JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

Johns Hopkins University, Dept. of Biostatistics Working Papers

8-16-2010

SURROGATE SCREENING MODELS FOR THE LOW PHYSICAL ACTIVITY CRITERION OF FRAILTY

Sandrah P. Eckel
Division of Biostatistics, Keck School of Medicine, University of Southern California, eckel@usc.edu

Karen Bandeen-Roche
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics and Center on Aging and Health, Department of Medicine

Paulo H.M. Chaves
Center on Aging and Health and Division of Geriatric Medicine and Gerontology, Department of Medicine, The Johns Hopkins Medical Institutions

Linda P. Fried
Mailman School of Public Health, Columbia University

Thomas A. Louis
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Surrogate screening models for the low physical activity criterion of frailty

Sandrah P. Eckel[1], Karen Bandeen-Roche[2,3], Paulo H.M. Chaves[3,4],

Linda P. Fried[5], and Thomas A. Louis[2]

[1]Division of Biostatistics, Keck School of Medicine, University of Southern California, Los Angeles, California.

[2]Department of Biostatistics, The Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

[3]Center on Aging and Health and [4]Division of Geriatric Medicine and Gerontology, Department of Medicine, The Johns Hopkins Medical Institutions, Baltimore, Maryland.

[5]Mailman School of Public Health, Columbia University, New York, New York.

CORRESPONDING AUTHOR

Sandrah P. Eckel, Ph.D.

Division of Biostatistics,

Keck School of Medicine,

1540 Alcazar Street, CHP-220,

Los Angeles, CA 90089-9011.

Telephone: (323) 442-2030

Fax: (323) 442-2349

Email: eckel@usc.edu

ALTERNATE CORRESPONDING AUTHOR

Thomas A. Louis, Ph.D.

Email: tlouis@jhsph.edu

RUNNING HEAD

Surrogate low physical activity models

WORD COUNT

Abstract: 246

Text : 3,601
(excludes figures, tables and references)

FIGURES/TABLES

Figures: 2

Tables: 3

Abstract

Background and Aims. Low physical activity, one of five criteria in a validated clinical phenotype of frailty, is assessed by a standardized questionnaire on up to 20 leisure time activities. Because of the time demanded to collect the interview data, it has been challenging to translate to studies other than the Cardiovascular Health Study (CHS), for which it was developed. Considering subsets of activities, we identified and evaluated streamlined surrogate assessment methods and compared them to one implemented in the Women's Health and Aging Study (WHAS).

Methods. Using data on men and women ages 65 and older from the CHS, we applied logistic regression models to rank activities by "relative influence" in predicting low physical activity. We considered subsets of the most influential activities as inputs to potential surrogate models (logistic regressions). We evaluated predictive accuracy and predictive validity using the area under receiver operating characteristic curves and assessed criterion validity using proportional hazards models relating frailty status (defined using the surrogate) to mortality.

Results. Walking for exercise and moderately strenuous household chores were highly influential for both genders. Women required fewer activities than men for accurate classification. The WHAS model (8 CHS activities) was an effective surrogate, but a surrogate using 6 activities (walking, chores, gardening, general exercise, mowing and golfing) was also highly predictive.

Conclusions. We recommend a 6 activity questionnaire to assess physical activity for men and women. If efficiency is essential and the study involves only women, fewer activities can be included.

Keywords

Frail Older Adults, Logistic Regression, Physical Activity, Screening.

Introduction

Frailty is largely considered to be a medical syndrome characterized by multi-system decline. It has been operationalized into a clinical definition by identifying and integrating 5 characteristics that appear consistent with relevant hypotheses to constitute a clinical syndrome: weakness (as measured by grip strength), poor endurance, unintended weight loss, low physical activity and slow walking speed (1). Individuals are considered frail if they satisfy 3 or more criteria, prefrail if they satisfy 1 or 2 and robust if they satisfy none. Using this definition, frailty in older adults has been found to be associated with important increased risks of falls, hospitalization, worsening mobility, worsening disability and death (1, 2).

The 5 criteria characterization of frailty – originally developed and validated using data from the Cardiovascular Health Study (CHS) – is being increasingly applied in the gerontological literature. However, the low physical activity criterion has proven to be particularly challenging to assess and to translate to other studies and potentially to clinical settings because of the time required. In the CHS, physical activity is ascertained using a modified version of the Minnesota Leisure Time Activities (MLTA) Questionnaire (3) assessing participation in up to 20 activities. In other studies, including the Women's Health and Aging Study (2), investigators have applied modified definitions (4, 5, 6, 7, 8, 9, 10) or have excluded the criterion altogether (11, 12). The CHS questionnaire is time-consuming to administer and presents considerable participant and analyst burden. A streamlined version would be of interest to the gerontological community and could be included in a simplified screening tool for frailty.

