nature portfolio

Peer Review File

Sustained physical activity offers benefits beyond activity volume in chronic disease prevention

Corresponding Author: Dr Mingyang Song

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you for the opportunity to review this interesting study. In this paper, the authors utilized HPFS, NHS and NHSII data to examine the association between long-term physical activity patterns and the incidence of major chronic diseases. I found the topic engaging, and the analyses appear generally sufficient. The findings have the potential to provide valuable insights for public health. However, I have several concerns that I believe should be addressed to strengthen the manuscript.

One major concern pertains to the generalizability of the study's findings, as the sample consists solely of health professionals—a population with substantially greater knowledge of disease prevention and heightened health awareness compared to the general public. This limits the extent to which the results can be extrapolated to broader populations. Additionally, the manuscript lacks sufficient detail regarding the minimum threshold of self-reported physical activity data required across multiple waves to define consistency in physical activity. Clarification on this point is essential to assess the robustness of the longitudinal analysis.

A second major concern involves the incomplete adjustment for key confounding variables. There is substantial evidence indicating that socioeconomic status is a significant determinant of major chronic conditions. The omission of socioeconomic indicators from the analytical models raises important questions regarding the validity and robustness of the reported associations, as these relationships may be confounded by unmeasured socioeconomic factors. The authors are encouraged to incorporate relevant covariates or, at minimum, provide a rationale for their exclusion.

The manuscript does not appear to follow any established reporting checklist. Given that this is an observational epidemiological study, adherence to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist would substantially enhance the transparency, completeness, and replicability of the research. The authors are strongly encouraged to revise the manuscript in accordance with STROBE guidelines.

Some minor comments below:
1. Line 270, Page 13: The numerical values presented in the text do not fully align with those shown in the flow chart (eFigure 1). Please ensure consistency between the narrative and visual representation of the study population.
2. Lack of Pre-Registered Protocol: The study does not include an online pre-registered analysis protocol. Pre-registration is an important step in enhancing transparency and reducing potential bias in observational research.
3. Covariate Selection (Page 15-16): No justification is provided for the selection of covariates included in the models. Key socioeconomic variables—such as educational attainment and income level—that are known to influence both physical activity and health outcomes appear to be omitted. The use of a directed acyclic graph (DAG) to guide confounder selection would strengthen the analytical approach.
4. Definition and Handling of Physical Activity Data: It remains unclear what the minimum requirement is for the number of waves of physical activity data needed to classify participants as consistently active. Additionally, the manuscript does not specify how missing data were handled in the analysis. Clarification on both points is necessary to assess the reliability of the exposure classification and the validity of the results.
5. Line 114, Page 6: The meaning of "range: 0-33%" is unclear, particularly as this figure does not appear in Table 2. Please clarify what this range refers to and ensure consistency between the text and tables.
6. Lines 122-135: The argument that "maintaining a volume of low physical activity level on average but consistently

meeting the recommendation was related to greater reduction in disease incidence" is not well-supported. The confidence intervals reported are wide and show substantial overlap, which undermines the strength of the conclusion. The authors should interpret these findings with greater caution and consider revising the language to reflect the statistical uncertainty.

Thank you again for allowing me to review this manuscript. I hope my comments are helpful in refining your work further.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
Authors have assessed the longitudinal associations of physical activity with different health outcomes. The dataset is interesting and the results show what we already know. Novelty needs to be better characterized. Also, statistical analysis requires further refinements. Some specific comments:
1. There are longitudinal studies, but these were not capture in the review done by authors, please include these references.
2. what is the novelty of the study? is it the trajectory analysis done? or the joint analysis? this needs to be clarified
3. how these findings will be able to inform nuances in the current guidelines? nevertheless, this evidence is too preliminary.
4. the validity of the trajectory analysis is somewhat weak. Normally these kind of analysis are exploratory and not definitive.
5. Cis are too narrow in the trajectory analysis - why?
6. Marginal structural models and other causal inference models - Target trial emulation are better suited to explore the research question, I believe.
Conclusion needs to be assessed within caveats of the study

(Remarks on code availability)
code does not seem to be available?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I appreciate the author's thoughtful and comprehensive responses to my comments. The revised manuscript demonstrates clear improvements in clarity and overall presentation.

My only concern pertains to the newly introduced Target Trial Emulation, which was not part of the originally registered protocol. Any deviation from the pre-specified protocol should be explicitly acknowledged in the manuscript, and a strong justification needs to be provided for why they were made.

Additionally, I noticed a minor inconsistency in the manuscript. On Line 291, Page 13, the manuscript states that 121,700 female nurses were included in the NHS cohort, whereas Supplementary Figure 1 reports a total of 121,701. Please ensure that these figures are consistent throughout the manuscript and supplementary materials.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
Authors have responded sufficiently to all my queries- no further comments

(Remarks on code availability)

<u>Response to Reviewer #1's comments</u>

Thank you for the opportunity to review this interesting study. In this paper, the authors utilized HPFS, NHS and NHSII data to examine the association between long-term physical activity patterns and the incidence of major chronic diseases. I found the topic engaging, and the analyses appear generally sufficient. The findings have the potential to provide valuable insights for public health. However, I have several concerns that I believe should be addressed to strengthen the manuscript.

