ARIC Manuscript Proposal # 1801

PC Reviewed: 6/14/11  Status: A  Priority: 2
SC Reviewed: _________  Status: ____  Priority: ____

1.a. Full Title: Genome-Wide Association Study of Sedentary Behavior

b. Abbreviated Title (Length 26 characters): GWAS of Sedentary Behavior

2. Writing Group:
Marcel den Hoed, Ph.D., MRC Epidemiology Unit, Cambridge, UK
Keri Monda, Ph.D., University of North Carolina at Chapel Hill
Kari North, Ph.D., University of North Carolina at Chapel Hill
Marilyn Cornelis, Harvard School of Public Health
Other authors welcome

Lead Author**
**This will be a multi-first authored (starred) effort, but our working lead is listed here:
Marcel den Hoed, Ph.D., MRC Epidemiology Unit, the Institute of Metabolic Science

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

First author:
Name: Marcel den Hoed
Address: MRC Epidemiology Unit
Institute of Metabolic Science
Box 285
Addenbrooke’s Hospital
Hills Road, Cambridge, CB2 0QQ, United Kingdom
Telephone: +44 (0)1223 330315
Fax: +44 (0)1223 330316
Marcel.denHoed@mrc-epid.cam.ac.uk

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: Keri Monda
Address: 137 E. Franklin St., Ste 306, Chapel Hill, NC, 27514
Phone: 919-966-8491  Fax: 919-966-9800
E-mail: monda@unc.edu
3. **Timeline:**

*Expected meta-Analysis completion = July 1, 2011; Additional analyses (if necessary) complete = August 1, 2011; Manuscript drafting / submission = September 15, 2011*

4. **Rationale:**

Several studies have recently reported associations of sedentary behaviors (TV viewing, time spent driving a car, time spent sitting) with all-cause and CVD mortality\(^1\)\(^-\)\(^4\). In other studies, sedentary behaviors showed association with obesity, type 2 diabetes, hypertension and incident CVD\(^5\)\(^,\)\(^6\). These associations were generally independent of the total level of physical activity and/or the time spent on moderate and vigorous activities.

Evidence suggests that physical activity level is a heritable trait, with estimates ranging from low to moderate levels of heritability (17-62\%) due to genetic factors\(^7\)\(^-\)\(^13\). Many of the heritability estimates have been derived from twin studies which have found that genetic factors play a role in adult exercise participation\(^10\)\(^,\)\(^14\), suggesting that the remaining heritability is accounted for by environmental factors. A number of genome-wide linkage studies\(^15\)\(^-\)\(^17\) and candidate gene association studies\(^10\)\(^,\)\(^18\)\(^-\)\(^20\) have been conducted as well as one genome-wide association study of exercise behavior\(^21\); however, although a number of suggestive loci have been discovered, no loci have yet been incontrovertibly associated with either physical activity, the time spent on moderate and vigorous intensity activities, or sedentary behavior.

Further preliminary evidence from an ongoing study with objective measures of physical activity and sedentary behavior (Actiheart) in 773 monozygotic and dizygotic twin pairs, initiated by the MRC-Epidemiology Unit, suggests that the heritability of the time spent sedentary (~46\%) is similar to that of physical activity-related energy expenditure (~49\%) and the time spent on moderate and vigorous intensity activities (~43\%). Furthermore, the correlation between the time spent sedentary and the time spent on moderate and vigorous intensity activities is rather low (r=0.27). Together these results suggest that: 1) the inter-individual variation in the time spent sedentary can at least partly be explained by genetic factors; 2) the genetic etiology of sedentary behavior may be largely independent from that of the time spent on moderate and vigorous intensity activities.

So far, no loci have been incontrovertibly associated with either physical activity, the time spent on moderate and vigorous intensity activities, or sedentary behavior. Our aim is to increase our knowledge on the genetic etiology of physical activity related traits and sedentary behavior by means of a large-scale meta-analysis of genome-wide association studies.

5. **Main Hypothesis/Study Questions:**

We hypothesize that an individual’s sedentary behaviors are at least partly genetically determined and driven by an individual’s innate propensity. We propose to
perform a genome-wide association (GWA) study to identify loci associated with sedentary behaviors in individuals of European descent.

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).