We investigated whether responses on a subset of activities in the current questionnaire could be used to accurately determine physical activity status as defined by the current criterion. This work introduces and validates streamlined surrogate assessments, informs on the validity of modified low physical activity criteria and offers guidance for future simplifications.

Methods

Study Population and Measures

We used data from the CHS, a longitudinal study of 5,888 older adults originally designed to study the epidemiology of cardiovascular disease in older adults (13). CHS participants were recruited from age- and gender-stratified random samples of Medicare eligibility lists in 4 U.S. communities. Beginning in 1989, CHS participants in the original cohort (Cohort 1, N=5201) underwent annual clinical examinations. Due to inadequate minority representation, a second cohort of African Americans (Cohort 2, N= 687) was added to the CHS starting in 1992. As presented in detail elsewhere (1, 13), extensive self-reported and clinical information was gathered at each examination.

Physical activity questionnaire.–Physical activity was assessed using a modified version of the Minnesota Leisure Time Activities (MLTA) Questionnaire (1, 3) on up to 20 activities. Fifteen activities were pre-specified: walking for exercise, moderately strenuous household chores, mowing the lawn, raking the lawn, gardening, hiking, jogging, biking, exercise cycling, dancing, aerobics/aerobic dancing, bowling, golfing, calisthenics/general exercising and swimming. Respondents could report on up to 2 additional activities of his/her choice. Three other activities (doubles and singles tennis and racquetball) were included in the 1989–1990 questionnaire, but excluded in subsequent years due to low response rates. Participants who had engaged in an activity during the prior 2 weeks were further questioned on the number of sessions, the average session duration and the seasonality of his/ her participation. The questionnaire was administered at baseline (1989-1990 for Cohort 1 and 1992-1993 for Cohort 2) and at up to 2 follow up visits (1992-1993 for Cohort 1 and at 1996-1997 for both cohorts).

Low physical activity determination.–Kilocalorie per week (kcal/week) expenditures were calculated for each activity using its metabolic equivalent (MET) score (a measure of exercise intensity) (14,15):

$$\begin{aligned}&\text{activity-specific MET (kcal/(kg} \times \text{hour))} \times \text{body weight (kg)}\\&\quad \times \text{activity duration (minutes)/60}\\&\quad \times \text{number of sessions in past two weeks/2}\\&\quad \times \text{number of months per year/12.}\end{aligned}$$

Previous CHS kcal/week computations implicitly assumed each participant weighed 60 kg. We accounted for differences in body weight by including the multiplicative factor (body weight in kg)/60. Recent work in the CHS updated kcal/week expenditures by assigning MET scores to free response activities, deleting unlikely values for duration or frequency for all activities and filling in missing items using single regression imputation based on age, gender, race, body mass index and corresponding activity data (16). Total kcal/week expenditure on leisure time activities was obtained by summing expenditures over all activities (15 pre-specified (18 for Cohort 1 at baseline) and the up to 2 free response) using the updated activity-specific expenditures. After pooling the cohorts, we defined participants with the gender-specific lowest quintile of total kcal/week expenditure at baseline to have low physical activity (1). The baseline threshold was applied at follow up. We had access to the updated total kcal/week expenditures from (16) (which we adjusted for body weight) and questionnaire response information. Activity-specific kcal/week expenditures were not available, so we reproduced the updated calculation for the 15 activities.

Pre-existing modified criterion.–In the Women's Health and Aging Study (WHAS), physical activity was assessed by a questionnaire on 6 activities, 5 of which are standard CHS activities: walking, doing strenuous household chores, dancing, bowling and exercise. We assumed that the sixth activity, "doing strenuous outdoor chores," corresponds with 3 CHS activities: mowing, raking and gardening. Low physical activity was assigned using a threshold on total kcal/week expenditures from these activities 1/3 the value of the gender-specific CHS threshold from (1).

Exclusions.–We applied the CHS baseline exclusion criteria for frailty to remove participants with diseases potentially producing frailty-like characteristics (1). We excluded 550 participants at baseline: 47 with "a history of Parkinson's disease," 249 with adjudicated stroke, 74 with "Mini-Mental scores < 18," and 235 "who were taking Sinemet, Aricept, or antidepressants"(1). We further excluded 45 participants with incomplete baseline data on low physical activity or on the 15 leisure time activities. Participants missing 3 or 4 of the other frailty criteria (6 participants at baseline) were assigned missing frailty status and were included in physical activity analyses, but excluded from analyses involving frailty status.

