**Number and Timing of Ambulatory Blood Pressure Monitoring Measurements**

Byron C. Jaeger, PhDa, Oluwasegun P. Akinyelure, MDb, Swati Sakhuja, MPHb, Joshua D. Bundy, PhD, MPHc, Cora E. Lewis, MD MSPHb, Yuichiro Yano, MD, PhDd, George Howard, DrPHa, Daichi Shimbo, MDe, Paul Muntner, PhDb, Joseph E. Schwartz, PhDe,f

1. Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL
2. Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL
3. Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA
4. Department of Medicine, Duke University, Durham, NC
5. Department of Medicine, Columbia University Irving Medical Center, New York City, NY
6. Department of Psychiatry, Stony Brook School of Medicine, Stony Brook, NY

Word count of the manuscript text: 2,844

Word count of the abstract: 250

Address for correspondence: Byron C. Jaeger, 327M Ryals Public Health Building, 1665 University Blvd Birmingham, AL. Email: [bcjaeger@uab.edu](mailto:bcjaeger@uab.edu)

Short title: Number and Timing of ABPM

**Abstract**

Ambulatory blood pressure (BP) monitoring (ABPM) may cause sleep disturbances. Some home BP monitoring (HBPM) devices are able to obtain a limited number of readings at specific times during sleep. Determining the optimal number and timing of BP measurements required to obtain an accurate estimate of sleep BP can inform the use of these devices. We used data from the Jackson Heart Study (JHS) and the Coronary Artery Risk Development in Young Adults (CARDIA) study to evaluate 74 different approaches to sample BP measurements during sleep. Approaches included 2 to 4 BP measurements obtained at specific clock times and times relative to the start of sleep. We assessed chance-corrected agreement (i.e., Kappa statistic) for classification of nocturnal hypertension (i.e., mean asleep systolic/diastolic BP ≥ 120/70 mm Hg) between each BP sampling approach and sampling BP throughout sleep. We computed a concordance (C-) statistic for left ventricular hypertrophy and albuminuria using a model that included asleep BP according to each BP sampling approach, separately. Sampling BP at 1, 2, 4, and 5 hours after falling asleep provided the highest overall Kappa statistic (Overall: 0.836; CARDIA: 0.839, JHS: 0.832). There was no evidence of a difference in C-statistics for left-ventricular hypertrophy or albuminuria when BP was sampled at 1, 2, 4, and 5 hours after falling asleep compared to sampling BP throughout sleep (p = 0.67). These results suggest that using four BP measurements at 1, 2, 4, and 5 hours after falling asleep provides high agreement with sampling BP throughout sleep.

**Introduction**

Higher blood pressure (BP) levels during sleep have been associated with an increased risk for cardiovascular disease (CVD) and target organ damage, independent of BP measured in a clinical setting.1–6 Ambulatory BP monitoring (ABPM) measures BP, typically every 15 to 30 minutes, throughout the day and night.7 Although most participants find ABPM acceptable, it may cause sleep disturbances for some individuals.8–10 Home BP monitoring (HBPM) is another approach for measuring BP outside of the office setting. HBPM devices have been developed that can be programmed to measure BP at pre-specific time periods including when someone is asleep. For instance, Kario et. al., reported on an HBPM device that is put on before going to sleep and measures BP at 2:00, 3:00, and 4:00 AM.11 Also, some BP monitoring devices allow users to specify daytime and nighttime periods (e.g., daytime is from 6am to 10pm and nighttime is from 10pm to 6am) as well as time intervals between BP measurements (i.e., 15, 20, 30, or 60-minute intervals).

Obtaining fewer BP readings during sleep using a HBPM versus ABPM device may reduce discomfort and disrupted sleep.12 However, few studies have considered the number and timing of BP measurements required to obtain an estimate of BP during sleep similar to that obtained by a full ABPM recording (i.e., using ABPM throughout sleep). Yang et. al., and Rinfret et. al., independently examined the question of how many readings should be collected in order to obtain a reasonably accurate estimate of mean daytime and nighttime BP or mean BP using HBPM.13,14 However, these analyses examined scenarios where BP measurements were obtained randomly.