We sincerely thank the reviewer for recognizing the rigor, thoroughness, and added knowledge of our study, as well as for all the helpful comments. We have revised the manuscript accordingly and provided point-by-point responses below.

One major concern pertains to the generalizability of the study's findings, as the sample consists solely of health professionals—a population with substantially greater knowledge of disease prevention and heightened health awareness compared to the general public. This limits the extent to which the results can be extrapolated to broader populations.

We agreed with the reviewer regarding generalizability and acknowledged this in the Discussion: '*Secondly, our participants are health professionals in the US, which limits the generalizability of findings.*'.

At the same time, we note that the relatively homogeneous occupation of this cohort reduces potential confounding by socioeconomic status compared to studies conducted in the general population. Furthermore, because occupational physical activity is largely similar and generally low across our study participants, our study design provides a unique strength in assessing the impact of leisure-time activity, as the minimal variation in occupational activity among participants helps reduce a potential source of bias that often arises in studies of the general population.

Additionally, the manuscript lacks sufficient detail regarding the minimum threshold of self-reported physical activity data required across multiple waves to define consistency in physical activity. Clarification on this point is essential to assess the robustness of the longitudinal analysis.

We thank the reviewer for this valuable comment and would like to provide additional technical details. Because participants' physical activity and other data were assessed at each questionnaire wave, we formatted the data into a counting-process structure, with one row per person-time, which is well-suited for survival analysis and accommodates time-varying exposures and covariates in a longitudinal setting. We then fit Cox proportional hazards models stratified by questionnaire wave (every two years) as a proxy for calendar time, providing finer control for time. Since the models were estimated by wave and the missing values were imputed by carrying forward the most recent non-missing data, it was unnecessary to restrict the analysis to participants with at least a certain number of waves. Furthermore, we defined consistency in two ways to assess robustness: one as the percentage of follow-up years meeting the recommended level (relative measure; Table 2) and the other as the cumulative number of years meeting the recommended level (absolute measure; Supplementary Table 4). The results were similar across both definitions.

We have clarified our approach in the Methods section.

> (Methods, Statistical Analysis, Page 17)
> Cox proportional-hazards models stratified by age (months), questionnaire cycle (two-year interval), and cohort (only in pooled analyses) were used with the counting process data structure (one row per person-questionnaire cycle) to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) for the risk of major chronic diseases. Missing values in one questionnaire cycle were imputed by carrying forward the most recent non-missing data.

A second major concern involves the incomplete adjustment for key confounding variables. There is substantial evidence indicating that socioeconomic status is a significant determinant of major chronic conditions. The omission of socioeconomic indicators from the analytical models raises important questions regarding the validity and robustness of the reported associations, as these relationships may be confounded by unmeasured socioeconomic factors. The authors are encouraged to incorporate relevant covariates or, at minimum, provide a rationale for their exclusion.

We have conducted additional adjustment of socioeconomic variables, and the results are essentially unchanged (please see the table below). This is expected since within each cohort (HPFS: male physicians; NHS and NHSII: female nurses), socioeconomic status is relatively homogenous. Participants share broadly similar levels of education, professional training, occupational stability, and social standing, which well reduces the confounding by socioeconomic status.

In the HPFS, direct measures of socioeconomic status were not collected. However, it is reasonable to assume that all participants hold MD or MD-equivalent degrees, given this is required for their profession, and physicians generally occupy high socioeconomic status in the US. There are two measures available in both NHS and NHSII cohorts: partner/spouse's education (<high school, high school, college, graduate school) and the participant's subjective social status, the latter assessed using the MacArthur Scale (a 10-rung ladder representing relative standing in terms of income, education, and occupation). We compared results with and without adjustment for these two variables in the nurse cohorts, and the findings were essentially unchanged, suggesting that residual confounding by socioeconomic status is unlikely to substantially affect our results.

Table. Hazard ratios and 95% confidence intervals for major chronic diseases incidence in women

	Consistency (percentage of follow-up years) of physical activity meeting the recommended level			
	Q2	Q3	Q4	Per 50% increment
Main model [a]	0.96 (0.94, 0.98)	0.88 (0.86, 0.91)	0.78 (0.76, 0.80)	0.88 (0.87, 0.89)
Main model w/ SES adjustment [b]	0.98 (0.96, 1.00)	0.90 (0.87, 0.92)	0.78 (0.77, 0.80)	0.89 (0.87, 0.90)
	Cumulative years meeting the recommended level			
	Q2	Q3	Q4	Per 10-y increment
Main model [a]	0.94 (0.92, 0.97)	0.87 (0.84, 0.89)	0.78 (0.76, 0.80)	0.88 (0.87, 0.90)
Main model w/ SES adjustment [b]	0.96 (0.93, 0.98)	0.88 (0.86, 0.91)	0.79 (0.77, 0.82)	0.89 (0.88, 0.90)
	Cumulative average physical activity volume, MET-hours/week			
	Q2	Q3	Q4	Per 15 MET-hr/wk
Main model [a]	0.91 (0.89, 0.93)	0.85 (0.83, 0.87)	0.79 (0.77, 0.81)	0.91 (0.90, 0.92)
Main model w/ SES adjustment [b]	0.92 (0.90, 0.94)	0.86 (0.84, 0.88)	0.80 (0.78, 0.82)	0.91 (0.90, 0.92)

[a] Main model is the Cox proportional hazards model stratified by age (months), questionnaire cycle (two-year interval), and cohort, adjusted for covariates including race, smoking status and packyears, alcohol consumption, Alternative Healthy Eating Index-2010, regular aspirin use, regular NSAIDs use, multivitamin use, family history of diabetes mellitus/cardiovascular disease/cancer, menopausal status and hormone use.

