**ARIC Manuscript Proposal #4117**

| | | |
|---|---|---|
| **PC Reviewed: 9/13/22** | **Status: _____** | **Priority: 2** |
| **SC Reviewed: _____** | **Status: _____** | **Priority: ____** |

**1.a. Full Title**: The effect of mid-life physical activity on frailty-associated plasma proteins: a trial emulation approach

**b. Abbreviated Title (Length 26 characters)**: Physical Activity and Frailty

**2. Writing Group**:
Writing group members: Fangyu Liu, John Jackson, Keenan Walker, Jeremy Walston, Rasika Mathias, Michael Griswold, Priya Palta, Melinda Power, Josef Coresh, Beverly Gwen Windham, Jennifer Schrack (others welcome; order of authorships is to be determined)

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _____ **[please confirm with your initials electronically or in writing]**

**First author**: **Fangyu Liu**
Address: 2024 E Monument Street, Suite 2-700,
Baltimore, MD 21287
E-mail: fliu28@jhu.edu

**ARIC author** to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).
Name: **Jennifer Schrack**
Address: Department of Epidemiology,
Johns Hopkins Bloomberg School of Public Health,
615 N Wolfe St, E7144,
Baltimore, MD 21205
E-mail: jschrack1@jhu.edu

**3. Timeline**:
This proposal is for dissertation research. We aim to complete analyses and manuscript by April 2024.

**4. Rationale**:

A few trials investigated whether physical activity, mostly through structured exercise programs, can reduce frailty score[1–8] and incidence of frailty[9–11]. Though encouraging results were found in most trials, one challenge is that physical activity was a component of the frailty outcome

measure in some studies but these studies did not examine whether the effect was limited to that one component.[3,11,12] Perhaps for this reason, some trials reported null findings.[8,10] Understanding whether and how physical activity alters the underlying pathophysiology of frailty will enhance our knowledge on physical activity as a strategy to prevent and treat frailty. Physical activity is believed to simultaneously upregulate many systems, such as metabolism, inflammation, and mitochondrial function.[13,14] However, most studies of the molecular benefits of physical activity to date have mainly focused on individual systems rather than comprehensive panels of markers that may better determine whether physical activity affects multiple systems simultaneously.[14] The studies also focused on structured exercise programs that target one or more physical fitness components (e.g., aerobic training for cardiorespiratory endurance, resistance training for muscle strength, and/or balance training).[14,15] Physical activity in daily living usually consists of a mixture of these types of activities, performed in a less structured manner in terms of intensity, duration, and frequency. Therefore, physical activity performed in daily living may not elicit the same biological benefits as those structured programs, yet may be more acceptable to many older adults. Indeed, one trial randomized participants to receive an accelerometer and encouraged participants to increase their steps/day based on the feedback from the device. Though the frailty score decreased significantly post intervention, the score did not differ between the intervention and control groups.[8] Therefore, much more remains to be learned about the effect of physical activity performed in daily living on frailty and its underlying biological mechanisms.

Two recent studies examined the effects of physical activity in daily living on a large number of proteins.[16,17] Ubaida-Mohien and colleagues found that 1019 proteins from a muscle biopsy were associated with higher self-reported moderate-to-vigorous physical activity (MVPA) among 58 healthy adults aged 20-87 years. These proteins were involved in mitochondria biogenesis and function, immune response, proinflammatory response, DNA damage, and senescence.[16] These proteins also differed between older and younger adults in the same sample,[18] but their relationships with measures of muscle quality such as muscle mass and strength, and with frailty were not examined. Stattin and colleagues took a more targeted proteomic approach and found MVPA affected 75 out of 184 plasma proteins specifically related to cardiovascular health.[17] Although cardiovascular disease can increase the risk of frailty and vice versa, these two conditions are overlapping yet distinct from each other.[19] Taken together, these two studies[16,17] only provide limited knowledge about PA-induced benefits on frailty.

In a previous proposal (#3961), we proposed to examine the proteins measured in mid-life that may have a causal relationship with frailty in later life and therefore potential targets for early preventions of frailty. The objective of current proposal builds upon this ongoing effort to investigate whether exercise in daily living during mid-life can influence the expressions of these proteins. Specifically, we will estimate the effect of maintaining the CDC recommended 150 minutes/week of MVPA on levels of proteins that may cause frailty among primarily middle-aged adults by emulating a trial, where middle-aged adults are randomized to receive the intervention upon enrollment into the Atherosclerosis Risk in Community Study (ARIC). The per-protocol effect of long-term adherence to this intervention on the expressions of causal proteins identified in MP #3961 in mid-life (Visit 1 and 3) will be estimated by inverse probability weighting (IPW) method and g-formula via iterated conditional expectations (ICE)[20].