We evaluated a total of 74 different variations on sampling BP during sleep, each using a specific number and timing of measurements. We assessed each BP sampling variations’ chance-corrected agreement (i.e., Kappa statistic) with sampling BP throughout sleep for classification of nocturnal hypertension (i.e., mean asleep systolic/diastolic BP ≥ 120/70 mm Hg).15 We also computed a concordance (C-) statistic for prediction of left ventricular hypertrophy and albuminuria using a model that included sleep BP according to each BP sampling approach, and tested whether this C-statistic was different from that of a model using sleep BP according to full ABPM. Last, we assessed the association (i.e., prevalence ratio) of systolic BP (SBP) and diastolic BP (DBP) during sleep according to each BP sampling variation with left-ventricular hypertrophy and albuminuria.

**Methods**

*Study population*

We used data from participants in the Jackson Heart Study (JHS) and the Coronary Artery Risk Development in Young Adults (CARDIA) study who underwent 24‐hour ABPM.

The JHS, a community-based prospective cohort study, was designed to evaluate the etiology of CVD among African Americans living in or near Jackson, MS.16 The JHS enrolled a total of 5,306 non-institutionalized African Americans aged ≥21 years between 2000 and 2004. At the baseline JHS visit, 1,146 participants elected to undergo ABPM. The CARDIA study was designed to examine the development and determinants of clinical and subclinical CVD and their risk factors.17 The CARDIA study recruited 5,115 participants, 18 to 30 years of age, at four field centers in the United States (Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA) in 1985-1986. During the Year 30 Exam (2015–2016), 831 CARDIA participants enrolled in an ABPM ancillary study conducted in the Birmingham, AL and Chicago, IL field centers.

We included participants from each study who slept ≥5 hours while wearing the ABPM device. We further restricted the analysis to participants with at least one valid BP reading during sleep within 30 minutes of all times required by the BP sampling variations we studied (N = 621 JHS and 458 CARDIA participants; **Table** **S1**). Conduct of the JHS and CARDIA study was approved by institutional review boards at the participating institutions and the current analysis of JHS and CARDIA data was approved by the University of Alabama at Birmingham Institutional Review Board. Written informed consent was obtained from all participants.

*Left ventricular hypertrophy and albuminuria*

Echocardiograms and urine specimens were assessed during the Year 30 Exam for CARDIA participants and during the baseline study visit for JHS participants. Left ventricular mass was determined and indexed to body surface area to obtain left ventricular mass index (LMVI) according to recommendations from the American Society of Echocardiography and European Association of Cardiovascular Imaging.18 Left ventricular hypertrophy (LVH) was defined as LVMI >95 g/m2 in women and >115 g/m2 in men. Urine specimens were used to measure urinary albumin and creatinine excretion, which were used to calculate the urine albumin-to-creatinine ratio (ACR). Albuminuria was defined as an ACR ≥30 mg/g. Albuminuria was quantified using a 24-hour urine sample in the JHS, if available. Otherwise, a spot urine sample was used. In CARDIA, a spot urine sample was collected for all willing participants.

*Ambulatory blood pressure monitoring*

In the JHS, ABPM was conducted using the SpaceLabs model 90207 device (SpaceLabs Healthcare, Snoqualmie, WA), which has been previously validated, and BP was measured every 20 minutes over a 24-hour period.19 JHS participants self-reported the times they went to sleep and woke up while wearing the ABPM device. In CARDIA, ABPM was conducted using the SpaceLabs OnTrak model 90227 device (SpaceLabs Healthcare, Snoqualmie, WA), which has also been previously validated, with an appropriately sized cuff andBP was measured every 30 minutes over a 24-hour period.20 CARDIA participants also wore an Actiwatch activity monitor (Philips Respironics, Murrysville, PA) on the wrist of their non-dominant arm. For CARDIA participants, awake and asleep time periods were determined using the activity monitor data in conjunction with participants’ self-reported awake and asleep times. Nocturnal hypertension was defined as a mean SBP ≥ 120 mm Hg or mean DBP ≥ 70 mm Hg during sleep.