[b] Socioeconomic status (SES) was measured by partner/spouse's education and the participant's subjective social status in MacArthur Scale (a 10-rung ladder representing relative standing in terms of income, education, and occupation).

The manuscript does not appear to follow any established reporting checklist. Given that this is an observational epidemiological study, adherence to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist would substantially enhance the transparency, completeness, and replicability of the research. The authors are strongly encouraged to revise the manuscript in accordance with STROBE guidelines.

Thanks for pointing this out. We have provided STROBE Checklist (cohort study) in the revision.

Some minor comments below:

1. Line 270, Page 13: The numerical values presented in the text do not fully align with those shown in the flow chart (eFigure 1). Please ensure consistency between the narrative and visual representation of the study population.

We have verified the numerical values reported in the text ("A total of 45 426 men from the HPFS and 186 062 women from the NHS and NHSII were included") and confirmed that they are consistent with the number of participants remaining in the cohorts after exclusions, as shown in Supplementary Fig. 1. To avoid confusion with the number of participants before exclusions, we have clarified in the text accordingly.

> (Results, Population Characteristics, Page 4)
> A total of 45 426 men from the HPFS and 186 062 women from the NHS and NHS II were included after applying the exclusion criteria (Supplementary Fig. 1).

2. Lack of Pre-Registered Protocol: The study does not include an online pre-registered analysis protocol. Pre-registration is an important step in enhancing transparency and reducing potential bias in observational research.

We appreciate the reviewer's careful consideration. To ensure scientific rigor and reproducibility, the cohorts (HPFS, NHS, NHSII) have established a review process for every study using the data[1]. This process includes submission and presentation of an analysis proposal detailing the scientific background, study hypotheses, design, exposure definitions, outcomes, covariates, and a brief statistical analysis plan. Proposals are presented at biweekly cohort meetings to the research community for feedback and input. We presented our analysis proposal on February 15, 2024, and have provided it for the reviewer's reference.

3. Covariate Selection (Page 15-16): No justification is provided for the selection of covariates included in the models. Key socioeconomic variables—such as educational attainment and income level—that are known to influence both physical activity and health outcomes appear to be omitted. The use of a directed acyclic graph (DAG) to guide confounder selection would strengthen the analytical approach.

We selected confounders based on the standard definition—factors that are common causes of both physical activity and major chronic diseases (i.e., variables with direct arrows pointing to both physical activity and major chronic diseases on the DAG)—as well as data availability. As noted above, participants in the HPFS are all male physicians in the US, and participants in the NHS and NHSII are all female nurses in the US. In pooled analyses, we conducted Cox models stratified by cohort, and socioeconomic status is unlikely to vary substantially within cohort because participants within cohort share the same occupation and therefore broadly similar levels of education, professional training, occupational stability, and social standing. Accordingly, socioeconomic status is not expected to bias results and they are typically not adjusted as confounders in studies based on these cohorts.

We have clarified our approach to confounder selection in the Methods section and provided the rationale for not adjusting for socioeconomic status in the Discussion section.

> (Methods, Statistical Analysis, Page 17)
> We selected confounders based on the standard definition—factors that are common causes of both physical activity and major chronic diseases (i.e., variables with direct arrows pointing to both physical activity and major chronic diseases on the DAG)—as well as data availability.
>
> (Discussion, Page 13)
> Secondly, our participants are health professionals in the US, which may limit the generalizability of findings. However, the similar occupations of our study participants reduced variation in occupational physical activity and socioeconomic status, thus lowering potential confounding compared to studies in the general population.

4. Definition and Handling of Physical Activity Data: It remains unclear what the minimum requirement is for the number of waves of physical activity data needed to classify participants as consistently active. Additionally, the manuscript does not specify how missing data were handled in the analysis. Clarification on both points is necessary to assess the reliability of the exposure classification and the validity of the results.

Thanks for the comments. As responded to the point above, we formatted the data into a counting-process structure, with one row per person-time, and then fit Cox proportional hazards models stratified by questionnaire wave (every two years) as a proxy for calendar time, providing finer control for temporal trends. Since the models were estimated by wave and missing values were imputed by carrying forward the most recent non-missing data, it was unnecessary to restrict the analysis to participants with a minimum number of waves. Furthermore, we defined consistency in two ways to assess robustness: one as the percentage of follow-up years meeting the recommended level (relative measure; Table 2) and the other as the cumulative number of years meeting the recommended level (absolute measure; Supplementary Table 4). The results were similar across both definitions.

We have clarified our approach in the Methods section.

> (Methods, Statistical Analysis, Page 17)
> Cox proportional-hazards models stratified by age (months), questionnaire cycle (two-year interval), and cohort (only in pooled analyses) were used with the counting process data structure (one row per person-questionnaire cycle) to estimate the hazard ratios (HRs) and 95% confidence intervals (CIs) for the risk of major chronic diseases. Missing values in one questionnaire cycle were imputed by carrying forward the most recent non-missing data.

