Commun*. 2018;9(1):5257. doi:10.1038/s41467-018-07743-4
33. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B Methodol*. 1972;34(2):187-220.
34. Easton DF, Peto J, Babiker AG. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. *Stat Med*. 1991;10(7):1025-1035. doi:10.1002/sim.4780100703
35. Kirkwood BR, Sterne JAC. *Essential Medical Statistics*. John Wiley & Sons, Incorporated; 2003.
36. Noyce AJ, Bestwick JP, Silveira-Moriyama L, et al. PREDICT-PD: Identifying risk of Parkinson's disease in the community: methods and baseline results. *J Neurol Neurosurg Psychiatry*. 2014;85(1):31-37. doi:10.1136/jnnp-2013-305420
37. Reeve A, Simcox E, Turnbull D. Ageing and Parkinson's disease: Why is advancing age the biggest risk factor? *Ageing Res Rev*. 2014;14(100):19-30. doi:10.1016/j.arr.2014.01.004
38. Turrisi TB, Bittel KM, West AB, et al. Seasons, weather, and device-measured movement behaviors: a scoping review from 2006 to 2020. *Int J Behav Nutr Phys Act*. 2021;18(1):24. doi:10.1186/s12966-021-01091-1
39. Najafi F, Mansournia MA, Abdollahpour I, Rohani M, Vahid F, Nedjat S. Association between socioeconomic status and Parkinson's disease: findings from a large incident case-control study. *BMJ Neurol Open*. 2023;5(1):e000386. doi:10.1136/bmjno-2022-000386
40. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982;69(1):239-241. doi:10.1093/biomet/69.1.239
41. Postuma RB, Berg D. Advances in markers of prodromal Parkinson disease. *Nat Rev Neurol*. 2016;12(11):622-634. doi:10.1038/nrneurol.2016.152
42. Dommershuijsen LJ, Boon AJW, Ikram MK. Probing the Pre-diagnostic Phase of Parkinson's Disease in Population-Based Studies. *Front Neurol*. 2021;12. doi:10.3389/fneur.2021.702502
43. Weir CB, Jan A. BMI Classification Percentile And Cut Off Points. In: *StatPearls*. StatPearls Publishing; 2023. Accessed January 16, 2024. <http://www.ncbi.nlm.nih.gov/books/NBK541070/>
44. Plummer M, Carstensen B. Lexis: An R Class for Epidemiological Studies with Long-Term Follow-Up. *J Stat Softw*. 2011;38:1-12. doi:10.18637/jss.v038.i05
45. Hong LS, Lewington S. Lexis Expansion- Age-at-Risk Adjustment for Survival Analysis. Published online 2013.

46. Snyder P, Tummalacherla M. MJFF Levodopa Response Study. Published online 2019. doi:10.7303/SYN20681023
47. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *The Lancet*. 2007;370(9596):1453-1457. doi:10.1016/S0140-6736(07)61602-X

ARTICLE IN PRESS

Table 1: Count by Demographic Characteristics in the UK Biobank. TDI=Townsend deprivation index,