Data Analysis

We used baseline data to develop and evaluate potential surrogate models for low physical activity. Surrogates were constructed using logistic regression. Covariates eligible for inclusion in the surrogate models were the kcal/week expenditures from each of the 15 activities (henceforth referred to in the model as “activities”). Analyses were performed for all participants and for subsets defined by gender to study possible gender differences. For all-participant models, we included the main effect for an indicator variable for men and two-way interaction terms between the indicator and all included activities. The following steps were taken to produce surrogates using a subset of the original activities.

Relative influence.–To calculate the “relative influence” of each activity, we fit a model including all 15 activities and recorded the deviance, a measure of goodness of fit (17). We then removed an activity, refit the model, and recorded the resultant change in deviance. We refer to this as the relative influence of the activity in predicting low physical activity. We repeated this procedure to obtain relative influences for all activities. We normalized the relative influences to sum to 100.

Potential surrogates.–We ranked activities by relative influence and proposed potential surrogates based on nested subsets of the most relatively influential activities. We constructed potential surrogates including the 1, 2, 3, 4, 5, 6, 8, 10, 12 and 15 most relatively influential activities, respectively. We compared predictive accuracy of the potential surrogates using area under the receiver operating characteristic curve (AUC). AUC ranges from 0.5 to 1.0, with larger values signifying better discrimination between low and not low physical activity individuals (18). Using baseline AUC, we selected several all-participant surrogates for further evaluation. We only considered all-participant surrogates because using the same activities for men and women allows for a unified assessment method.

Cross-validation at baseline.–To avoid overfitting surrogate models to baseline CHS data, we calculated the relative influence and baseline AUC using 10 fold cross-validation (19). In this procedure, the data is randomly split into 10 equally sized subsamples and the model is fit with 9/10 of the data. Predictions are obtained from the resultant model for the remaining 1/10 of the data. The process is repeated until predictions have been obtained for all 10 subsamples. The predictions are pooled and the relative influence or AUC is calculated.

Predictive validity.–We assessed predictive validity of the surrogates developed at baseline by comparing cohort-specific (cross-validated) AUC at baseline to the corresponding AUC calculated at follow up. We compared results for the surrogates developed in this study to those for the 8 CHS activity WHAS surrogate.

Discrimination threshold and misclassification.–We imposed gender-specific worst quintile thresholds to assign low physical activity status at baseline according to the surrogates. Hence for the logistic regression models, participants with the 20% highest predicted probabilities of low physical activity for their gender were assigned low physical activity. Using the surrogate physical activity classification, we created a new frailty status classification (robust/prefrail/frail). We then calculated the percent of participants misclassified in terms of low physical and frailty status when using each surrogate instead of the original criterion.

Criterion validity.–We performed criterion validation using a previously developed proportional hazards model (1) that relates baseline frailty status to mortality after 3 years of follow up. We examined the raw association of baseline frailty with mortality and the association after adjusting for covariates previously found to be associated with the risk of mortality in this cohort: age; gender; cohort; self-reported: general health, income, smoking status, difficulty with instrumental activities of daily living (IADL); brachial blood pressure; ankle arm index; fasting glucose; albumin; creatinine; maximum percentage stenosis of the internal carotid artery (by ultrasound); adjudicated history of congestive heart failure (CHF); cognitive function (by Mini-Mental State Examination); any major electrocardiogram abnormality; use of diuretics without a history of hypertension or CHF; any difficulty with IADL; and modified CES-D depression score (excluding those questions in the "exhaustion" frailty criterion) (1, 20).

We fit separate proportional hazards models for frailty defined using the following low physical activity assessment methods: the original criterion, the WHAS surrogate, the 6 activity, 2 activity, and 1 activity surrogates, and excluding low physical activity (where the number of criteria for each frailty category is the same, but participants have at most 4 criteria). We evaluate the impact of applying surrogates for low physical activity in these models using Akaike's Information Criteria (AIC) (21), a measure of goodness of fit for which smaller values imply superior fit, and the hazard ratios comparing prefrail to robust and frail to prefrail.

RESULTS

At baseline, the study included 5,293 participants, with a mean age of 73 years (interquartile range: 68 to 76 years), of whom 58% were women and 15% were African American. Moderately strenuous household chores and walking for exercise were the most common leisure time activities, with more than half of all participants having participated in each at baseline (Table 1). Activity participation varied by gender, with higher participation rates for men in mowing and raking the lawn, golfing and biking, and a higher participation rate for women in chores. For all activities except chores, the median kcal/week expenditures for those who participated were greater for men than for women. The median expenditure on golfing was markedly higher than any of the other activities.

[Table 1 about here.]

