



Published in final edited form as:

*Med Sci Sports Exerc.* 2014 August ; 46(8): 1570–1577. doi:10.1249/MSS.0000000000000278.

## Convergent Validity of a Brief Self-reported Physical Activity Questionnaire

Kelley Pettee Gabriel<sup>1</sup>, Stephen Sidney<sup>2</sup>, David R. Jacobs Jr.<sup>3</sup>, Charles P. Quesenberry Jr.<sup>2</sup>, Jared P. Reis<sup>4</sup>, Sheng-Fang Jiang<sup>2</sup>, and Barbara Sternfeld<sup>2</sup>

<sup>1</sup>Division of Epidemiology, Human Genetics and Environmental Sciences; University of Texas Health Science Center in Houston, School of Public Health – Austin Regional Campus; Austin, TX

<sup>2</sup>Division of Research, Kaiser Permanente Northern California; Oakland, CA

<sup>3</sup>Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN

<sup>4</sup>Division of Cardiovascular Sciences, Program in Prevention and Population Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD

### Abstract

**Purpose**—To determine if summary estimates of a self-report physical activity questionnaire that does not specifically assess frequency or duration [the Coronary Artery Risk Development in Young Adults (CARDIA) Physical Activity History (PAH)] differs from the summary estimates of one that does (CARDIA Supplemental Questionnaire).

**Methods**—Following the Year 25 exam (2010–11), 203 CARDIA black and white men and women (aged  $50.3 \pm 3.6$  years) at the Oakland, CA site participated in this comparison study. The between-questionnaire association and agreement was determined for continuous and categorical estimates based on (1) quartiles and (2) meeting 2008 Physical Activity Guidelines. Differences in participant characteristics by concordance/discordance status were also examined. Finally, receiver operating characteristic (ROC) curves were computed to determine the accuracy of the PAH compared to the supplemental questionnaire.

**Results**—Reported physical activity levels were high and varied significantly by race and sex (all  $p < 0.01$ ). Between-questionnaire estimates were significantly correlated ( $\rho = 0.75$  to  $0.90$ ; all  $p < 0.001$ ) and had high agreement ( $\kappa = 0.51$  to  $0.80$ ) across all race/sex groups. A higher proportion of women than men were classified as concordant by quartile of vigorous intensity ( $p = 0.001$ ), but no other participant characteristics were associated with concordant/discordant quartile ranking. Participants classified as concordant based on PA guidelines had lower body mass index than those classified as discordant (both  $p < 0.05$ ). The AUC was 0.95 suggesting that the PAH has high accuracy for classifying individuals as meeting physical activity guidelines.

---

Corresponding Author: Kelley Pettee Gabriel, Assistant Professor, University of Texas Health Science Center at Houston, School of Public Health - Austin Regional Campus, Division of Epidemiology, Human Genetics, and Environmental Sciences, 1616 Guadalupe Street, Suite 6.300, Austin, TX USA 78701, Office: 512-391-2525, Fax: 512-482-6185, Kelley.P.Gabriel@uth.tmc.edu.

**Conflict of Interest/Financial Disclosure:** None

**Conclusions**—Although it is inconvenient that the PAH is not expressed in more standard units, these findings support the practice of not directly assessing frequency and duration, which are frequent sources of reporting error.

### Keywords

self-report; convergent validity; physical activity assessment; components of physical activity behavior

---

## INTRODUCTION

Physical activity estimates obtained from questionnaires are typically the product of participant reported information related to: [1] intensity level obtained via specific activities or within broad intensity categories, [2] duration and [3] frequency of physical activity completed within a pre-specified recall period (15). Together, these physical activity characteristics provide an estimate of physical activity dose or volume. Physical activity dose can be further described in a number of ways, including activity dose within specific domains (leisure-time, occupation, transportation and domestic/self-care) or intensity categories (moderate- and/or vigorous- intensity physical activity). Physical activity dose estimates are also often used to classify respondents based on meeting (or not meeting) current public health guidelines for physical activity (16).