## 5. Main Hypothesis/Study Questions:

Aim: To examine the effect of long-term adherence to the CDC recommended level of MVPA in mid-life on frailty-associated proteins

Hypothesis: Adhering to the recommended MVPA is causally associated with protective expressions of frailty-associated proteins

## 6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study design: An emulation of target trial using longitudinal observational data from the ARIC. Approximately 12,000 participants had self-reported MVPA at Visit 1 and 8,200 participants at Visit 3. Under the trial emulation framework, the protocols of the target trial are specified as if such a trial is feasible and ethical, and how observational data can mimic this trial is described. This framework allows critical evaluation of the observational studies.[21] The table outlines the components of the target trial we will emulate using observational data from ARIC.

| Component | Target Trial | Emulation using observational data |
|---|---|---|
| Eligibility criteria | Adults aged 45-65 years who are free of functional limitations that prevent them from engaging in physical activity. | Same.<br><br>Whether a participant can engage in physical activity was not directly measured in ARIC. We will approximate this using the functional status[a] measured during annual follow-up since 1993 (around Visit 3). However, this measurement was not available around Visit 1. It can be argued that functional limitation preventing physical activity is rare at Visit 1 due to the relatively young age.<br><br>Additional exclusion criteria specific to ARIC are participants who were neither Black nor white, and non-whites in Washington Co. and Minneapolis. Due to small number of participants of these races, we may not have sufficient overlap of other confounders in the |

| | | |
|---|---|---|
| | | exchangeability assumption or there is not enough variability in physical activity. Therefore, the causal effect cannot be reliably estimated for these race groups. In a sensitivity analysis, we will include all races. |
| Intervention strategy | Intervention: achieving and maintaining the CDC recommended 150 minutes/week of MVPA.[22]<br>Control: performing MVPA as the participants choose. Some individuals may have ≥150 minutes/week of MVPA. | The long-term adherence to the recommendation is approximated by two measurements of time spent in MVPA per week at Visits 1 and 3 (See Measurement of MVPA for details). |
| Treatment assignment | At the enrollment into the study (Visit 1). | Same. |
| Follow-up | Starts at enrollment and ends at 6 years after enrollment, death, or loss to follow-up, whichever occurs first. | Same, except enrollment is taken as the date of Visit 1 (1987-1989) and 6-year follow-up is taken as Visit 3 (1993-1995). The exact follow-up time is approximately 6 years but may vary among participants. |
| Outcome | Frailty-associated plasma protein levels at the end of 6-year follow-up. | Same, except that the exact follow-up time may vary around 6 years.<br><br>Frailty-associated plasma proteins at Visit 3 will be identified from the results of Manuscript Proposal #3961. |
| Causal contrast of interest | Per-protocol effect, i.e., the effect of adhering to the assigned intervention/control during the 6-years of follow-up. | Same |
| Statistical analysis | Per-protocol effect estimated via comparison between the same two groups with adjustment for pre- and post-baseline prognostic factors associated with adherence to the intervention and loss to follow-up including death. | Same, specifically g-formula via iterated conditional expectations (ICE) and marginal structural model (MSM) by inverse probability weighting (IPW) methods. |

Note. [a] Functional status was assessed based on 4 questions on the ability to do usual activities, walk up and down stairs without help, do heavy housework, and walk half a mile without help. Having difficulty with any of these items indicate functional limitation.[23]

Intervention:
The intervention is an example of a representative intervention as described by Young et al,[24] which is a type of intervention that ensures a treatment is maintained above a threshold, e.g., at least 150 minutes of MVPA. Instead of fixing everyone's time spent doing MVPA to a fixed minute, e.g., 150 minutes, a representative intervention allows individuals to have different time spent in MVPA. This is achieved by assigning a new MVPA time that is ≥150 minutes. The assigned value is obtained from a random draw from a distribution of MVPA time that is defined by those who sustain MVPA of 150 minutes or greater at visits 1 and 3. For example, one participant may be assigned a value of 165 minutes, and another participant 170 minutes. The effect of a representative intervention can be estimated using g-formula or inverse probability weight of marginal structural models.[24,25]

In this analysis, we define the intervention as: all participants, regardless of their observed MVPA minutes/week, will be intervened at Visit 1 and Visit 3 to a MVPA minute ≥150 minutes/week randomly drawn from participants who had the same MVPA history, light PA history, and confounder history under the intervention scenario. Under the no intervention scenario, participants follow their observed MVPA minutes/week. This comparison allows us to estimate the risk reduction in the entire population (only some of whom have MVPA minute ≥150 minutes/week) if the intervention were enacted (where everyone would have MVPA minute ≥150 minutes/week), which is sometimes referred to as a measure of generalized impact, a generalization of population attributable risk, which is arguably relevant for policymaking. This comparison has been made for survival among nonmetastatic prostate cancer.[25]

Unlike real-life trials where the exchangeability between invention and control arms were achieved by randomization, the exchangeability is achieved in the emulated trial by reweighting the individuals in the study sample for both intervention and control scenarios. The estimation procedures of g-formula and IPW are described below.