Baseline lowest quintile thresholds on total kcal/week expenditures (men: 332, women: 154) were different from those previously reported in the CHS (men: 383, women: 270) (1) because total expenditures were adjusted for body weight and included free response activities. Low physical activity prevalence was higher for Cohort 2 and increased with time. For Cohort 1, the prevalence was 0.18, 0.24 and 0.29 at 1989-1990, 1992-1993 and 1996-1997, respectively. For Cohort 2, the prevalence was 0.35 and 0.37 at 1992-1993 and 1996-1997, respectively.

Surrogate models

Relative influence.–The most (walking and chores) and least (hiking and jogging) relatively influential activities were similar for men and women (Figure 1). Activity ranking by relative influence was similar to ranking by percent participation (Table 1). Gender-related differences in relative influence ranking were particularly evident with chores, mowing, golfing, raking, dancing and aerobics. For men, a large number of activities had non-negligible relative influence while for women, most of the relative influence belonged to 4 activities: chores, walking, general exercise and gardening.

[Figure 1 about here.]

Baseline AUC.–Baseline AUC for potential surrogates was generally high (Table 2), with all-participant surrogates that include more than 1 activity having AUC ≥0.920. AUC was generally larger in models that included more activities. Models for men had smaller AUC than models for women that included the same number of activities (the activities included may have differed).

[Table 2 about here.]

Surrogate selection.–We selected the following all-participant surrogate models for further evaluation. The 6 activity surrogate (walking, chores, gardening, general exercise, mowing and golfing) had high baseline AUC of 0.975 and included the top 4 activities for men and women, in terms of relative influence. To investigate whether simple surrogates validated well, we also selected the 1 and 2 activity surrogates (walking, and walking and chores, respectively).

Determination of physical activity by the surrogate.–To calculate low physical activity status using the streamlined surrogate: (i) substitute appropriate activity-specific kcal/week expenditures in the following logistic regression equations, (ii) sum to obtain the predicted log odds of low physical activity, x, (iii) transform to the probability scale by calculating exp(x)/(1+exp(x)) and (iv) classify probabilities greater than or equal to the discrimination threshold (6 activity surrogate: men 0.537, women 0.547; 2 activity surrogate: men 0.498, women 0.567) as low physical activity.

Logistic regression coefficients for the 6 activity surrogate:

Men: 1.885 – 0.009 walking – 0.008 chores – 0.011 gardening – 0.011 exercising – 0.010 mowing – 0.018 golfing

Women: 2.966 – 0.025 walking – 0.026 chores – 0.027 gardening – 0.028 exercising – 0.035 mowing – 0.129 golfing

Logistic regression coefficients for the 2 activity surrogate:

Men: – 0.010 – 0.007 walking – 0.006 chores

Women: 1.419 – 0.020 walking – 0.019 chores

WHAS vs. 6 activity surrogates.–There was considerable overlap in the activities included in the 6 activity surrogate and in the 8 CHS activities included in the WHAS surrogate. Notable differences in activity inclusion occurred with bowling and dancing in the WHAS surrogate, each of which had relatively low participation rates, and with golfing in the 6 activity surrogate, for which participations rates were low, but kcal/week expenditures were high.

Follow up AUC.–AUC of the surrogate models constructed at baseline remained relatively constant at follow up

(Table 3), despite increasing prevalences of low physical activity. The WHAS and 6 activity surrogates displayed the best AUC, with values ranging from 0.965 to 0.978 over the cohorts and study years. The 1 and 2 activity surrogates had lower and more variable AUC, with AUC for the 2 activity surrogate ranging from 0.898 to 0.921 and AUC for the 1 activity surrogate ranging from 0.689 to 0.768.

[Table 3 about here.]

Discrimination threshold.–For the WHAS surrogate, the gender-specific worst quintile thresholds on total expenditures for the 8 activities (105 kcal/week for women and 148 for men) were higher than the thresholds previously applied in the WHAS (90 kcal/week for women and 128 kcal/week for men, had men been included) (2). Had we applied the WHAS thresholds, low physical activity prevalence at baseline in CHS would have been 0.18 for women and 0.16 for men.

Misclassification.–Surrogates including more activities misclassified fewer participants in terms of low physical activity at baseline. Low physical activity misclassification rates were 6.8%, 7.2%, 15.4% and 32.8%, respectively, for the WHAS, 6, 2 and 1 activity surrogates. Misclassification rates were lower for frailty status: 4.0%, 4.3%, 9.6% and 20.8%, respectively. Frailty status misclassification from the 1 activity surrogate was greater than if the low physical activity criterion had been excluded altogether (11.3%).

[Figure 2 about here.]