BMI=Body Mass Index, SD=standard deviation

Characteristic	Median daily step count					All
	<6,276	6,276-8,145	8,146-9,966	9,967-12,368	12,369+	
No. of Participants	18,933	18,939	18,941	18,942	18,941	94,696
Demographic factors						
Age, mean (SD), (years)	63.3 (7.96)	62.6 (7.89)	62.3 (7.85)	62.1 (7.74)	61.4 (7.65)	62.3 (7.84)
Female, No. (%)	10,792 (57.0%)	10,916 (57.6%)	10,780 (56.9%)	10,727 (56.6%)	10,158 (53.6%)	53,373 (56.4%)
White ethnicity, No. (%)	18,288 (96.6%)	18,321 (96.7%)	18,377 (97.0%)	18,416 (97.2%)	18,406 (97.2%)	91,808 (97.0%)
Socio-economic factors, No. (%)						
Most deprived UK population TDI Quintile (≥ 2.516) ¹⁴	2,242 (11.8%)	1,711 (9.0%)	1,720 (9.1%)	1,759 (9.3%)	1,881 (9.9%)	9,313 (9.8%)
No education/professional qualifications	2,157 (11.4%)	1,643 (8.7%)	1,489 (7.9%)	1,305 (6.9%)	1,270 (6.7%)	7,864 (8.3%)
Not employed	8,418 (44.5%)	7,271 (38.4%)	7,133 (37.7%)	6,793 (35.9%)	6,498 (34.3%)	36,113 (38.1%)
Recruitment Centre						
England	16,979 (89.7%)	16,997 (89.7%)	17,079 (90.2%)	16,961 (89.5%)	16,950 (89.5%)	84,966 (89.7%)
Scotland	1,147 (6.1%)	1,220 (6.4%)	1,170 (6.2%)	1,270 (6.7%)	1,357 (7.2%)	6,164 (6.5%)
Wales	807 (4.3%)	722 (3.8%)	692 (3.7%)	711 (3.8%)	634 (3.3%)	3,566 (3.8%)
Behavioural factors						
Alcohol 3+ times/week, No. (%)	7,732 (40.8%)	8,918 (47.1%)	9,506 (50.2%)	9,988 (52.7%)	10,234 (54.0%)	46,378 (49.0%)
Never smoker, No. (%)	9,971 (52.7%)	10,827 (57.2%)	11,025 (58.2%)	11,214 (59.2%)	11,057 (58.4%)	54,094 (57.1%)
Non coffee drinker, No. (%)	4,168 (22.0%)	3,751 (19.8%)	3,611 (19.1%)	3,555 (18.8%)	3,516 (18.6%)	18,601 (19.6%)
Accelerometer-derived Daily Sleep Duration, mean (SD), (hours)	9.28 (1.53)	8.97 (1.22)	8.82 (1.13)	8.69 (1.09)	8.50 (1.04)	8.85 (1.24)
BMI, mean (SD), (kgm ⁻²)	28.6 (5.58)	27.1 (4.45)	26.5 (4.16)	25.9 (3.88)	25.5 (3.68)	26.7 (4.53)
Prior disease, No. (%)						
Type 2 Diabetes	1,533 (8.1%)	857 (4.5%)	611 (3.2%)	535 (2.8%)	474 (2.5%)	4,010 (4.2%)
Depression	7,376 (39.0%)	6,534 (34.5%)	6,217 (32.8%)	5,928 (31.3%)	5,487 (29.0%)	31,542 (33.3%)
Constipation	1,020 (5.4%)	707 (3.7%)	652 (3.4%)	627 (3.3%)	548 (2.9%)	3,554 (3.8%)
Bladder Dysfunction	177 (0.9%)	122 (0.6%)	115 (0.6%)	89 (0.5%)	80 (0.4%)	583 (0.6%)
Season of device wear, No. (%)						
Spring	4,113 (21.7%)	4,160 (22.0%)	4,432 (23.4%)	4,413 (23.3%)	4,604 (24.3%)	21,722 (22.9%)
Summer	4,334 (22.9%)	4,536 (24.0%)	4,989 (26.3%)	5,378 (28.4%)	5,897 (31.1%)	25,134 (26.5%)
Autumn	5,675 (30.0%)	5,830 (30.8%)	5,663 (29.9%)	5,459 (28.8%)	5,305 (28.0%)	27,932 (29.5%)
Winter	4,811 (25.4%)	4,413 (23.3%)	3,857 (20.4%)	3,692 (19.5%)	3,135 (16.6%)	19,908 (21.0%)