The Physical Activity History (PAH) questionnaire was developed for the Coronary Artery Risk Development in Young Adults (CARDIA) study, an on-going prospective study of the development of cardiovascular disease risk and subclinical atherosclerosis in black and white men and women aged 18–30 years at baseline (1985–86), to provide reliable information on typical, rather than recent, physical activity. Estimates of total activity, as well as moderate and vigorous intensity physical activity, can be computed (6). However, the PAH does not specifically address frequency (in terms of frequency per week or month) or duration (in terms of actual duration per session or week) of physical activity, two necessary components for estimating physical activity dose in minutes or MET-minutes. Questions related specifically to these characteristics were excluded from the PAH based on the notion that they were major contributors to measurement error, which could ultimately lead to information bias and threats to internal study validity. Also, given the limited amount of time scheduled for exam visits, it was deemed less important to collect these less reliable data. Although previous studies have shown that the derived summary estimate variables from the PAH, expressed as Exercise Units (EU), demonstrate reasonable reliability and validity for ranking individuals (7), it is unknown whether EUs are reflective of actual physical activity volume. This inability to directly quantify physical activity dose greatly impacts the ability to characterize individual- and population-level physical activity levels in CARDIA. Further, the use of EUs greatly limits the ability to discuss adequately the clinical and public health significance of physical activity- related CARDIA findings.

In direct response to these concerns, an ancillary study was conducted at Year 25 among a sub-sample of CARDIA participants at the Kaiser Permanente (Oakland, CA) field center which added a more detailed supplemental physical activity questionnaire. The overall

objective of this ancillary study was to compare the summary estimates obtained from the PAH and the supplemental questionnaire to gain a better understanding of the degree of error in the PAH, due to the omission of specific quantification of frequency and duration. To accomplish this goal, the summary physical activity estimates obtained from the two questionnaires were compared, using a variety of approaches to determine the correlation, agreement, and predictive accuracy of the PAH using the supplemental questionnaire as the criterion measure.

## METHODS

### Design Overview and Study Participants

CARDIA includes 5,115 adults who were between the ages of 18 and 30 years at the baseline examination in 1985–86. The CARDIA study has been described in detail previously (3). Briefly, participants were recruited from four geographical locations (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA) and were re-examined 2 (1987–88), 5 (1990–91), 7 (1992–93), 10 (1995–96), 15 (2000–01), 20 (2005–06) and 25 (2010–11) years after baseline (3). Beginning about two-thirds through the Year 25 exam, participants at the Oakland, CA site were approached sequentially to participate in this ancillary study until the desired sample size (i.e.,  $n=50$  from each of the four race and sex groups) or end of the exam cycle was achieved. These supplemental physical activity questions were asked after all other core exam components, including the PAH had been completed. The administration of the supplemental physical activity questions took place approximately 2–3 hours after the PAH. All participants provided informed consent and the institutional review board at each CARDIA field center and coordinating center approves the study annually.

### Data Collection

**Participant Characteristics**—At the Year 25 core exam, standardized questionnaires were used to assess participant characteristics including, age, educational attainment (high school degree, associate’s degree, or bachelor’s degree), marital status (married/living with a significant other or other), children or step children (yes or no), full-time job (yes or no), home owner (yes or no), difficulty paying for basic needs (not very hard or other), and self-rated health (fair/poor, good, or excellent). Height (stadiometer) and weight (balance-beam scale) were measured with the participant lightly clothed and without shoes (4) and body mass index (BMI) was calculated as weight (kg) divided by height ( $m^2$ ).

**Self-reported Physical Activity**—The CARDIA Physical Activity History (PAH) is an interviewer-administered questionnaire designed to assess usual physical activity levels (6, 14). Participants reported the total number of months during the past year that they participated in 13 different categories or activities (12 leisure time, one occupational). Duration of activity was collected by asking how many months each class of activity had been performed for at least one hour during the month. For each activity that was performed for at least one hour in at least one month, respondents were asked for the number of months s/he performed the activity frequently. The term “frequently” or “frequent participation” was operationalized individually for each activity and ranged from 2 to 5 hours per week. The