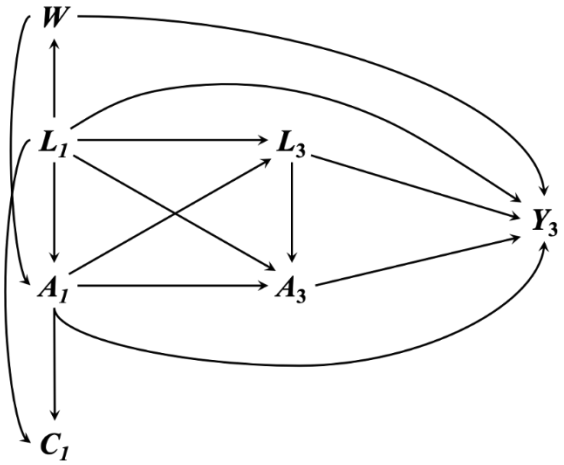
Measurement of MVPA
We will use MVPA variables derived from the sports and exercise section by Dr. Palta and colleagues.[26] A metabolic equivalent of task (MET) value was assigned to each sport or exercise type reported by the participants based on the 2011 Compendium of Physical Activities.[27] Activities were classified into light (1-3 METs) and moderate-to-vigorous physical activity (>3 METs) categories. For each category, minutes/week was estimated by multiplying the frequency and the duration for each activity and summing across all activities of the same category. Participants who reported no sports or exercise activities were assigned a 0 minute/week to both categories.

Covariates:
1. Time-fixed covariates measured at Visit 1: age, sex, race-center, education (less than high school; high school/GED/vocational school; or any college), and family income (<$25,000 per year; $25,000 to <$50,000 per year; and ≥$50,000 per year).
2. Time-varying covariates measured at Visits 1 and 3: BMI, smoking status, alcohol use, total cholesterol, eGFR, comorbidity (including hypertension, diabetes, CHD, heart failure, cancer, and COPD), and minutes of light physical activity .

The eGFR was estimated using serum creatinine, age, sex and race.[28] Hypertension was defined as a systolic blood pressure >140 mmHg or a diastolic blood pressure >90 mmHg or use of hypertensive medication. Diabetes was defined as meeting one of the following criteria: a fasting glucose ≥126 mg/dL, non-fasting glucose ≥200 mg/dL, current use of diabetes medication, and self-reported physician diagnosis. CHD was defined as a self-reported history or medical record evidence of myocardial infarction, coronary artery bypass graft or angioplasty, or myocardial infarction determined by ECG adjudication. Heart failure was identified based on self-reported heart failure medication use within the past two weeks. Cancer was ascertained by linkage with cancer registry, medical records and hospital discharge summaries, and death certificates.[29] Chronic lung disease was ascertained using self-reported physician diagnoses.

Statistical analysis:

Below causal Directed Acyclic Graph (DAG) aids the discussion on the statistical methods.




The notations used in the DAG and throughout the methods section are as follows:

- $A_{1,3}$: Physical activity intervention. A = 1 if MVPA ≥150 minutes/week, and 0 otherwise
- $Y_3$: Plasma protein levels at Visit 3
- $L_1$: Age, sex, race-center, education, family income, BMI, smoking status, alcohol use status, total cholesterol, eGFR, comorbidities (hypertension, diabetes, CHD, heart failure, cancer, and COPD), and light physical activity at Visit 1.
- $L_3$: BMI, smoking status, alcohol use status, total cholesterol, eGFR, comorbidities, and light physical activity at Visit 3.
- $C_1$: Loss to follow up before Visit 3. $C_1$ = 0 if the person is not lost, and 1 otherwise.
- W: Unmeasured confounders

G-formula via ICE and MSM by IPW methods are chosen to appropriately handle the time-varying confounders of $A_3$ and $Y_3$ but are affected by previous exposure, $A_1$. Traditional methods such as regression models conditioning on the time-varying confounders or propensity score matching cannot properly handle such confounding.

