Recalled Early Life Socioeconomic Status (SES): Assessment of Data Quality and Estimates of Impact of Recall Error on Associations with Cardiovascular Disease

b. Abbreviated Title (Length 26 characters):
Sources of Early Life SES

2. Writing Group:
Writing group members:
Kathryn Rose, Heejung Bang, Gerardo Heiss, Jay Kaufman, Cindy Owens, Mehul Patel

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

First author: Kathryn Rose

Address:
Bank of America Center
137 E Franklin St
Chapel Hill, NC 27514

Phone: 919-966-1967 Fax: 919-966-9800
E-mail: Kathryn_rose@unc.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: Kathryn Rose

Phone: 919-966-1967 Fax: 919-966-9800
E-mail: Kathryn_rose@unc.edu

3. Timeline:
4. Rationale:

Background

As part of ARIC ancillary study 2003.07 ((R01HL081627, K Rose, PI), we searched for parental SES (occupation) of ARIC decedents using birth certificates, declassified census records, and city directories. We recently completed our data collection and located parental SES data from at least one of these historical records for 3141 (91%) of 3444 decedent records searched. One of the specific aims of our study is to assess the comparability of this data (considered the gold standard) to that recalled by ARIC participants in mid to later adulthood. The recalled SES data were collected on ~12,700 ARIC survivors in 2001-2002 during annual telephone follow-up as part of an ARIC ancillary study [Life Course SES, Social Context, and Cardiovascular Disease (LC-SES), R01HL080287, G. Heiss, PI]. Approximately 1200 LC-SES participants have since died; thus, they have parental occupational data obtained from both historical records and self-report.

Rationale

In the life course epidemiologic literature, early life SES is often ascertained via self-report from adult participants. To date, the accuracy of data obtained using this approach has not been systematically investigated and there is some concern that recalled early life SES may underestimate true early life SES associations with chronic disease outcomes in adulthood. A review of the literature on SES and cause-specific mortality concluded that studies using objective measures of childhood SES reported stronger associations with mortality than those using recalled childhood SES (1). Also, in a Finish cohort, lower childhood socioeconomic position (SEP) was associated with an increase rate of coronary events when SEP was derived from historical school records but not when based on SEP information recalled in adulthood (2). Although the SEP measures differed across sources (the recalled SEP was based on reports of parents’ occupation and education and the historical SEP was derived from items recorded by a school nurse on hygiene and participation in school lunch programs), the results suggest that recall error may result in an underestimation of the impact of childhood SES. This illustrates importance of incorporating validation efforts into studies that rely on recalled childhood SES. We are aware of only one published validation study based on two small samples (N = 22 and N = 35) of British men, with reported concordance rates of 66% and 80%, respectively (3).

Given the limited literature assessing the validity of recalled early life SES, and the suggestion that such information may under-estimate true early life SES, we aim to assess the concordance and accuracy of recalled early life SES to that obtained from historical records. In addition to providing needed validation data on recalled life course SES measures, our work will estimate the magnitude and direction of recall error, and its potential impact on early life SES- chronic disease associations.

Main Hypothesis/Study Questions:

1. What is the level of agreement between early life (parental) occupations recorded on historical records (gold standard) to that recalled by participants in mid to later life.
   I. Does agreement vary by birth cohort, race, gender and source of historical data?
   II. Are recalled parental occupations more favorable (e.g., higher percentage of nonmanual or professional occupations) than those recorded on historical records?
2. What is the sensitivity and specificity of recalled early life SES (assuming that early life SES from historical records is the gold standard)?
   I. What is the estimated magnitude and direction of bias due to recall error in a simple analysis of early life SES- CVD associations?
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).

Our participants will include ~1200 ARIC decedents that participated in the LC-SES study and have since died. Their historical parental occupation data was collected over the past three years from birth certificates, declassified census records, and city directories. Occupations will be grouped into 6 standard census categories, as well as into categories commonly used in SES studies (e.g., manual and nonmanual occupational groups).

We will cross-tabulate the parental occupation data recalled by participants with those obtained from historical records to calculate % concordance. Kappa Statistics will be calculated [standard version as well as prevalence-adjusted, bias-adjusted version (PABAK)]. PABAK will be informative given the unbalanced distributions of parental occupations (favoring nonprofessional and manual occupations) (4,5,6,7). Accuracy will be assessed using sensitivity and specificity estimates (assuming that historical records are gold standard). We will also descriptively quantify the direction of differences between the recalled and historical record-derived parental occupations (e.g., difference in the proportion of fathers classified to nonmanual occupations). Such information will be used to inform simple sensitivity analyses (e.g., 8) to illustrate the potential impact (and range) of recall error on early life SES with CVD-related associations.

7.a. Will the data be used for non-CVD analysis in this manuscript?  ____ Yes  ____ No

This is a methodologic paper. However, papers based on this work will be used to address questions related to socioeconomic disparities in CVD and associated risk factors.

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?  ____ Yes  ____ No

(This file ICTDER03 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  ____ 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”?

 ____ 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 list 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)?
MS 1595. Another related MS was approved by ARIC but it is not currently listed on the ARIC website. It’s number is MS #970 and it is published:


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* 2003.07; 1998.02) also need to find aric # for LC-SES
   _ 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 at [http://www.cscce.unc.edu/aric/forms/](http://www.cscce.unc.edu/aric/forms/)

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.

OK

References