PAH was scored in EUs which represents a weighted sum, based on intensity of activity (defined in metabolic equivalents ranging from 3 to 8) and number of months of less frequent participation (months during which the activity was performed for at least 1 hour, but less than the specified hours per week), plus three times the number of months of “frequent participation” (14). Three-times the number of months was selected based on the notion that the duration of activity in months with more frequent participation is higher than the duration in months with less frequent participation (6). The total physical activity summary estimate reflects all 13 included activities or activity categories. The vigorous intensity summary estimate includes: [1] jogging or running, [2] vigorous racket sports, [3] bicycling faster than 10 miles per hour or exercising hard on the exercise bike, [4] swimming, [5] vigorous exercise class or dancing, [6] home or leisure activity (e.g., shoveling snow, moving heavy objects, or weight lifting), [7] vigorous job activity (e.g., lifting, carrying, or digging), and [8] strenuous sports (e.g., basketball, football, skating, or skiing). The moderate intensity summary score includes: [9] non-strenuous sports (e.g., softball, shooting baskets, volleyball, ping pong, or leisurely jogging, swimming, or biking), [10] take walks or hikes or walk to work, [11] bowling or golf, [12] home exercises, calisthenics and [13] home maintenance or gardening (e.g., carpentry, painting, raking, or mowing). The PAH is available on the CARDIA website (<http://www.cardia.dopm.uab.edu/>) (3).

The structure of the supplemental questionnaire was consistent with the PAH. For each of the 13 categories or activities included on the PAH in which a participant engaged in during the previous year, additional information was related to the frequency (days per week) and duration (minutes each time). Using jogging or running as an example, for “frequent participation” the following questions were asked, “for the XX months that you jogged or ran for at least 2 hours per week, how many days per week did you do it? For how many minutes did you do it each time?”. For “infrequent participation”, the question wording changed slightly, “for the XX months that you jogged or ran for at least an hour total time in the month, how many times per month or per week did you do it? For how many minutes did you do it each time?”. Physical activity estimates were scored as the product of reported frequency and duration ( $\text{min}\cdot\text{wk}^{-1}$ ) and summed across all 13 categories or activities (total physical activity) and within vigorous- (i.e., items 1–8 above) and moderate- (items 9–13 above) intensity categories.

## Statistical Methods

First, univariate analyses were conducted on measured parameters and all variables were assessed for normality. Normally distributed variables were reported as mean and standard deviation, non-normally distributed variables as medians with 25<sup>th</sup> and 75<sup>th</sup> percentiles, and proportions were noted for categorical variables. Differences in measured parameters between race and sex groups were determined using analysis of variance (ANOVA), Kruskal-Wallis, or chi-square tests. Second, Spearman rank order correlation coefficients were used to examine the between-questionnaire correlation between continuous [1] moderate-intensity, [2] vigorous-intensity, and [3] total physical activity summary estimates in the entire analytic sample and by race and sex groups. Third, the physical activity summary scores from both questionnaires were categorized based on quartiles (of entire

analytic sample) and agreement was assessed via weighted kappa statistics in the entire analytic sample and by race and sex groups. Fourth, for each physical activity summary estimate, participants were then classified as concordant if they were categorized in the same quartile for both questionnaires and discordant if quartile rankings differed by questionnaire. Differences in participant characteristics were then determined by concordance status using chi-square tests or Student t-tests. Fifth, participants were classified as meeting 2008 Physical Activity Guidelines (16) based on the total physical activity estimate (moderate-plus vigorous- intensity), which was defined for the PAH as 300 EUs (2, 11) and as 150 minutes per week from the supplemental questionnaire. The 300 EU threshold was first proposed by Parker et al. (11) to approximate the American College of Sports Medicine recommendations for the amount of physical activity needed to support weight loss (i.e., 1,500 kcals per week) (12). Participants were then identified as concordant if they were classified as meeting or not meeting physical activity guidelines on both questionnaires and discordant if they were classified as meeting or not meeting guidelines with one questionnaire, but not the other. Differences in participant characteristics were determined by concordance status using chi-square tests or Student t-tests. Finally, receiver operating characteristic (ROC) curves where sensitivity (y-axis) is plotted as a function of a false-positive rate (1 – specificity; x-axis) were used to evaluate the 300 EU threshold value. Here, the supplemental questionnaire served as the criterion measure. The area under the curve (AUC) was also computed. AUC values range from 0 to 1, with a value closer to 1 indicating a perfectly accurate tool and 0.5 indicative of what one would expect due to chance (10).