5. Line 114, Page 6: The meaning of "range: 0-33%" is unclear, particularly as this figure does not appear in Table 2. Please clarify what this range refers to and ensure consistency between the text and tables.

The range refers to the minimum and maximum values of the percentage of follow-up years in which physical activity met the recommended level within the lowest quartile. In Table 2, however, we presented the interquartile range (25^{th}–75^{th} percentile) rather than the minimum-maximum values. To avoid confusion, we have revised the text to ensure consistency with the table.

> (Results, Page 5)
> Compared with participants in the lowest quartile of long-term physical activity consistency (median: 0% of follow-up years meeting the recommended level of ≥ 7.5 MET-hours/week, interquartile range: 0–20%), …

6. Lines 122-135: The argument that "maintaining a volume of low physical activity level on average but consistently meeting the recommendation was related to greater reduction in disease

incidence" is not well-supported. The confidence intervals reported are wide and show substantial overlap, which undermines the strength of the conclusion. The authors should interpret these findings with greater caution and consider revising the language to reflect the statistical uncertainty.

Thanks for the helpful comment. We have revised the wording as per the reviewer suggested.

> (Results, Page 6)
> In general, maintaining a volume of less than 10 MET-hours/week on average and consistently meeting the recommended level of ≥ 7.5 MET-hours/week throughout the follow-up period (i.e., approximately 8–10 MET-hours/week on average for 100% of follow-up years), was related to a modestly lower major chronic diseases incidence than sporadic high-volume activity in some years mixed with inactivity in other years (i.e., > 23 MET-hours/week on average but less than 50% of follow-up years meeting the recommended level), with the HR (95% CI) of 0.84 (0.77, 0.91) and 0.88 (0.81, 0.97), respectively. Despite the overlapping confidence intervals, the differences in point estimates suggest that consistently medium-volume activity may provide slightly greater benefit than sporadic high-volume activity.

Response to Reviewer #2's comments

Authors have assessed the longitudinal associations of physical activity with different health outcomes. The dataset is interesting and the results show what we already know. Novelty needs to be better characterized. Also, statistical analysis requires further refinements. Some specific comments:

We appreciate the reviewer's insightful comments that have helped us improve the manuscript quality. We have revised the manuscript accordingly and provided point-by-point responses below.

1. There are longitudinal studies, but these were not capture in the review done by authors, please include these references.

In response to the helpful suggestion, we have re-conducted literature search and added several references with physical activity measurements at more than one time point in the Introduction section to more clearly situate our study within the broader body of longitudinal research.

> (Introduction, Page 3-4)
> Although numerous studies have followed participants over time, they typically rely on a single baseline assessment of physical activity and therefore cannot capture long-term patterns or changes across decades. Studies with repeated assessments are often limited by few measurement time points, short time periods, relatively small sample sizes, restriction to adolescents or older adults, and a focus on describing trajectories without assessing the impact on long-term health outcomes[2-7]. Two recent studies in the general population modeled physical activity trajectories, including one across four age periods from adolescence to midlife and the other across four questionnaire waves over 10 years, and examined their associations with mortality[8,9]. They indicated the importance of remaining physically active and preventing declines in midlife to prevent premature deaths. However, these studies were still limited by a small number of repeated measurements and focused on general trends of maintenance, increase, or decrease without clarifying the relative contributions of volume versus consistency. Few studies have investigated whether physical activity requires consistent adherence to the recommended level, or whether sporadic high activity interspersed with inactivity can confer sustained health benefits.
>
> In this work, leveraging the physical activity data obtained through repeated assessments for 32 years in three large US prospective cohorts (Health Professionals Follow-up Study (HPFS), Nurses' Health Study (NHS) and NHS II), we systematically examined different long-term physical activity patterns during adulthood in relation to the incidence of major chronic diseases. Our findings provide critical insights into the relative importance of pursuing high-volume activity versus maintaining consistent yet medium-volume activity over decades.

2. What is the novelty of the study? is it the trajectory analysis done? or the joint analysis? this needs to be clarified

We thank the reviewer for raising this important point. The novelty and major additions of this study are summarized as follows.

Focus on long-term patterns rather than a snapshot of physical activity

While numerous studies have established the health benefits of physical activity, most have focused on baseline activity volume, different intensities, or short-term behaviors (e.g., timing within a day, sessions per week). Such approaches provide valuable evidence but utilize only a snapshot of lifestyle. In contrast, chronic diseases develop over prolonged periods, and physical activity fluctuates across the life course. It is therefore critical to understand whether sustained engagement is necessary for risk reduction, or whether sporadic high activity interspersed with inactivity may confer comparable protection.

Address key barriers to analyze long-term physical activity patterns

Studying long-term patterns of physical activity is challenging because of both (a) limited availability of repeated assessments in large cohorts and (b) the statistical complexity as physical activity levels can change continuously over several decades. To address these challenges, we utilized cohorts with repeated physical activity data over 30 years and applied a joint analysis framework that integrates cumulative average volume with consistency of meeting recommended activity levels. This approach provides a practical way to capture both independent and combined effects of volume and consistency. We further complemented this analysis with trajectory and lagged analyses to corroborate and extend the insights.