Criterion validity.–The 6 activity and WHAS surrogates produced models with hazard ratios (adjusted and unadjusted) larger in absolute value than those resulting from use of the original criterion and AIC smaller than from the original criterion (Figure 2). Hazard ratios resulting from the 1 and 2 activity surrogates or from excluding the low physical activity criterion were biased towards the null of 1. Although the 1 activity surrogate performed more poorly than the 2 activity surrogate in terms of AUC and misclassification, it produced hazard ratios closer to those from the original criterion.

Discussion

We determined that the low physical activity criterion of frailty can be effectively assessed using a subset of the up to 20 activities in the current questionnaire. We recommend including 6 activities (walking, chores, gardening, general exercise, mowing and golfing) in a surrogate. This surrogate has strong validity because: (1) it was developed using baseline CHS data and has similar predictive accuracy at follow up, (2) the resultant reclassifications in terms of low physical activity status and categorical (robust/prefrail/frail) frailty status are relatively low and (3) frailty status – as defined by the surrogate – has an association with the risk of mortality in older adults similar to that found with the original frailty criteria. Use of this streamlined surrogate will ease participant burden, simplify the assignment of low physical activity status and speed frailty assessment.

We aimed to simplify assessment, so our ultimate recommendation uses the same activities for both men and women. However, we found that women required fewer activities than men for accurate assessment, so studies involving only women can reasonably use a smaller set of activities (chores, walking, general exercising and gardening were the 4 activities with highest relative influence for women).

We confirmed that the WHAS surrogate includes key activities for determining physical activity status in an older adult population similar to the CHS. The surrogate corresponds well with the original criterion and our cross-validation using data from the CHS provides evidence to support its continued application. For future applications of the WHAS surrogate, we recommend that investigators consider applying the higher threshold on kcal/week expenditures (105 kcal/week for women and 148 for men) for assignment of low physical activity status. The original WHAS threshold produces a lower prevalence of low physical activity and might attenuate frailty-related findings.

A limitation of our study lies in the selection of the number of activities to include in a surrogate. We informally weighed the tradeoff between efficiency (fewer activities) and good predictive accuracy (more activities). If additional information had been available on the relative costs of false positives and false negatives in low physical activity misclassification and on the cost of administering the current questionnaire, a cost-benefit analysis could have been performed to determine the optimal number of activities and the optimal discrimination threshold. It is unclear how streamlining will affect the face validity of physical activity assessment. Participants may miss the opportunity to respond on activities of their choice. Applications of the 6 activity surrogate in ongoing or new studies will provide insight.

Our work assumes that the current standardized MLTA questionnaire and criterion for low physical activity is a gold standard. This may not be true. For the CHS study population, the WHAS and 6 activity surrogates produced frailty status hazard ratios for 3 year mortality that were larger than for frailty as originally defined. This provides evidence that other physical activity measures may potentially produce a refined frailty categorization.

In this study, we had a narrowly defined goal of identifying and evaluating surrogates based on subsets of activities in the current questionnaire. In related work (22), we investigated whether surrogates can be constructed from other physical activity information available in the CHS that was not used in the original criterion, including: the number of city blocks walked outside the home in the last week, usual walking pace outside the home, the number of flights of stairs climbed in the last week, and the number of hours spent seated or lying down in a typical 24 hour period. Neither of these studies answers the broader and important question of how to best assess low physical activity for frailty, but they do produce information that can be used to address this question in the future. Physical activity is challenging to assess in older adults, as evinced by the large body of literature on the subject and the range of assessment methods applied in other studies to determine the low physical activity criterion of frailty (23, 24, 25, 26, 27). The assessment of physical activity for frailty differs from much of this work because it aims only to identify older adults at the lowest end of the energy expenditure spectrum. Future work on developing alternative assessments might consider focusing on a small number of age-appropriate low to moderate energy activities (not limited to leisure-time activities) that effectively distinguish between low and not low physical activity adults. It is also of interest to develop a screening tool that can be applied to older adults who are not living in the community but, for example, who are hospitalized or institutionalized. Differences in culture, socioeconomic status and the built environment may require additional consideration.

FUNDING

This work was supported by grants from the National Institute on Aging (NIA) at the National Institutes of Health (NIH): NIA T32 AG00247, WHAS1-N01AG12112, WHAS2 - R01AG11703, and WHAS3 (MERIT Award) - 1R37AG19905; and from the National Institute of Diabetes and Digestive and Kidney Diseases at NIH: R01 DK061662.

ACKNOWLEDGMENTS

The Cardiovascular Health Study was supported by contract numbers N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, grant number U01 HL080295 from the National Heart, Lung, and Blood Institute, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of participating CHS investigators and institutions can be found at http://www.chs-nhlbi.org.

References

1. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *Journal of Gerontology: Medical Sciences*. 2001;56A:M146–M156.

2. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of Frailty: characterization in the Women's Health and Aging Studies. *Journal of Gerontology: Medical Sciences*. 2006;61:262–266.

3. Taylor HL, Jacobs DR, Schucker B, Knudsen J, Leon AS, Debacker G. A questionnaire for the assessment of leisure-time physical activities. *Journal of Chronic Diseases*. 1978;31:745–755.

4. Mohr BA, Bhasin S, Kupelian V, Araujo AB, O'Donnell AB, McKinlay JB. Testosterone, sex hormonebinding globulin, and frailty in older men. *Journal of the American Geriatrics Society*. 2007;55:548–555.

5. Purser JL, Kuchibhatla MN, Fillenbaum GG, Harding T, Peterson ED, Alexander KP. Identifying frailty in hospitalized older adults with significant coronary artery disease. *Journal of the American Geriatrics Society*. 2006;54:1674–1681.

6. Gill TM, Gahbauer EA, A HG, Han L. Transitions between frailty states among community-living older persons. *Archives of Internal Medicine*. 2006;166(4):418–423.

7. Faber MJ, Bosscher RJ, Chin A Paw MJ, Wieringen PC. Effects of exercise programs on falls and mobility in frail and pre-frail older adults: a multicenter randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*. 2006;87:885–896.

8. Bartali B, Frongillo EA, Bandinelli S, et al. Low nutrient intake is an essential component of frailty in older persons. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2006;61:589–593.

9. Woods NF, LaCroix AZ, Gray SL, et al. Frailty: emergence and consequences in women aged 65 and older in the Women's Health Initiative Observational Study. *Journal of the American Geriatrics Society*. 2005;53:1321–1330.

10. Puts MT, Lips P, Deeg DJ. Sex differences in the risk of frailty for mortality independent of disability and chronic diseases. *Journal of the American Geriatrics Society*. 2005;53:40–47.

11. Ottenbacher KJ, Ostir GV, Peek M, Snih SA, Raji MA, Markides KS. Frailty in older Mexican Americans. *Journal of the American Geriatrics Society*. 2005;53:1524–1531.

12. Ostir GV, Ottenbacher KJ, Markides KS. Onset of frailty in older adults and the protective role of positive affect.

Psychology and Aging. 2004;19:402–408.

13. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Annals of Epidemiology*. 1991;3:263–276.

14. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*. 1993;25:71–80.

15. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of Physical Activities: an update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise*. 2000;32:S498–S516.

16. van Stralen KJ, Doggen CJ, Lumley T, et al. The relationship between exercise and risk of venous thrombosis in elderly people. *Journal of the American Geriatrics Society*. 2008;56:517–522.

17. McCullagh P, Nelder JA. Generalized linear models. *Chapman & Hall/CRC*. 1989.

18. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143:29–36.

19. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. *Springer Verlag*. 2001.

20. Fried LP, Kronmal RA, Newman AB, et al. Risk factors for 5-year mortality in older adults: the Cardiovascular Health Study. *Journal of the American Medical Association*. 1998;279:585–592.

21. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974;19:716–723.

22. Eckel SP. Quantifying individual and city level modification of the health effects of air pollution in older adults. PhD Thesis. Johns Hopkins University. 2009.

23. Sallis JF, Haskell WL, Wood PD, et al. Physical activity assessment methodology in the five-city project. *American Journal of Epidemiology*. 1985;121:91–106.

24. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). *Medical Care*. 1992:473–483.

25. Washburn RA, Smith KW, Jette AM, Janney CA. The physical activity scale for the elderly (PASE): development and evaluation. *Journal of Clinical Epidemiology*. 1993;46:153–162.

26. Stel VS, Smit JH, Pluijm SMF, Visser M, Deeg DJH, Lips P. Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. *Journal of Clinical Epidemiology*. 2004;57:252–258.

27. Anderson GL, Manson J, Wallace R, et al. Implementation of the Women's Health Initiative study design. *Annals of Epidemiology*. 2003;13:5–17.

Figure 1: Relative influence of kilocalorie/week expenditures from the 15 leisure time activities, as determined by logistic regression for: (a) all participants, (b) men and (c) women.

Figure 2: Association of baseline frailty status with mortality over 3 years of follow up, according to different definitions of low physical activity. Proportional hazards models estimate the raw association (gray) and the covariate adjusted association (black). Hazard ratios compare prefrail to robust participants (dashed 95% confidence intervals) and frail to robust participants (solid 95% confidence intervals). Thin vertical lines demarcate point estimates from the original criterion.