## RESULTS

The analytic sample (n=203) represents approximately 5.8% and 26.2% of all CARDIA- (n=3,499) and Oakland site- (n=774) participants that attended the Year 25 exam, respectively. There were no significant differences in sex, race, age, or BMI between the analytic sample and other CARDIA participants at Year 25. However, the analytic sample had significantly higher reported moderate intensity and total physical activity levels from the PAH at Year 25 than other CARDIA participants [moderate: median (25<sup>th</sup>, 75<sup>th</sup> percentile): 147.0 (72.0, 228.0) EUs versus 116.0 (48.0, 196.0) EUs; total: median (25<sup>th</sup>, 75<sup>th</sup> percentile): 312.0 (156.0, 540.0) EUs versus 273.0 (124.0, 484.0) EUs; both p<0.01]. Further, educational attainment (p=0.054) and reported vigorous intensity activity (p=0.058) were higher in the analytic sample when compared to other participants; however, these differences were borderline significant. There were no significant differences in any of these factors between the analytic sample and Oakland site participants.

The descriptive characteristics of the analytic sample (n=203), including differences by race and sex group, are shown in Table 1. Of the 203 participants, 37 (18.2%) were black men (BM), 49 (24.1%) were white men (WM), 60 (29.6%) were black women (BW), and the remaining (n=57; 28.1%) were white women (WW). The Year 25 mean  $\pm$  standard deviation of age and BMI was 50.3  $\pm$  3.6 years and 30.6  $\pm$  7.3 kg/m<sup>2</sup>, respectively, with WM being the oldest (p=0.04) and BW having the highest BMI (p<0.001). In general, most participants had a bachelor's degree or higher (55.2%), were married or living with a significant other (68.5%), had children or stepchildren (78.3%), were home owners (76.4%), and reported

excellent health status (53.7%). There were also significant differences in educational attainment, marital status, full-time job and home owner status, and level of difficulty paying for basic needs by race and sex groups (all  $p < 0.05$ ).

### Physical Activity Scores as Continuous Estimates

Self-reported physical activity estimates from both questionnaires for the entire analytic sample and after stratification by race and sex are also shown in Table 1. For both questionnaires, there were significant differences in all three physical activity summary estimates by race and sex group (all  $p < 0.01$ ; Table 1). BW reported the lowest physical activity levels on all summary estimates, regardless of questionnaire. WM reported the highest level of vigorous intensity physical activity on both questionnaires, but WW reported the highest values for moderate intensity and total physical activity on the PAH while WM reported the highest level for those variables on the supplemental questionnaire only.

The between-questionnaire correlation coefficients for all three physical activity estimates ranged from  $\rho = 0.79$  to  $0.86$  (all  $p < 0.001$ ) (Table 2), and the strength of these associations did not vary substantially by race and sex (all  $p < 0.001$ ).

### Physical Activity Scores as Categorical Estimates

**Physical Activity Quartiles**—Quartile cut-point values per physical activity estimate are shown in Table 1 (see median and 25<sup>th</sup> and 75<sup>th</sup> percentile values). Weighted kappa statistics are shown in Table 3, in the entire analytic sample and then stratified by race and sex groups. In all participants, the weighted kappa statistics showing level of agreement between questionnaires for classifying participants based on quartiles of physical activity ranged from  $\kappa = 0.60$  to  $0.65$ , suggesting moderate agreement between questionnaires (Table 3). Although there were differences by race/sex group, all values suggested moderate or higher agreement between questionnaires. WM had the lowest agreement for moderate intensity and total activity ( $\kappa = 0.54$  and  $\kappa = 0.51$ , respectively), while BM had the highest agreement for moderate intensity ( $\kappa = 0.68$ ) and BW had the highest agreement for total activity ( $\kappa = 0.65$ ). Further, BM and BW had the lowest and highest agreement for vigorous intensity categories ( $\kappa = 0.52$  and  $\kappa = 0.80$ , respectively).