***g-formula via ICE****:* the g-formula for the frailty-associated protein level at Visit 3 is:
Under the scenario when all participants adhere to the recommendation at both Visits 1 and 3:

$$E\left(Y_3^{g=1}\right) = \sum_{\bar{k}_3} \sum_{\bar{l}_3} E\left(Y_3 \middle| \bar{K}_3 = \bar{k}_3, \bar{L}_3 = \bar{l}_3, C_1 = 0\right) \times \prod_{j=1}^{3} \left\{ f\left(l_j \middle| \bar{k}_{j-2}, \bar{l}_{j-2}, c_{j-2} = 0\right) f^{obs}\left(k_j \middle| A_j = 1, \bar{k}_{j-2}, \bar{l}_j, c_{j-2} = 0\right) \right\}$$

Under the no intervention scenario:

$$E(Y_3^{g=0}) = \sum_{\bar{k}_3} \sum_{\bar{l}_3} E(Y_3|\bar{K}_3 = \bar{k}_3, \bar{L}_3 = \bar{l}_3, C_1 = 0) \times \prod_{j=1}^{3} \{f(l_j|\bar{k}_{j-2}, \bar{l}_{j-2}, c_{j-2} = 0) f^{obs}(k_j|\bar{k}_{j-2}, \bar{l}_j, c_{j-2} = 0)\}$$

where $g = 1\ or\ 0$ denotes intervention and control scenarios; $j = 1\ or\ 3$ denotes Visit 1 and Visit 3; $Y_3$ is frailty-associated protein levels at Visit 3; $K_j$ is time in MVPA as a continuous variable; $A_j$ is the binary indicator of MVPA ≥150 minutes; $L_j$ is measured confounders including time-fixed confounders at baseline and time-varying confounders; $C_1$ is the indicator of censoring after Visit 1. The overbar represents the history of the variable through $j^{th}$ visit, e.g., $\bar{L}_3 = (L_1, L_3)$. No information from Visit 2 is used since physical activity was not measured at Visit 2. When subscript is negative, e.g, $\bar{A}_{j-2}$ when $j$ = 1, the variable is an empty set.

The following steps will be implemented to estimate the levels of frailty-associated proteins when all participants adhered at both visits:

1. Among participants who had observed MVPA ≥150 minutes/week at both Visit 1 and Visit 3 ($A_1$=$A_3$=1), a linear regression model of frailty-associated proteins at Visit 3 ($Y_3$) will be fitted with time in MVPA ($K_1$, $K_3$), time-fixed confounders and time-varying confounders ($L_1$, $L_3$) at both visits. All confounders will be modelled flexibly with non-linear terms to minimize model misspecification.
2. Among the same participants as in Step 1, a linear model of time in MVPA at Visit 3 ($K_3$) will be fitted with flexible modeling of time in MVPA at Visit 1 ($K_1$), LPA at both visits, and other confounders at both visits ($L_1$, $L_3$).
3. For all participants who had observed MVPA ≥150 minutes/week at Visit 1 ($A_1$=1) and were not lost to follow up before Visit 3 ($C_1$=0), regardless of their observed MVPA at Visit 3, a new value of the time in MVPA at Visit 3 will be randomly drawn from the model in Step 2.
4. Using the randomly drawn time in MVPA at Visit 3, predicted levels of frailty-associated proteins will be obtained using model from Step 1 for the same participants in Step 3. This step averages the protein levels over the distribution of $K_3$ and $L_3$.
5. Among the same participants of Steps 3 and 4, a linear model of the predicted protein levels from Step 4 will be fitted with the flexible modeling of time in MVPA ($K_1$), time in LPA and other confounders at Visit 1 ($L_1$).
6. Among participants who had observed MVPA ≥150 minutes/week at Visit 1 ($A_1$=1), regardless of whether they returned at Visit 3, a linear model of time in MVPA at Visit 1 ($K_1$) will be fitted with flexible modeling of LPA and other confounders at Visit 1 ($L_1$).
7. For all participants at Visit 1, regardless of their observed MVPA, a new time in MVPA at Visit 1 will be randomly drawn from the model in Step 6.
8. Using the randomly drawn time in MVPA at Visit 1, predicted levels of frailty-associated proteins will be obtained using model from Step 5 for all participants at Visit 1. The average of the predicted protein levels is the estimate of the outcome when all participants adhered to the intervention.