Figure 1

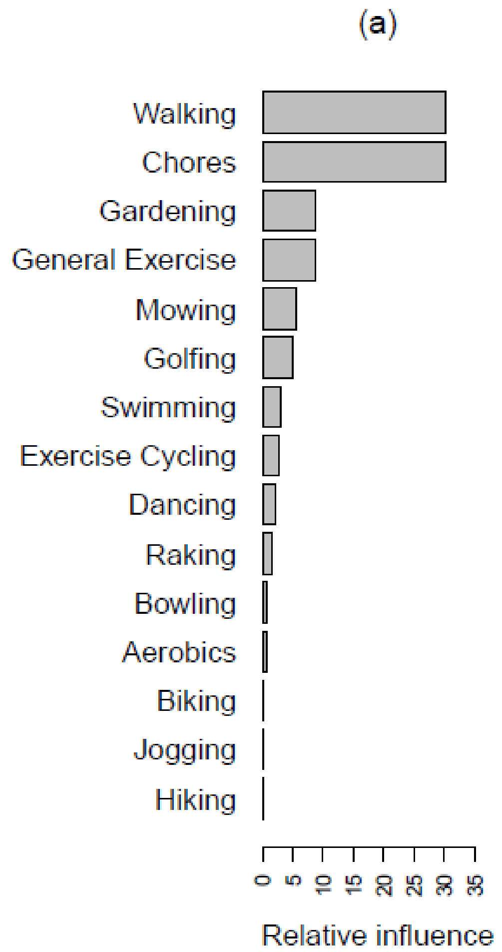
(a)
Walking
Chores
Gardening
General Exercise
Mowing
Golfing
Swimming
Exercise Cycling
Dancing
Raking
Bowling
Aerobics
Biking
Jogging
Hiking
0 5 10 15 20 25 30 35
Relative influence
(b)
Walking
Chores
Mowing
Golfing
Gardening
General Exercise
Exercise Cycling
Swimming
Raking
Dancing
Biking
Bowling
Aerobics
Jogging
Hiking
0 5 10 15 20 25 30 35
Relative influence
(c)
Chores
Walking
General Exercise
Gardening
Swimming
Dancing
Exercise Cycling
Golfing
Bowling
Mowing
Aerobics
Raking
Hiking
Biking
Jogging
0 5 10 15 20 25 30 35
Relative influence

Figure 2

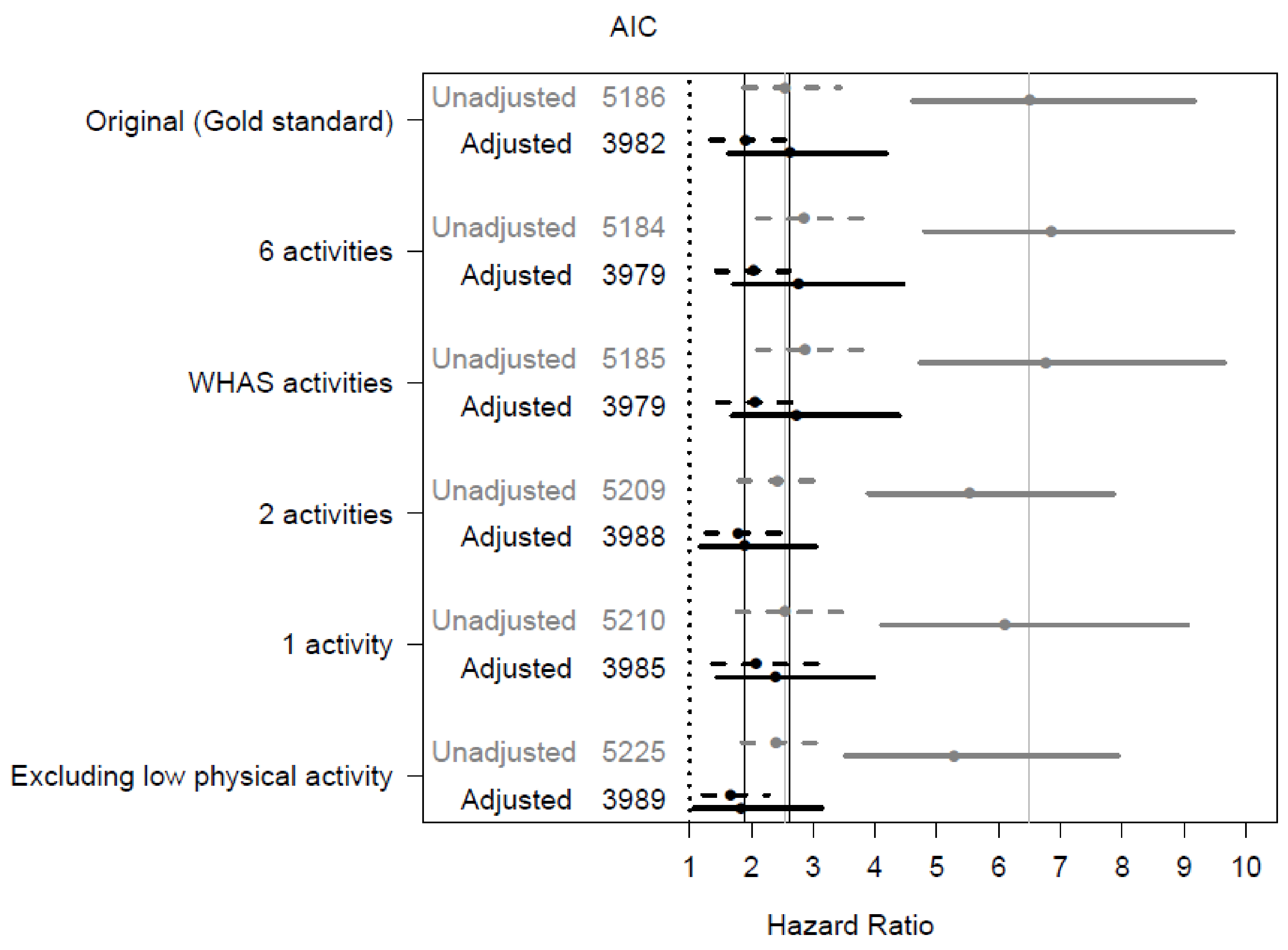
AIC
Original (Gold standard)
6 activities
WHAS activities
2 activities
1 activity
Excluding low physical activity
Unadjusted 5186
Adjusted 3982
Unadjusted 5184
Adjusted 3979
Unadjusted 5185
Adjusted 3979
Unadjusted 5209
Adjusted 3988
Unadjusted 5210
Adjusted 3985
Unadjusted 5225
Adjusted 3989
1 2 3 4 5 6 7 8 9 10
Hazard Ratio