Table 4 shows differences in participant characteristics by concordance status based on quartiles of the sample distribution for each of the physical activity estimates. A significantly greater proportion of women versus men had concordant results when vigorous intensity physical activity was categorized into quartiles ( $p < 0.001$ ). There were no significant differences by concordance status for any other participant characteristic for moderate-, vigorous- intensity, or total activity estimates. The proportion with children or step children by concordance status was of borderline statistical significance for the total activity estimate ( $p = 0.10$ ).

**Meeting Physical Activity Guidelines**—Table 5 shows differences in participant characteristics based on meeting guidelines for physical activity. Again, there were no significant differences by concordance status for most participant characteristics. However, BMI was significantly lower among those classified as concordant versus those with



discordant results for meeting guidelines ( $p = 0.02$ ). Also, the difference in self-rated health status by concordance status was of borderline statistical significance ( $p=0.054$ ), with those classified as concordant having a higher proportion of respondents reporting good or excellent health.

Based on ROC curve analyses, the ability of the PAH to classify participants as meeting physical activity guidelines (16) (as determined by the supplemental questionnaire) was high (AUC = 0.95; Figure 1). The accuracy of the PAH was highest in BM (AUC=0.98), followed by WM and BW (AUC=0.97 and 0.92, respectively), and was lowest in WW (AUC=0.92). In general, the sensitivity was lower, while the specificity was higher with greater PAH scores. The 300 EU threshold used in previous CARDIA analyses (2, 11) to define meeting guidelines had an associated predicted probability of 98.5% and sensitivity and specificity of 64.5% and 97.1%, respectively. At 150 EUs, the predicted probability was 70.3%, sensitivity 88.2%, and specificity 76.5%. At a score of 500 EUs, the predicted probability and specificity was 100%; however, the sensitivity was 0% (Figure 1).

## DISCUSSION

This study provided a practical example of the potential measurement issues that could arise when primary components of a summary score, reflecting physical activity volume or dose, are not directly quantified. However, findings from the current study generally support the ability of the CARDIA PAH, a self-report measure that does not require respondents to directly quantify duration or frequency, to provide reflective physical activity summary estimates when compared to a similarly structured, self-reported instrument that does prompt respondents to directly recall these physical activity characteristics. These analyses included a thorough evaluation of both the continuous and categorical physical activity estimates based on sample-determined quartiles and meeting physical activity guidelines (16).

These findings have several key implications for physical activity measurement by self-report. First, reported total activity levels, a combination of moderate- and vigorous-intensity summary scores, exceeded the 2008 Physical Activity Guidelines for Americans (16) for a large proportion of participants on both questionnaires. For the supplemental questionnaire, for example, the median reported total physical activity levels were over two-times [i.e., 381.0 (210.0, 716.0)  $\text{min}\cdot\text{wk}^{-1}$ ] the recommended physical activity level necessary to achieve health benefits. Similarly, using 300 EUs as a cut-point roughly equivalent to meeting physical activity guidelines (2, 11), the median PAH total physical activity estimate of 312.0 (156.0, 540.0) EUs is also suggestive of a highly active study sample, although to a lesser magnitude than that of the supplemental questionnaire. More specifically, 83.3% reported meeting physical activity guidelines by the supplemental questionnaire compared to only 51.2% with the PAH. While dependent on the selection of a threshold value to infer meeting guidelines (16), these findings suggest that the tendency to over-report physical activity may be amplified when respondents are asked to specifically recall and report frequency and duration and supports the initial concerns of CARDIA investigators during the development of the PAH.

It has been suggested that individuals recall activity frequency via two primary cognitive processes: episode enumeration or rate based estimation (5). Episode enumeration requires individuals to retrieve all discrete events within a specified time period and then count these episodes. The likelihood of using this technique for information retrieval is reduced with longer recall periods and more frequent activities. For this, rate-based estimation is often employed. Here, individuals estimate activity frequency based on usual participation and then multiply this estimated frequency by the length of the recall time period, which could result in either an over- or under- estimation of frequency. Given that the questionnaires utilize a past-year recall time frame, it is possible that participants used rate-based estimation and, in turn, over-estimated frequency. Recalling physical activity duration can be equally challenging. With the supplemental questionnaire, individuals were asked to recall the number of minutes s/he engaged in a particular activity each time. However, with this phrasing, some participants may have reported the corresponding duration from the most recent activity event; whereas, others may report an average duration from a series of discrete events that either occurred recently or sometime in the past, which could contribute to overall measurement error.