Narrow existing knowledge gaps by providing novel insights

- Based on the joint analysis, we found that greater consistency in meeting recommended levels was associated with lower disease risk within each tertile of cumulative average volume, and vice versa. This narrows an evidence gap regarding long-term maintenance of physical activity.
- Based on the joint analysis, we suggested that maintaining activity just above the recommended level is related to slightly greater benefit than sporadic high activity in some years mixed with inactivity in other years for CVD and cancer, whereas for type 2 diabetes, volume is more critical than consistency.
- Together with the trajectory analysis and lagged analysis, the overall body of findings provide new evidence that refines current physical activity guidelines by clarifying the relative importance of sustained consistency versus average volume over decades.
- Please also refer to the table summary below in responses to the next comment.

We have clarified these points in the revised manuscript to make the novelty of our study explicit.

(Introduction, Page 3) Although numerous studies have followed participants over time, they typically rely on a single baseline assessment of physical activity, …

(Discussion, Paragraph 2, Page 9-10) Studying the impact of long-term physical activity patterns on chronic disease presents challenges…

3. How these findings will be able to inform nuances in the current guidelines? nevertheless, this evidence is too preliminary.

Thanks for the comment, which allows us to further clarify the public health relevance of our findings. Current physical activity guidelines emphasize total volume and type of activity but do not address the role of consistency or long-term patterns, in large part because there has been little evidence from human cohorts due to challenges in collecting and analyzing repeated physical activity measures over decades. Our study systematically investigated long-term patterns of physical activity in relation to major chronic diseases, thereby helping to narrow this evidence gap.

While our findings may be viewed as preliminary, they provide a foundation for understanding whether sustained activity offers benefits beyond total volume alone. By extending the time horizon beyond a single snapshot of activity, our results suggest new dimensions of physical activity behavior that could eventually inform refinements to guidelines. To make this clearer, we have added a summary table showing how different analyses converge on consistent findings.

Key findings	Implications for the guideline and future research
Joint analysis (Figure 1): Greater consistency in meeting recommended levels was associated with lower disease risk within each tertile of cumulative average volume, and vice versa. Joint analysis (Figure 1): Maintaining activity just above the recommended level (7.5 MET-hours/week) was related to slightly greater risk reduction for major chronic diseases than engaging in sporadic high activity (triple the recommended level on average) interspersed with inactivity for over half of follow-up years. Trajectory analysis (Table 3): The medium-to-high-volume trajectory between age 40–60 was associated with greater benefit for major chronic disease prevention after age 60 than the high-to-medium-volume trajectory.	Guidelines should incorporate a component on sustained physical activity, emphasizing the importance of long-term consistency in addition to total volume. Even moderate but consistent physical activity may confer benefits comparable to, or greater than, sporadic high activity, highlighting the value of promoting sustained activity.

Joint analysis (Figure 1): For CVD and cancer, maintaining activity just above the recommended level (7.5 MET-hours/week) was related to slightly greater benefit than engaging in sporadic high activity (triple the recommended level on average) interspersed with inactivity for over half of follow-up years. However, for T2D, higher average volume appeared to be more important than consistency (HR (95% CI): 0.90 (0.79, 1.04) for consistently medium-volume activity vs. 0.76 (0.65, 0.89) for sporadic high-volume activity). Trajectory analysis (Table 3): The benefits of high activity volume were particularly pronounced for T2D (HR 0.74 (0.68, 0.81) for consistently medium-volume group vs HR 0.42 (0.35, 0.51) for consistently high-volume group). For T2D risk after age 60, the medium-to-high-volume trajectory in age 40-60 was associated with greater benefit than high-to-medium-volume trajectory, indicating the benefit decayed over time. However, for CVD and cancer risk after age 60, the medium-to-high-volume trajectory was associated with similar benefit as high-to-medium-volume trajectory.	Type 2 diabetes: High average volume is critical. Lower activity earlier in life might be offset by adopting high activity in late midlife, before the common age of T2D onset, suggesting flexibility in timing but a need to engage in high levels on average. CVD and cancer: Consistency matters more than high average volume. Sustained activity, even at moderate levels, is suggested to be more protective than sporadic bursts of high activity. Guideline relevance: Findings highlight the value of incorporating both volume and consistency into recommendations, with disease-specific nuances: emphasize long-term maintenance for CVD and cancer, and high volume (at least in later middle adulthood) for T2D prevention.

We have revised the Discussion section to explicitly highlight the added knowledge and how it may guide future research toward more nuanced recommendations.

(Discussion Page 11-12) Furthermore, we observed a variation in the relative importance of consistency and volume across individual diseases…

(Discussion, Paragraph 5, Page 12) Current physical activity guidelines emphasize total volume and type of activity but do not address the role of consistency or long-term patterns. By extending the time horizon beyond a single snapshot of activity, …

4. The validity of the trajectory analysis is somewhat weak. Normally these kinds of analysis are exploratory and not definitive.

Thanks for the careful consideration. We agree that trajectory analyses should be interpreted with caution. In our study, the primary evidence comes from the analyses of consistency, volume, and their joint effects, with the trajectory analysis serving as a secondary approach. To increase its validity, we used established group-based modeling approaches, selected the number of trajectories and the functional forms based on model fit criteria, and examined whether the identified groups were clinically interpretable. The results of the trajectory analysis were broadly in line with our joint analysis (please refer to the table above in response to comment #3), lending additional support to the robustness of our findings.

We have clarified this point in the Discussion by noting that the trajectory analysis should be viewed as exploratory but nevertheless provides insights into long-term patterns that are difficult to capture otherwise.