Under no intervention scenario, the modeling of time in MVPA among adherent participants and the random drawing of the values were omitted (Steps 2, 3, 6, and 7). Also, Steps 1, 4, and 5 will be performed among participants who had two visits regardless of their adherence status. Step 4 will use observed time in MVPA. Step 6 will be performed among participants who had Visit 1

regardless of their status at that visit. The average of the final predicted protein levels is the estimate of the outcome under no intervention scenario. The difference in the two estimates is the causal effect of interest. The 95% confidence interval can be obtained using percentile method from resampling individuals 500 times.

***MSM by IPW method**:*
G-formula method can be a more efficient estimator,[20] but it is subject to model misspecification.[24,30] Therefore, we will estimate the same effect using MSM by IPW. The stabilized weights for adhering to the intervention (i.e., MVPA ≥150 minutes/week) and for censoring (i.e., loss to follow-up) will be estimated separately for the intervention and control scenarios.

For the scenario that all participants adhere to the recommendation, stabilized weights for adherence will be estimated by

$$SW_A = \prod_{j=1}^{3} \frac{f\left(A_j = 1 \middle| \bar{A}_{j-2} = 1, \bar{C}_{j-2} = 0\right)}{f\left(A_j = 1 \middle| \bar{A}_{j-2} = 1, \bar{L}_j, \bar{K}_{j-2}, \bar{C}_{j-2} = 0\right)}$$

The numerator of the weight will be quantified by the predicted probability of adherence from a pooled logistic model conditional on visit number. The denominator will be quantified by another pooled logistic model additionally conditional on visit number, time in MVPA and time-varying confounders at the current and previous visit. At Visit 3, only participants who were adherent at Visit 1 are included in these models.

In addition, to account for loss to follow-up, we will also estimate stabilized weights for censoring. As participants can only be lost to follow-up between Visit 3 and Visit 1, the weights are estimated by:

$$SW_C = \frac{f(C_1 = 0 | A_1 = 1)}{f(C_1 = 0 | A_1 = 1, L_1, K_1)}$$

The numerator will be the marginal probability of being censored after Visit 1 among participants who were adherent at Visit 1. The model for the denominator will include time-fixed and time-varying confounders and time in MVPA.

The final weights under the adherence scenario will take the product of the $SW_A$ and $SW_C$ for each individual. Covariate balance will be checked using standardized mean difference between participants who adhere to the intervention and participants who did not adhere to the intervention.[31,32] The check will be performed at Visit 1 and Visit 3 separately. At Visit 3, the check will be performed among participants who were adherent at Visit 1.

Under no intervention scenario, every participant has a weight of 1 for adherence. The stabilized weights for censoring (also the final weights) will be estimated by:

$$SW_C = \frac{f(C_1 = 0)}{f(C_1 = 0 | L_1, K_1)}$$

Lastly, we will pool the data from people who are adherent to the intervention at both visits and their final weights under the intervention with the data from all participants and their final weights under the control scenario into a single dataset. The causal effect of the intervention will

be estimated from this pooled dataset by below saturated MSM model using linear regression, weighted by the final weights:

$$E\left(Y_3^g\right) = \alpha_0 + \alpha_1 I(G)$$

where G = 1 for adherence scenario, and G = 0 for no intervention scenario.

Assumptions:

For the estimates to be causal, three assumptions (i.e., exchangeability, positivity, and consistency) need to be met. The exchangeability assumption in our study is

$$Y_3^g \perp A_j \,|\bar{L}_j = \bar{l}_j, \bar{A}_{j-2} = \bar{a}_{j-2}^g$$

This assumption may be violated due to the presence of unmeasured confounders, $W$, as shown in the DAG. In a sensitivity analysis, we will exclude smokers and participants with comorbidity (including hypertension, diabetes, CHD, heart failure, cancer, and COPD), abnormal eGFR (<30 ml/min per 1.73 $m^2$), and BMI below 18.5 $kg/m^2$. Moreover, we will also estimate an E-value that measures how strong the unmeasured confounding needs to be to explain away the PA-protein associations.[33]

The positivity assumption in this study is

$$f_{A_j\,|\bar{L}_j,\bar{A}_{j-2}}\left(a_j^g \middle| \bar{l}_j, \bar{a}_{j-2}^g\right) \equiv f^{obs}\left(a_j^g \middle| \bar{l}_j, \bar{a}_{j-2}^g\right) \text{ with probability 1}$$

This assumption will be empirically checked in the data. Without conditioning on any covariates, approximately 2,100 participants had MVPA ≥150 minutes/week at both Visits 1 and 3 (17.5% of the sample at Visit 1, and 25.6% of the sample at Visit 3).