Second, the between-questionnaire association and agreement for the vigorous intensity estimate was higher than that for either moderate or total activity, which may be a result of the more structured nature of activities included in this estimate (14). It is well-supported that individuals are more able to accurately recall higher intensity, structured activities than lower intensity, unstructured activities (1, 9, 13). Vigorous intensity activities often have more noticeable physiological cues (e.g., rapid heart rate, sweating) than lower intensity activities, which may better facilitate memory retrieval (5). Further, men tend to overestimate participation in vigorous intensity activities when compared to women (8). In the current study, the between-questionnaire correlation coefficients and agreement were lower in men. This finding suggests that this tendency to over-report in men may be due to requesting specific quantitative information related to duration and frequency. The highest observed between-questionnaire association and agreement for the vigorous intensity estimate was found in BW. However, BW in our study had the lowest reported vigorous intensity levels of all the race and sex groups. Therefore, this suggests that individuals are more accurate when reporting “no” versus “any” activity participation.

Third, as shown with the categorical analysis (based on quartiles or meeting physical activity guidelines), the PAH does a reasonable job of ranking or categorizing individuals’ physical activity levels when compared to the supplemental questionnaire. Further, the PAH works as well as the supplemental questionnaire at ranking or categorizing individuals’ physical activity levels for participants of varying characteristics, with a few notable exceptions. For the quartile-specific analysis, women had a higher concordance rate for the vigorous intensity activity than men. Again, this was likely due to the sex-related differences with regards to engaging in and reporting vigorous intensity physical activity that were detailed above. However, when participants were cross-categorized based on meeting physical activity guidelines, BMI significantly differed by concordance status. More specifically, those classified as concordant had a lower BMI when compared to those categorized as discordant. These findings may be due to the fact that BMI is associated with physical



activity (16). Behaviors that are performed more routinely are often less cognitively challenging to recall when compared to more sporadic activities.

Fourth, the ROC curve analysis showed that the PAH has high accuracy when classifying participants as meeting or not meeting physical activity guidelines and that this precision did not vary substantially by sex or race. In physical activity related research, it is often advantageous to categorize individuals based on meeting guidelines. This is a challenging task when related summary estimates are not expressed as minutes per week, which is the case for the PAH. Previous CARDIA analyses have used a threshold of 300 EUs to define meeting guidelines (2, 11); however, this cut-point had not been rigorously evaluated until now. Since individuals tend to over-report physical activity levels (15), the most appropriate PAH score should maximize specificity rather than sensitivity. Specificity, in this context, would be operationalized as the proportion of respondents that did not meet physical activity guidelines who were correctly identified. In the current study, 300 EUs was associated with a sensitivity and specificity of 64.5% and 97.1%, respectively when compared to the supplemental questionnaire. Interestingly, the 250 EU cut-point was also associated with a specificity of 97.1%; however, sensitivity was higher at 70.4%. Given the nature of self-reported physical activity data, the 300 EU threshold provides a more conservative estimate of meeting physical activity guidelines and supports its continued use in future physical activity-related CARDIA analyses.

There are several limitations to consider when interpreting the results of the current study. First, accelerometer derived estimates of physical activity were not available at the Year 25 exam. Therefore, we are unable to determine if the strength of the association with the accelerometer estimates differs substantially by questionnaire type. Further, while the supplemental questionnaire was used as the criterion measure, it has never been formally evaluated for reliability or validity. Therefore, an important area for future research would involve a concurrent evaluation of the PAH and supplemental questions with accelerometer estimates. Second, the current analyses represented a small proportion of CARDIA participants that attended the 25 Year exam and only participants at the Oakland, CA field site were invited to participate. Therefore, results of current study may not be generalizable to other populations. While the analytic sample was more physically active than other CARDIA participants at Year 25, very few other statistically significant differences were noted, which enhances the overall external validity of these findings. Finally, the supplemental questions were always asked at the end of the core exam, during the same visit as the PAH. It is possible that participants could have recalled additional activities during the interim, which could have caused the derived estimates from the supplemental questionnaire to be inflated. Likewise, the relatively short interval between questionnaire administrations could have also resulted in the high concordance rate.