> (Discussion, Page 9-10)
> Most of the few longitudinal studies on physical activity have relied on trajectory analyses; however, such approaches are generally considered exploratory, as the validity of the identified groups depends heavily on modeling assumptions and may not reflect distinct underlying subpopulations[7-9]. To address these limitations, we utilized cohorts with repeated physical activity data over 30 years and applied a joint analysis framework that integrates cumulative average volume with consistency of meeting recommended activity levels. This approach provides a practical way to capture both independent and combined effects of volume and consistency. We further complemented this analysis with trajectory and lagged analyses, as well as target trial emulation to corroborate and extend the insights.

5. CIs are too narrow in the trajectory analysis - why?

Thanks for catching that. We suppose the reviewer is referring to Figure 2. The specific values for the mean and confidence intervals were provided in Supplementary Table 3. To ensure the trajectory of being consistently inactive was captured, we first identified a "consistently inactive" group as participants with physical activity < 7.5 MET-hours/week at all six age points. For this group, the mean was computed as the sample mean, and the 95% CI was obtained using SAS Proc Means under a t-distribution assumption. Given the large sample size in our cohorts, the standard error is small, which naturally results in narrow CIs.

For the other trajectory groups, both the mean and 95% CI were estimated using SAS Proc Traj as the estimated mean trajectory curve at each time point, with values stored in the OUTPLOT file. While the estimation of means and CIs from discrete mixture model is complicated, the narrow CIs in these groups are largely due to the large underlying sample size. Thus, the width of the CIs is consistent with expectations in large cohorts and reflects high statistical precision.

We have added technical details into the legend of Figure 2 - *Shaded bands represent 95% confidence intervals, which appear narrow due to the large sample size. For the "consistently inactive" group, means and CIs were calculated using SAS PROC MEANS, while for other groups they were derived from model-predicted values in SAS PROC TRAJ.*

6. Marginal structural models and other causal inference models - Target trial emulation are better suited to explore the research question, I believe. Conclusion needs to be assessed within caveats of the study.

We appreciate the reviewer's insightful comment. Target trial emulation provides a powerful way to quantify the effect of different courses of actions defined by researchers in eligible individuals. In response, we have incorporated target trial emulation into the revised analyses. Specifically, we emulated intervention strategies contrasting consistently maintaining physical activity above 7.5, 15, 30 MET-hours/week versus below 7.5 MET-hours/week. We believe the current study unpacking consistency, volume, and trajectory offers useful insights into long-term physical activity intervention strategy definitions and lay the groundwork for more complex trial emulations in future research.

At the same time, we recognize important limitations of target trial emulation. Its validity depends on several strong assumptions and analytic choices, including the correct specification of multiple parametric models to simulate covariates at each time interval and appropriate representation of causal pathways linking exposure, covariates, censoring, and outcomes. Like all observational approaches, it remains subject to unmeasured or residual confounding. In addition, the selection of participants into the analytic sample may introduce bias, as more stringent exclusion criteria are required—not only for missing baseline exposure but also for incomplete covariate information. These caveats underscore the importance of interpreting findings from target trial emulation cautiously and viewing them as complementary to other analytic approaches rather than definitive on their own.

We have added target trial emulation analyses and discussions in the manuscript and supplemental materials.

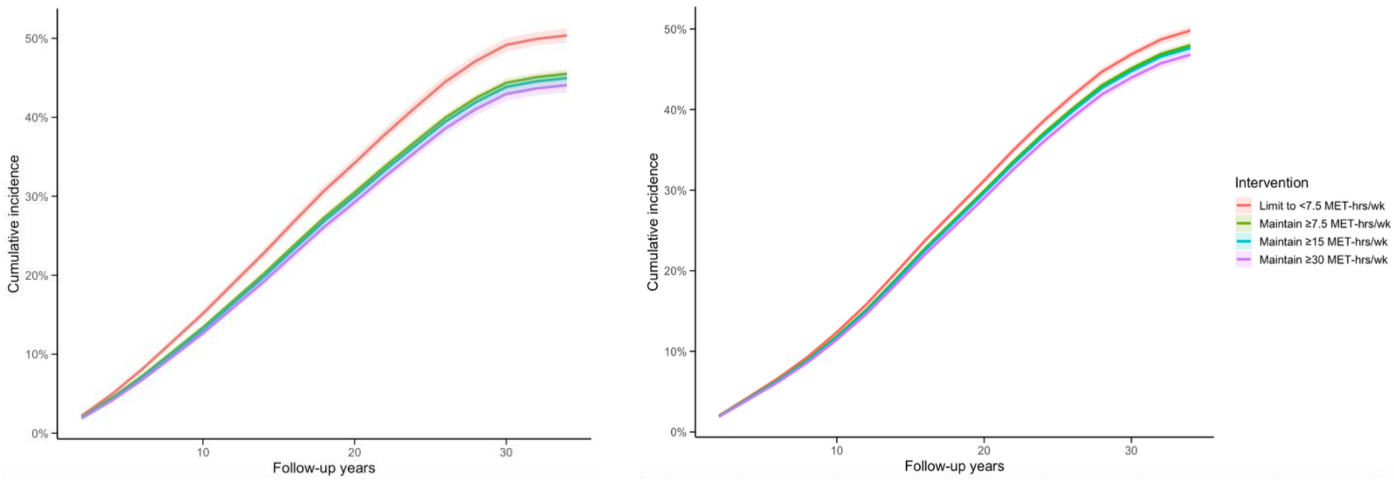
> (Methods, Page 19)
> We also employed target trial emulation as an approach to emulate a target trial (i.e., a hypothetical pragmatic trial that would have answered the causal question of interest) of sustained physical activity interventions over long periods for the prevention of major chronic diseases using observational data when randomized trials are unfeasible[10,11]. The key protocol components of the target trial are summarized in **Supplementary Table 8**. We used the parametric g-formula, a generalization of standardization to time-varying treatments and confounders, to estimate 32-year risk of major chronic diseases risks under each physical activity interventions[12,13].
>
> (Results, Page 8)
> In the target trial emulation analysis, we estimated 32-year risk of major chronic diseases under sustained physical activity strategies, adjusting for baseline and time-varying variables using the parametric g-formula. Compared with limiting physical activity to less than 7.5 MET-hours/week (32-year risk (95% CI): 50.3% (49.5%, 51.3%) in HPFS, 49.3% (48.5%, 50.2%) in NHS), interventions of maintaining ≥ 7.5 MET-hours/week of