The consistency assumption is an assumption that is particularly relevant for physical activity intervention. This assumption states that the value of the potential outcome when a person is assigned to treatment group is equal to the observed outcome when the person is in fact in the treatment group.[34] This assumption is often violated when there are multiple versions of treatment. There are at least two types of versions for our treatment, achieving ≥150 minutes/week of MVPA, in this analysis. The first version type is the actual minutes of MVPA above 150 minutes/week performed by participants. For example, some performed 160 minutes, and some 170 minutes. The other version type is the different types of sports in which the participants engage, and the types of sports (for example, jogging [endurance] vs weightlifting [strength]) may have different implications for frailty-associated proteins. The presence of multiple versions imposes two challenges: identification of causal effects and transportability of the effects. To identify the causal association between the treatment (150 minutes/week of MVPA) and plasma protein levels at Visit 3, it is assumed that there is no unmeasured confounder of treatment and versions that does not affect the outcome.[34] One example of the such confounders can be what type of exercise facilities the participants have in their neighborhood. We do not have this variable in ARIC. Therefore, we will empirically check if there are associations between type of sports/actual MVPA minutes and MVPA ≥150 minutes adjusting for the covariates. The challenge to transportability stems from the fact that other populations may have different distributions of MVPA minutes and type of the sports. Therefore, in a sensitivity analysis, we will estimate the effects of achieving 150 minutes/week of MVPA under different hypothetical distributions of sports types and MVPA minutes.

Because we have loss to follow-up between Visit 1 and Visit 3, the last assumption we make is that loss to follow-up is independent of potential outcome given the history of MVPA, light physical activity and confounders.

Limitations:
1. Plasma proteins are not measured at Visit 1, though the impact is mitigated through sensitivity analyses.
2. The focus on MVPA may overlook the health benefits of light and non-exercise physical activity.[35] We chose MVPA because the published guideline is based on MVPA. We could explore including light physical activity, but light physical activity may not be well measured by self-report questionnaires, and the threshold may be arbitrary without a guideline.
3. Physical activity is measured using self-report which is subject to recall and social desirability biases. Moreover, though the modified Baecke questionnaire also assessed leisure and work activity, these two domains were not included in the calculation of MET values.[26] This omission may underestimate the MVPA, and also limit our findings to sports instead of more general physical activity.
4. Lack of replication in external cohort. We are open to use the Baltimore Longitudinal Study of Aging (BLSA) as a potential external validation cohort. The advantage of the BLSA is that physical activity is measured objectively using accelerometers. Both limitation #2 and #3 can be addressed using accelerometers. The plan to replicate in the BLSA is not firm because we are uncertain how many BLSA participants, if any, have repeated physical activity measurement and proteomics data in their mid-life. This analysis in ARIC with a large sample size will provide first line of evidence, but replications are important when suitable cohorts become available.

**7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes __x__ No**

**b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No**
(The file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

**8.a. Will the DNA data be used in this manuscript? ____ Yes __x__ No**

**8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ____ Yes ____ No**

**9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status.**
ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: http://www.cscc.unc.edu/aricproposals/dtSearch.html

___x___ Yes _____ No

**10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?**

There are no prior proposals on physical activity plasma proteins, but below proposal are relevant.

MP 3438: Steve Nguyen, Kelley Pettee Gabriel, Anna Kucharska-Newton, B. Gwen Windham, Lisa Pompeii, Pamela L. Lutsey, Priya Palta, Ellen Demerath. Association of Mid- to Late-Life Physical Activity with Later Life Physical Function and Frailty

MP 3917: Proteomic Markers Linking Frailty and Heart Failure: The ARIC study. Diego Ramonfaur, Brian Claggett, Hicham Skali, Dalane Kitzman, Beverly Gwen Windham, Priya Palta, Jennifer Schrack, Fangyu Liu, Suma Konety, Chiadi Nduleme, Keenan Walker, Josef Coresh, Bing Yu, Amil Shah

MP 3609: Using Information from Multiple Cohort Studies to Simulate and Extend SPRINT-MIND: What Would the Effect of Intensive Blood Pressure Treatment on Prevention of Dementia Have Been If the Trial Had Not Been Terminated Early? : Sarah Ackley, Erin Bennett, Annette Fitzpatrick, Kan Z. Gianattasio, M. Maria Glymour, Rebecca Gottesman, Deborah A. Levine, Lenore Launer, Megha L. Mehrotra, Melinda C. Power (last), Jingkai Wei (first), Scott C. Zimmerman

**11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? __x__ Yes ____ No**

**11.b. If yes, is the proposal**

**___ A. primarily the result of an ancillary study (list number* ___2017.27; 2008.06_____)**

**___ B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____ _____ _____)**

*ancillary studies are listed by number https://sites.cscc.unc.edu/aric/approved-ancillary-studies