In summary, the results of these extensive analyses suggest that the PAH performs quite favorably when compared to the supplemental questionnaire and does not differ substantially by race and sex groups or by other important participant characteristics (e.g., educational attainment). This study provides evidence for using EUs derived from the PAH in a quantitative way to provide insight into the public health and clinical significance of study

findings and supports the continued use of this brief, self-reported questionnaire to assess physical activity within a large, diverse population-based study of cardiovascular health.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

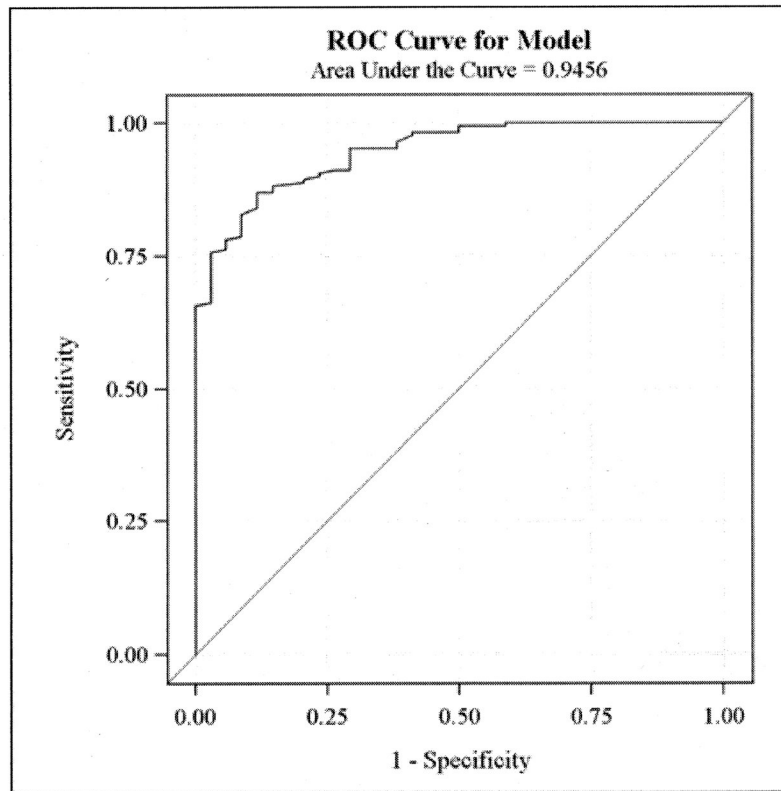
## Acknowledgments

The authors would like to acknowledge the CARDIA participants at the Oakland site that agreed to participate in this ancillary study. The Coronary Artery Risk Development in Young Adults Study (CARDIA) is supported by contracts HHSN268201300025C, HHSN268201300026C, HHSN268201300027C, HHSN268201300028C, HHSN268201300029C, and HHSN268200900041C from the National Heart, Lung, and Blood Institute (NHLBI), the Intramural Research Program of the National Institute on Aging (NIA), and an intra-agency agreement between NIA and NHLBI (AG0005). The results of the present study do not constitute endorsement by the American College of Sports Medicine.