physical activity reduced the absolute risk by 4.9% (4.1%, 5.8%) in HPFS and 2.3% (1.8%, 3.0%) in NHS (**Supplementary Fig. 5**).

Protocol component	Target trial specification	Target trial emulation
Eligibility criteria	No history of cancer, CVD, or T2D at baseline	Same as the target trial. We also require complete baseline data on physical activity, birth date, diet (missing FFQ items < 70, plausible energy intake reported [800 to 4200 kcal/d in males; 500 to 3500 kcal/d in females]) and other risk factors.
Physical activity intervention	Each individual is assigned to one of the following strategies at baseline: [1] Maintain ≥ 7.5 MET-hours/week of physical activity [2] Maintain ≥ 15 MET-hours/week of physical activity [3] Maintain ≥ 30 MET-hours/week of physical activity [4] Limit physical activity to < 7.5 MET-hours/week	Same as the target trial.
Strategy assignment	Individuals are randomly assigned to a strategy at baseline and are aware of their assigned strategy.	We assumed randomization conditional on the following baseline covariates: age, race, smoking history, overweight/obesity status, alcohol consumption, alternative healthy eating index-2010, regular aspirin use, regular NSAIDs use, multivitamin use, family history of diabetes mellitus, cardiovascular disease, or cancer; for females, also menopausal status and hormone use.
Outcome	Major chronic disease, defined as the first occurrence of T2D, major CVD (coronary heart disease and stroke), or total cancer (excluding nonmelanoma skin cancer and nonfatal prostate cancer).	Same as the target trial.
Follow-up	For each eligible individual, follow-up starts at strategy assignment (baseline) and ends upon loss to follow-up, death, the outcome of interest, or cutoff date, whichever happens first.	Same as the target trial. Loss to follow-up was defined as questionnaire nonresponse.
Statistical analysis	Intention to treat analysis. Per protocol analysis: Parametric g-formula to compare 32-year risk under each strategy with adjustment for pre- and postbaseline prognostic factors associated with adherence and loss to follow-up.	Same as the per-protocol analysis of the target trial.

Supplementary Fig. 5. Estimated 32-year risk of major chronic diseases under physical activity interventions in the Health Professionals Follow-up Study and Nurses' Health Study (1986-2018)

(A) Health Professionals Follow-up Study (B) Nurses' Health Study



	32-year risk [a], %	
Intervention strategy	HPFS	NHS
Limit to <7.5 MET-hours/week	50.3 (49.5, 51.3)	49.3 (48.5, 50.2)
Maintain ≥ 7.5 MET-hours/week	45.4 (44.7, 46.1)	46.9 (46.3, 47.5)
Maintain ≥ 15 MET-hours/week	45.0 (44.1, 45.6)	46.5 (45.7, 47.2)
Maintain ≥ 30 MET-hours/week	44.0 (43.1, 44.8)	45.7 (44.9, 46.5)

[a] Estimates were based on the parametric g-formula, adjusted for age, race, smoking packyears, alcohol consumption, body mass index, Alternative Healthy Eating Index-2010, regular aspirin use, regular NSAIDs use, multivitamin use, family history of diabetes mellitus, cardiovascular disease, or cancer; for females, also menopausal status and hormone use. A nonparametric bootstrapping procedure with 200 iterations was used to calculate percentile-based 95% confidence intervals for all estimates.

7. Code does not seem to be available?

The link to our codes on GitHub is https://github.com/gigifang925/PAPattern_ChronicDisease/tree/main

As part of the four-stage review process required by the cohorts (HPFS, NHS, NHSII) to ensure scientific rigor and reproducibility[1], our codes was reviewed by a dedicated expert cohort programmer prior to manuscript submission.

We initially chose uploading the code to the Code Ocean platform. However, after logging in and reviewing the tutorials, we found that SAS is not supported due to its proprietary nature. In addition, accessing the cohort data and running the analyses require approval through an application process and must be conducted on the Channing Division of Network Medicine's servers because of data privacy requirements and the large data size. As a result, Code Ocean was not a suitable option. We opted for an alternative by uploading the code to GitHub and submitting the link.