**12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.**

**12b. The NIH instituted a Public Access Policy in April, 2008** which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in http://www.cscc.unc.edu/aric/index.php, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

1. Chen R, Wu Q, Wang D, et al. Effects of elastic band exercise on the frailty states in pre-frail elderly people. *Physiother Theory Pract*. 2020;36(9):1000-1008. doi:10.1080/09593985.2018.1548673

2. Hsieh TJ, Su SC, Chen CW, et al. Individualized home-based exercise and nutrition interventions improve frailty in older adults: a randomized controlled trial. *Int J Behav Nutr Phys Act*. 2019;16(1):119. doi:10.1186/s12966-019-0855-9

3. Sadjapong U, Yodkeeree S, Sungkarat S, Siviroj P. Multicomponent Exercise Program Reduces Frailty and Inflammatory Biomarkers and Improves Physical Performance in Community-Dwelling Older Adults: A Randomized Controlled Trial. *Int J Environ Res Public Health*. 2020;17(11):E3760. doi:10.3390/ijerph17113760

4. Ng TP, Feng L, Nyunt MSZ, et al. Nutritional, Physical, Cognitive, and Combination Interventions and Frailty Reversal Among Older Adults: A Randomized Controlled Trial. *Am J Med*. 2015;128(11):1225-1236.e1. doi:10.1016/j.amjmed.2015.06.017

5. Tarazona-Santabalbina FJ, Gómez-Cabrera MC, Pérez-Ros P, et al. A Multicomponent Exercise Intervention that Reverses Frailty and Improves Cognition, Emotion, and Social Networking in the Community-Dwelling Frail Elderly: A Randomized Clinical Trial. *J Am Med Dir Assoc*. 2016;17(5):426-433. doi:10.1016/j.jamda.2016.01.019

6. Kim H, Suzuki T, Kim M, et al. Effects of exercise and milk fat globule membrane (MFGM) supplementation on body composition, physical function, and hematological parameters in community-dwelling frail Japanese women: a randomized double blind, placebo-controlled, follow-up trial. *PloS One*. 2015;10(2):e0116256. doi:10.1371/journal.pone.0116256

7. Cesari M, Vellas B, Hsu FC, et al. A physical activity intervention to treat the frailty syndrome in older persons-results from the LIFE-P study. *J Gerontol A Biol Sci Med Sci*. 2015;70(2):216-222. doi:10.1093/gerona/glu099

8. Nagai K, Miyamato T, Okamae A, et al. Physical activity combined with resistance training reduces symptoms of frailty in older adults: A randomized controlled trial. *Arch Gerontol Geriatr*. 2018;76:41-47. doi:10.1016/j.archger.2018.02.005

9. Oliveira JS, Pinheiro MB, Fairhall N, et al. Evidence on Physical Activity and the Prevention of Frailty and Sarcopenia Among Older People: A Systematic Review to Inform the World Health Organization Physical Activity Guidelines. *J Phys Act Health*. Published online August 11, 2020:1-12. doi:10.1123/jpah.2020-0323

10. Trombetti A, Hars M, Hsu FC, et al. Effect of Physical Activity on Frailty: Secondary Analysis of a Randomized Controlled Trial. *Ann Intern Med*. 2018;168(5):309-316. doi:10.7326/M16-2011

11. Serra-Prat M, Sist X, Domenich R, et al. Effectiveness of an intervention to prevent frailty in pre-frail community-dwelling older people consulting in primary care: a randomised controlled trial. *Age Ageing*. 2017;46(3):401-407. doi:10.1093/ageing/afw242

12. Cesari M, Vellas B, Hsu FC, et al. A physical activity intervention to treat the frailty syndrome in older persons-results from the LIFE-P study. *J Gerontol A Biol Sci Med Sci*. 2015;70(2):216-222. doi:10.1093/gerona/glu099

13. Fried LP. Interventions for Human Frailty: Physical Activity as a Model. *Cold Spring Harb Perspect Med*. 2016;6(6). doi:10.1101/cshperspect.a025916

14. Angulo J, El Assar M, Álvarez-Bustos A, Rodríguez-Mañas L. Physical activity and exercise: Strategies to manage frailty. *Redox Biol*. 2020;35:101513. doi:10.1016/j.redox.2020.101513

15. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. *Lancet Lond Engl*. 2019;394(10206):1376-1386. doi:10.1016/S0140-6736(19)31785-4