## References

1. Ainsworth BE, Caspersen CJ, Matthews CE, Masse LC, Baranowski T, Zhu W. Recommendations to improve the accuracy of estimates of physical activity derived from self report. *J Phys Act Health*. 2012; 9(Suppl 1):S76–84. [PubMed: 22287451]
2. Anderssen N, Jacobs DR Jr, Sidney S, Bild DE, Sternfeld B, Slattery ML, Hannan P. Change and secular trends in physical activity patterns in young adults: a seven-year longitudinal follow-up in the Coronary Artery Risk Development in Young Adults Study (CARDIA). *Am J Epidemiol*. 1996; 143(4):351–362. [PubMed: 8633619]
3. Coronary Artery Risk Development in Young Adults (CARDIA) Study Website [Internet]. Birmingham, AL: CARDIA Coordinating Center; [cited 2013 Sept 26]. Available from: <http://www.cardia.dopm.uab.edu/>
4. Cutter GR, Burke GL, Dyer AR, Friedman GD, Hilner JE, Hughes GH, Hulley SB, Jacobs DR Jr, Liu K, Manolio TA, et al. Cardiovascular risk factors in young adults. The CARDIA baseline monograph. *Control Clin Trials*. 1991; 12(1 Suppl):1S–77S. [PubMed: 1851696]
5. Durante R, Ainsworth BE. The recall of physical activity: using a cognitive model of the question-answering process. *Med Sci Sports Exerc*. 1996; 28(10):1282–1291. [PubMed: 8897386]
6. Jacobs DR. Validity and reliability of short physical activity history: CARDIA and the Minnesota Heart Health Program. *Journal of Cardiopulmonary Rehabilitation*. 1989; 9:448–59.
7. Jacobs DR Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc*. 1993; 25(1):81–91. [PubMed: 8423759]
8. Klesges RC, Eck LH, Mellon MW, Fulliton W, Somes GW, Hanson CL. The accuracy of self-reports of physical activity. *Med Sci Sports Exerc*. 1990; 22(5):690–697. [PubMed: 2233209]
9. Kriska AM, Caspersen CJ. Introduction to a Collection of Physical Activity Questionnaires. *Med Sci Sports Exerc*. 1997; 29(6):S5–S9.
10. Park SH, Goo JM, Jo CH. Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean J Radiol*. 2004; 5(1):11–18. [PubMed: 15064554]
11. Parker ED, Schmitz KH, Jacobs DR Jr, Dengel DR, Schreiner PJ. Physical activity in young adults and incident hypertension over 15 years of follow-up: the CARDIA study. *Am J Public Health*. 2007; 97(4):703–709. [PubMed: 17329668]
12. Pollock ML, Gaesser GA, Butcher J, Despres JP, Dishman RK, Franklin B, Garber CE. ACSM Position Stand: The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory fitness and muscular fitness, and flexibility in healthy adults. *Medicine & Science in Sports & Exercise*. 1998; 30(6):975–991. [PubMed: 9624661]

13. Sallis JF, Haskell WL, Wood PD, Fortmann SP, Rogers T, Blair SN, Paffenbarger RS Jr. Physical activity assessment methodology in the Five-City Project. *Am J Epidemiol.* 1985; 121(1):91–106. [PubMed: 3964995]
14. Sidney S, Jacobs DR Jr, Haskell WL, Armstrong MA, Dimicco A, Oberman A, Savage PJ, Slattery ML, Sternfeld B, Van Horn L. Comparison of two methods of assessing physical activity in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Epidemiol.* 1991; 133(12):1231–1245. [PubMed: 2063831]
15. Troiano RP, Pettee Gabriel KK, Welk GJ, Owen N, Sternfeld B. Reported physical activity and sedentary behavior: why do you ask? *J Phys Act Health.* 2012; 9(Suppl 1):S68–75. [PubMed: 22287450]
16. U.S. Department of Health and Human Services. Physical Activity Guidelines for Americans. [Internet]. Washington, D.C: U.S. Department of Health and Human Services; 2008. [cited 2013 September 26]. Available from: [www.health.gov/paguidelines](http://www.health.gov/paguidelines) [Accessed October 10]



PAH Score	Predicted Probability	Sensitivity	Specificity
100	0.439	95.3	61.8
150	0.703	88.2	76.5
200	0.877	81.1	91.2
250	0.956	70.4	97.1
<b>300</b>	<b>0.985</b>	<b>64.5</b>	<b>97.1</b>
500	1.000	0.0	100.0

**Figure 1.** Receiver operating characteristic (ROC) curves comparing the CARDIA Physical Activity History (PAH) to the Supplemental Questionnaire in a sub-group of CARDIA participants (n=203) at the Year 25 Exam