Reference

1 Stopsack, K. H. *et al.* Promoting Reproducibility and Integrity in Observational Research: One Approach of an Epidemiology Research Community. *Epidemiology* **34**, 389-395 (2023). https://doi.org/10.1097/ede.0000000000001599

2 Husøy, A. *et al.* Longitudinal changes in device-measured physical activity from childhood to young adulthood: the PANCS follow-up study. *International Journal of Behavioral Nutrition and Physical Activity* **21**, 29 (2024). https://doi.org/10.1186/s12966-024-01578-7

3 Corder, K. *et al.* Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. *British Journal of Sports Medicine* **53**, 496-503 (2019). https://doi.org/10.1136/bjsports-2016-097330

4 Nilsson, M., Ohlsson, C., Odén, A., Mellström, D. & Lorentzon, M. Increased physical activity is associated with enhanced development of peak bone mass in men: A five-year longitudinal study*. *Journal of Bone and Mineral Research* **27**, 1206-1214 (2012). https://doi.org/10.1002/jbmr.1549

5 Zhang, X. *et al.* Longitudinal association between physical activity and health-related quality of life among community-dwelling older adults: a longitudinal study of Urban Health Centres Europe (UHCE). *BMC Geriatrics* **21**, 521 (2021). https://doi.org/10.1186/s12877-021-02452-y

6 Wanigatunga, A. A. *et al.* Daily Physical Activity Patterns as a Window on Cognitive Diagnosis in the Baltimore Longitudinal Study of Aging (BLSA). *Journal of Alzheimer's Disease* **88**, 459-469 (2022). https://doi.org/10.3233/jad-215544

7 Lounassalo, I. *et al.* Distinct trajectories of physical activity and related factors during the life course in the general population: a systematic review. *BMC Public Health* **19**, 271 (2019). https://doi.org/10.1186/s12889-019-6513-y

8 Saint-Maurice, P. F. *et al.* Association of Leisure-Time Physical Activity Across the Adult Life Course With All-Cause and Cause-Specific Mortality. *JAMA Network Open* **2**, e190355-e190355 (2019). https://doi.org/10.1001/jamanetworkopen.2019.0355

9 Mok, A., Khaw, K.-T., Luben, R., Wareham, N. & Brage, S. Physical activity trajectories and mortality: population based cohort study. *BMJ* **365**, l2323 (2019). https://doi.org/10.1136/bmj.l2323

10 Hernán, M. A. & Robins, J. M. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* **183**, 758-764 (2016). https://doi.org/10.1093/aje/kwv254

11 Young, J. G., Stensrud, M. J., Tchetgen Tchetgen, E. J. & Hernán, M. A. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med* **39**, 1199-1236 (2020). https://doi.org/10.1002/sim.8471

12 Robins, J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical modelling* **7**, 1393-1512 (1986).

13 Taubman, S. L., Robins, J. M., Mittleman, M. A. & Hernán, M. A. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *International journal of epidemiology* **38**, 1599-1611 (2009).

Response to Reviewer #1's comments

I appreciate the author's thoughtful and comprehensive responses to my comments. The revised manuscript demonstrates clear improvements in clarity and overall presentation.

We are grateful for the reviewer's helpful and thoughtful feedback, which has enhanced the rigor and clarity of our work.

1. My only concern pertains to the newly introduced Target Trial Emulation, which was not part of the originally registered protocol. Any deviation from the pre-specified protocol should be explicitly acknowledged in the manuscript, and a strong justification needs to be provided for why they were made.

We appreciate this point. The target trial emulation was conducted in response to comments from another reviewer. We have uploaded the study proposal to the project GitHub page, described the internal registration and presentation under the cohort policy, and disclosed the deviations from the study proposal at the end of Methods section.

(Methods, Page 20)

To ensure scientific rigor and reproducibility, the cohorts (HPFS, NHS, NHS II) have established a review process for every study using the data. This process includes submission and presentation of an analysis proposal detailing the scientific background, study hypotheses, design, exposure definitions, outcomes, covariates, and a brief statistical analysis plan. Proposals are presented at biweekly cohort meetings to the research community for feedback and input. We presented our study proposal on February 15, 2024, which can be found at https://github.com/gigifang925/PAPattern_ChronicDisease/tree/main. Compared to the planned analyses described in the study proposal, we modified the definition of physical activity consistency and reported results based on the original definition in Supplementary Table 4 as a sensitivity analysis. We additionally incorporated trajectory analysis to strengthen the validity of findings. In response to the reviewer's request, we conducted a target trial emulation, an advanced causal inference approach by emulating a hypothetical pragmatic trial that would have answered the causal question of interest, to estimate the causal effect of sustained physical activity interventions on risk of major chronic diseases using observational data.

2. Additionally, I noticed a minor inconsistency in the manuscript. On Line 291, Page 13, the manuscript states that 121,700 female nurses were included in the NHS cohort, whereas Supplementary Figure 1 reports a total of 121,701. Please ensure that these figures are consistent throughout the manuscript and supplementary materials.

Thanks for catching this. We have corrected the typo in Supplementary Figure 1 to ensure consistency with the main article.

Response to Reviewer #2's comments

Authors have responded sufficiently to all my queries- no further comments.

We sincerely appreciate the reviewer's valuable comments, which have improved the quality of the manuscript.