Analytic Plan:

<table>
<thead>
<tr>
<th>Exclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Exclude individuals with missing information on physical activity variables of interest or covariates</td>
</tr>
<tr>
<td>• For studies with multiple time points, select baseline/earliest time point</td>
</tr>
<tr>
<td>• Analyses will be restricted to European Americans</td>
</tr>
</tbody>
</table>

Trait Creation

The goal of the GWAS is to capture all aspects of sedentary behavior through the definition of four traits: three dichotomous traits, reflecting sedentary behavior at work, home and during commuting, and one continuous trait: TV viewing. ARIC will contribute data to the analysis of sedentary behavior at work based on the variable RPAA40: V1: FREQ SIT AT WORK.

Sedentary behaviour at work (sedwork) will be defined as follows:

**Sedentary cases (1) vs non-sedentary controls (0)** in individuals who are employed and have data on sitting at work:

- **Sedwork=1:** Individuals who answer 4 (often) or 5 (always) to the above question.
- **Sedwork=0:** Individuals who answer 1 (never), 2 (seldom), or 3 (sometimes) to the above question.

**NOTE:** individuals who specifically report answer 1-3 to the question ‘At work I sit’ 3 will have a sedwork of 0. Do not confuse with individuals with missing information. These individuals should be excluded from the analysis.

Covariate Creation

1. **age:** continuous
2. **sex:** dichotomous, 0=male, 1=female
3. **study_site**
4. **EV1, EV2, EV3:** top eigenvectors will be included to adjust for population substructure
Running the Analyses

An additive genetic model will be used. For imputed data, use allele dosages. Otherwise, use measured genotype. Code reference allele as you wish (each study will document the reference allele in the results file).

- Model 1:
  Logistic regression of sedwork on allele dosage/genotype:
  \[ \text{trait} \sim \mu + \text{SNP} + \text{age} + \text{age}^2 \text{study_site} + \text{EVs} \]

- Model 2:
  Logistic regression of sedwork on allele dosage/genotype:
  \[ \text{trait} \sim \mu + \text{SNP} + \text{age} + \text{age}^2 + \text{study_site} + \text{EVs} + \text{BMI} \]

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>marker</td>
<td>SNP id as rs number.</td>
</tr>
<tr>
<td>chr</td>
<td>Chromosome number. Use symbols X and Y for non-autosomal markers.</td>
</tr>
<tr>
<td>position</td>
<td>Physical position for the reference sequence (indicate build 35/36 in readme file).</td>
</tr>
<tr>
<td>effect_allele</td>
<td>The ‘effect-allele’, also referred to as ‘coded-allele’ or ‘risk-allele’. For example: AA=0, AG=1 and GG=2, the effect_allele is G.</td>
</tr>
<tr>
<td>other_allele</td>
<td>The alternate allele.</td>
</tr>
<tr>
<td>strand</td>
<td>+ or -, representing either the positive/forward strand or the negative/reverse strand of the human genome reference sequence.</td>
</tr>
<tr>
<td>beta</td>
<td>Beta estimate from genotype-phenotype association --‘NA’ if not available.</td>
</tr>
<tr>
<td>stderr</td>
<td>Standard error of beta estimate --‘NA’ if not available.</td>
</tr>
<tr>
<td>pvalue</td>
<td>p-value for test statistic --‘NA’ if not available.</td>
</tr>
<tr>
<td>freq</td>
<td>Allele frequency for the effect_allele --‘NA’ if not available.</td>
</tr>
<tr>
<td>hwe</td>
<td>Exact test Hardy-Weinberg equilibrium p-value—only directly typed SNPs, ‘NA’ for imputed.</td>
</tr>
<tr>
<td>n</td>
<td>Total sample with phenotype and genotype for SNP.</td>
</tr>
<tr>
<td>imputed</td>
<td>1=imputed, 0=if directly typed. Use ‘0’ coding when imputation was used to only ‘fill-in’ the missing genotypes (i.e &lt;10% of subjects/snp).</td>
</tr>
<tr>
<td>oevr_imp</td>
<td>Observed divided by expected variance for imputed allele dosage. Also referred to as ‘Rsq’ in MACH. If some other quality measure was used for imputed SNPs, please describe in readme file. ‘NA’ or 1 if directly genotyped.</td>
</tr>
</tbody>
</table>

7.a. Will the data be used for non-CVD analysis in this manuscript?
   ____ Yes   ____ No
b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = “CVD Research” for non-DNA analysis, and for DNA analysis RES_DNA = “CVD Research” would be used?  N/A

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

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = “No use/storage DNA”?  __X__ 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/ARIC/search.php  __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)?  
   - None

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  
   __X_ A. primarily the result of an ancillary study (list number #2006.03 & 2007.02)

12. 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.

   Understood, and we will meet this deadline

References