16. Ubaida-Mohien C, Gonzalez-Freire M, Lyashkov A, et al. Physical Activity Associated Proteomics of Skeletal Muscle: Being Physically Active in Daily Life May Protect Skeletal Muscle From Aging. *Front Physiol*. 2019;10:312. doi:10.3389/fphys.2019.00312

17. Stattin K, Lind L, Elmståhl S, et al. Physical activity is associated with a large number of cardiovascular-specific proteins: Cross-sectional analyses in two independent cohorts. *Eur J Prev Cardiol*. 2019;26(17):1865-1873. doi:10.1177/2047487319868033

18. Ubaida-Mohien C, Lyashkov A, Gonzalez-Freire M, et al. Discovery proteomics in aging human skeletal muscle finds change in spliceosome, immunity, proteostasis and mitochondria. *eLife*. 2019;8:e49874. doi:10.7554/eLife.49874

19. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci*. 2004;59(3):255-263. doi:10.1093/gerona/59.3.m255

20. Wen L, Young JG, Robins JM, Hernán MA. Parametric g-formula implementations for causal survival analyses. *Biometrics*. 2021;77(2):740-753. doi:10.1111/biom.13321

21. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

22. 2018 Physical Activity Guidelines Advisory Committee. *2018 Physical Activity Guidelines Advisory Committee Scientific Report*. U.S. Department of Health and Human Services; 2018:779. Accessed July 13, 2020. https://health.gov/our-work/physical-activity/current-guidelines/scientific-report

23. Kucharska-Newton A, Griswold M, Yao ZH, et al. Cardiovascular Disease and Patterns of Change in Functional Status Over 15 Years: Findings From the Atherosclerosis Risk in Communities (ARIC) Study. *J Am Heart Assoc*. 2017;6(3):e004144. doi:10.1161/JAHA.116.004144

24. Young JG, Logan RW, Robins JM, Hernán MA. Inverse probability weighted estimation of risk under representative interventions in observational studies. *J Am Stat Assoc*. 2019;114(526):938-947. doi:10.1080/01621459.2018.1469993

25. Dickerman BA, Giovannucci E, Pernar CH, Mucci LA, Hernán MA. Guideline-Based Physical Activity and Survival Among US Men With Nonmetastatic Prostate Cancer. *Am J Epidemiol*. 2019;188(3):579-586. doi:10.1093/aje/kwy261

26. Palta P, Sharrett AR, Gabriel KP, et al. Prospective Analysis of Leisure-Time Physical Activity in Midlife and Beyond and Brain Damage on MRI in Older Adults. *Neurology*. 2021;96(7):e964-e974. doi:10.1212/WNL.0000000000011375

27. Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc*. 2011;43(8):1575-1581. doi:10.1249/MSS.0b013e31821ece12

28. Muntner P, He J, Astor BC, Folsom AR, Coresh J. Traditional and nontraditional risk factors predict coronary heart disease in chronic kidney disease: results from the atherosclerosis risk in communities study. *J Am Soc Nephrol JASN*. 2005;16(2):529-538. doi:10.1681/ASN.2004080656

29. Joshu CE, Barber JR, Coresh J, et al. Enhancing the Infrastructure of the Atherosclerosis Risk in Communities (ARIC) Study for Cancer Epidemiology Research: ARIC Cancer. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol*. 2018;27(3):295-305. doi:10.1158/1055-9965.EPI-17-0696

30. McGrath S, Young JG, Hernán MA. Revisiting the g-null Paradox. *Epidemiol Camb Mass*. 2022;33(1):114-120. doi:10.1097/EDE.0000000000001431

31. Jackson JW. Diagnosing Covariate Balance Across Levels of Right-Censoring Before and After Application of Inverse-Probability-of-Censoring Weights. *Am J Epidemiol*. 2019;188(12):2213-2221. doi:10.1093/aje/kwz136

32. Schnellinger EM, Valeri L, Jackson JW. DATA VISUALIZATION TOOLS FOR CONFOUNDING AND SELECTION BIAS IN LONGITUDINAL DATA: THE %LENGTHEN, %BALANCE, AND %MAKEPLOT (CONFOUNDR) MACROS AND R PACKAGE. *Am J Epidemiol*. 2020;189(12):1633-1636. doi:10.1093/aje/kwaa143

33. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607

34. VanderWeele TJ, Hernán MA. Causal Inference Under Multiple Versions of Treatment. *J Causal Inference*. 2013;1(1):1-20. doi:10.1515/jci-2012-0002

35. Liu F, Wanigatunga AA, Schrack JA. Assessment of Physical Activity in Adults using Wrist Accelerometers. *Epidemiol Rev*. Published online July 2, 2021:mxab004. doi:10.1093/epirev/mxab004