Table 1: Leisure time activities over the 2 weeks prior to baseline, summarized by percent participation and the median and interquartile range (IQR) of kilocalorie/week expenditure for those who participated.

	% Participation in activity			Kilocalorie/week expenditure on activity					
	All Participants (N=5293)	Men (N=2224)	Women (N=3069)	All Participants Median	(IQR)	Men Median	(IQR)	Women Median	(IQR)
Chores	63.3	51.3	72.0	242	(480)	191	(315)	286	(549)
Walking	55.8	58.1	54.2	388	(671)	530	(824)	311	(544)
Gardening	26.8	31.0	23.8	262	(656)	385	(878)	190	(427)
General Exercise	22.5	22.4	22.5	262	(385)	306	(432)	230	(355)
Mowing	21.4	40.2	7.8	237	(315)	263	(319)	144	(232)
Raking	16.8	23.2	12.1	70	(146)	82	(163)	59	(123)
Exercise Cycling	12.3	12.3	12.3	142	(287)	164	(318)	122	(270)
Dancing	7.2	8.0	6.6	169	(618)	172	(749)	169	(485)
Golfing	7.1	12.9	2.9	1535	(2147)	1742	(2310)	1052	(1389)
Swimming	5.1	6.1	4.4	270	(658)	353	(764)	214	(515)
Bowling	3.4	3.8	3.1	444	(461)	473	(566)	441	(415)
Biking	2.4	4.2	1.1	318	(758)	385	(835)	161	(251)
Aerobics	2.3	0.9	3.3	432	(618)	575	(1292)	423	(593)
Hiking	2.0	3.2	1.1	280	(507)	335	(769)	241	(344)
Jogging	0.9	1.4	0.5	459	(670)	680	(863)	335	(461)

Table 2: Cross-validated area under the ROC curve (AUC) for potential surrogate models at baseline, indexed by the number (#) of activities included. The subset of activities included depends on the corresponding relative influence ranking (all participants, men or women).

# Activities	All Participants	Men	Women
15	0.990	0.976	0.995
12	0.989	0.981	0.994
10	0.987	0.977	0.992
8	0.982	0.969	0.989
6	0.975	0.957	0.986
5	0.969	0.939	0.981
4	0.960	0.917	0.980
3	0.944	0.895	0.967
2	0.920	0.835	0.952
1	0.753	0.751	0.849

Table 3: Area under the ROC curve (AUC) at baseline (cross-validated) and at follow up for selected surrogates.

	WHAS Activities		6 Activities		2 Activities		1 Activity	
Study Year	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2	Cohort 1	Cohort 2
1989-1990	0.978		0.976		0.921		0.765	
1992-1993	0.976	0.977	0.977	0.965	0.918	0.898	0.768	0.689
1996-1997	0.971	0.978	0.972	0.966	0.898	0.907	0.759	0.761