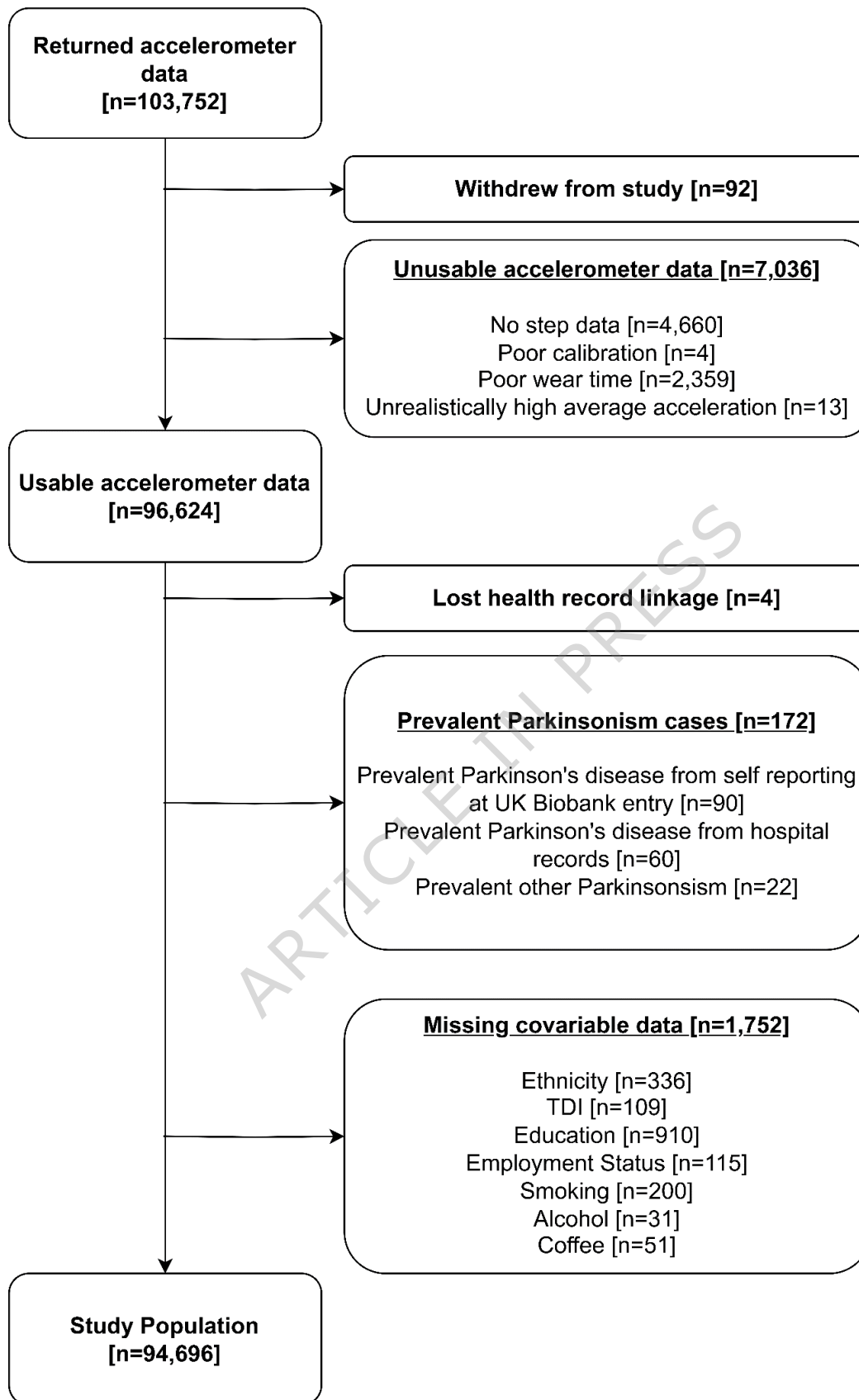


Figure 1: Exclusion flow diagram for analysis of UK Biobank Physical Activity Monitoring population.

TDI = Townsend deprivation index.

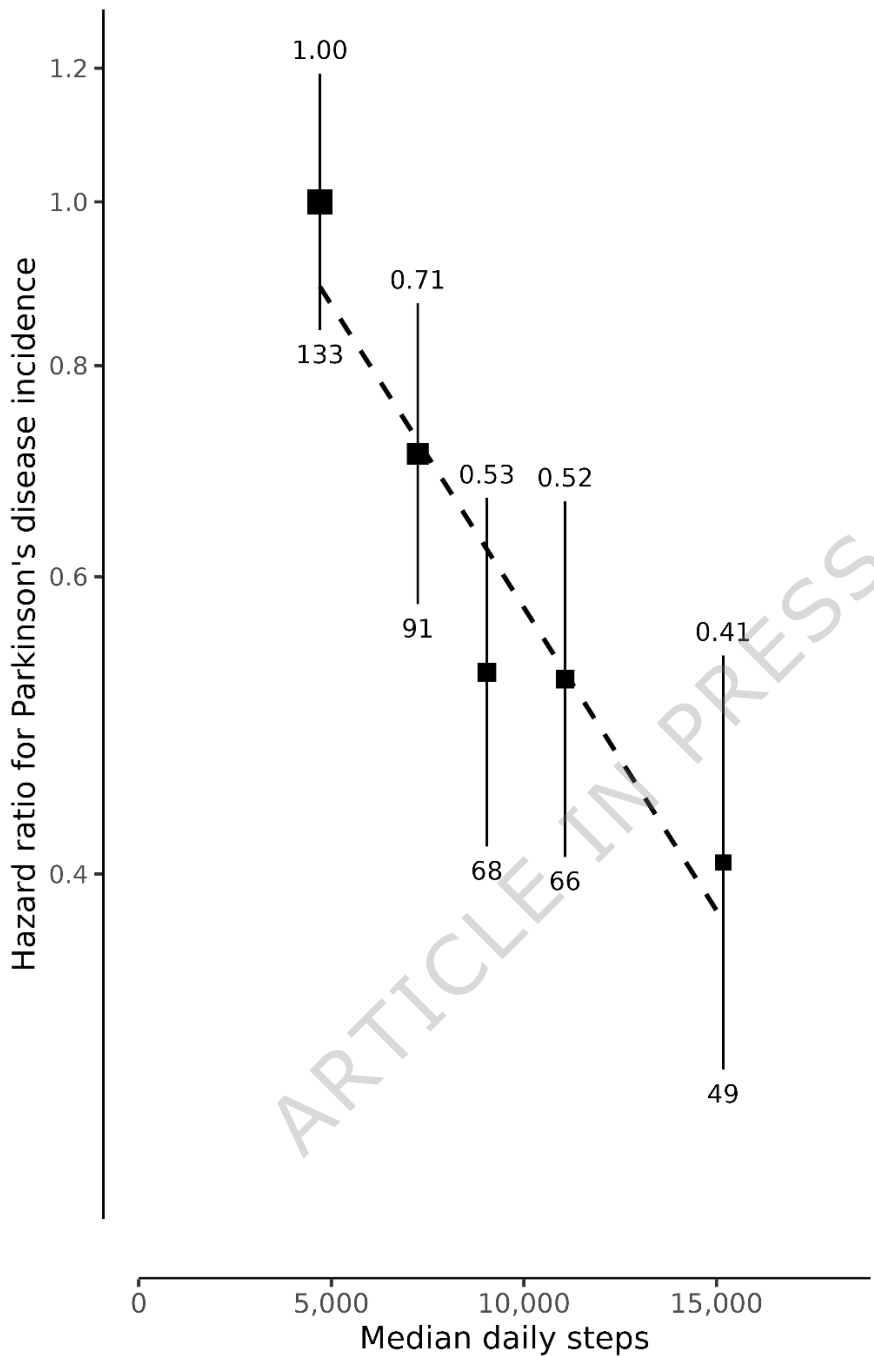


Figure 2: Association between quintiles of median daily step count and incident Parkinson's disease. Hazard ratios (HR) and 95% group-specific confidence intervals (CI) were calculated using age as a timescale, adjusted for season of wear, sex, ethnicity, Townsend deprivation index, geographic region, education, employment, alcohol intake, smoking status and coffee intake. HR is above, and the number of events is plotted below each data point. The HR (95% CI) per 1,000 additional daily steps is 0.92 (0.89 – 0.94).

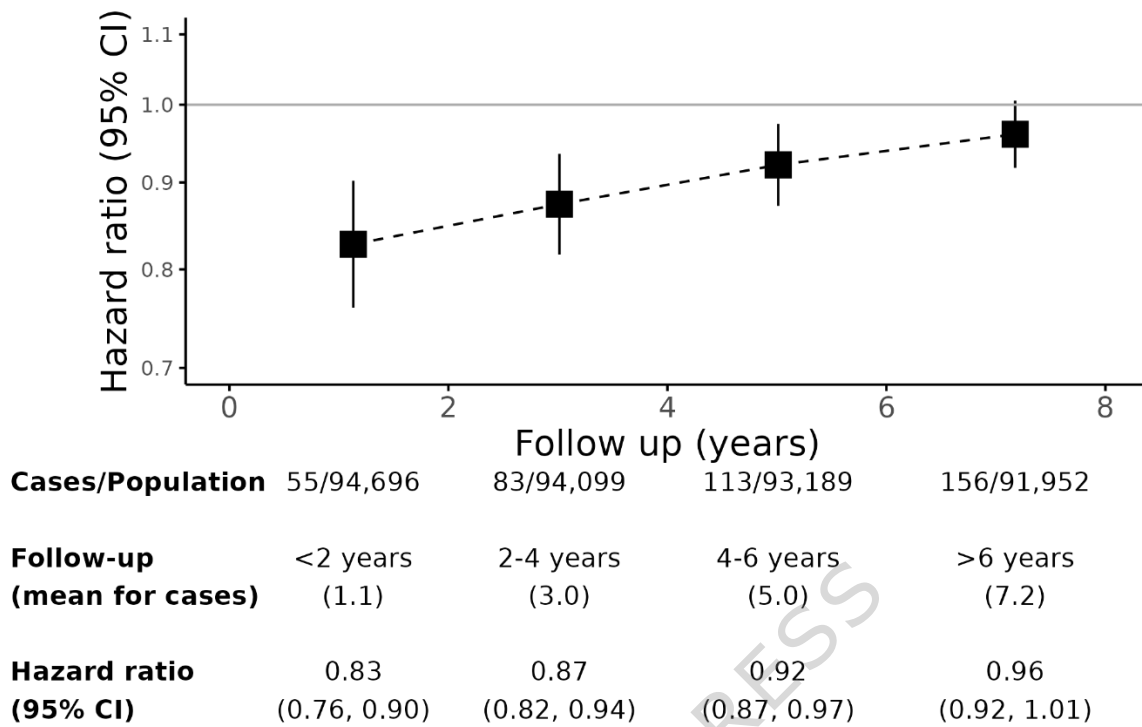


Figure 3: Hazard ratio for incident Parkinson's disease by period of follow-up per 1,000 median daily steps higher. Models use age as timescale, and are adjusted for season of wear, sex, Townsend deprivation index, geographic region of assessment centre, educational attainment, employment, smoking, alcohol and coffee consumption. CI=confidence interval.