*Blood pressure sampling strategies*

We considered ‘distributed’ and ‘consecutive’ strategies to sample BP during sleep. The distributed strategy used BP with intervals between measurements spanning at least 1 hour, whereas the consecutive strategy used consecutive measurements of BP (**Figure 1**). We considered 25 and 12 variations of the distributed and consecutive strategies, respectively, and implemented each variation according to 2 separate time structures, time since midnight and time since falling asleep. In total, we evaluated (25 + 12) \* 2 = 74 different variations on sampling BP.

*Statistical analyses*

Participant characteristics were tabulated overall and by study cohort. Differences in mean SBP and DBP during sleep, LVH, and albuminuria were compared between study cohorts using t- and chi-square tests for continuous and categorical variables, respectively. Analyses were conducted using R version 4.0.0 or later.21–24

*Evaluation of blood pressure sampling variations*

We computed each BP sampling variations’ chance-corrected agreement (i.e., Kappa statistic) with sampling BP throughout sleep for classification of nocturnal hypertension. We also computed the mean absolute difference between mean SBP and DBP during sleep according to each BP sampling variation and full ABPM throughout sleep. Using linear regression, we tested whether the difference between asleep SBP and DBP between full ABPM and each BP sampling variation was associated with the magnitude of asleep SBP and DBP, separately.

*Group-wise ranking of blood pressure sampling variations*

We defined 12 groups of BP sampling variations (**Table S2**). Each group comprised BP sampling variations taking the same number of measurements (i.e., 2, 3, or 4) and using the same strategy (i.e., consecutive or distributed) and time structure (i.e., time since midnight or time since falling asleep). We ranked BP sampling variations within each group based on the overall Kappa statistic. Our primary analyses assessed the 12 BP sampling variations (1 from each group) exhibiting the highest overall Kappa statistics. Secondary analyses assessed all 74 BP sampling variants.

*Comparisons of blood pressure sampling variants*

We applied bootstrap resampling with bias correction and acceleration to estimate differences in Kappa statistics between BP sampling variants among JHS and CARDIA participants, separately.25 These comparisons were limited to the 12 BP sampling variants described above. Comparisons were also stratified by time structure (i.e., time since midnight or time since falling asleep).

*Consistency of results across studies*

To assess the consistency of our findings, we calculated the Spearman rank order correlation coefficient for rankings of BP sampling variations in the JHS and CARDIA study. A Spearman correlation of 1 indicates that the BP sampling variations were ranked in the exact same order in both studies.

*Asleep BP, left-ventricular hypertrophy, and albuminuria*

Poisson regression models with robust standard errors were applied to estimate associations of asleep BP with LVH and albuminuria.26 Models were fitted (1) using SBP and DBP according to full ABPM, (2) foregoing SBP and DBP, and (3) using SBP and DBP from each BP sampling variation, separately. Prevalence ratios and C-statistics were computed for each model overall and among JHS and CARDIA participants, separately. DeLong’s test was applied to assess whether any BP sampling variation changed the model’s C-statistic compared to measuring BP throughout sleep. All models included adjustment for age, sex, race (for CARDIA participants only; all JHS participants are black), smoking status, diabetes, antihypertensive medication use, and sleep duration.

*Missing data*

The count and percent of missing values for each study variable were examined. As there were no missing values for primary study variables (asleep SBP and DBP), we did not impute missing data.

**Results**

Among participants included in the current analysis, the mean (standard deviation; SD) age was 57.1 (8.6) years. Additionally 32.0% of participants were male and 81.0% were black. Among JHS and CARDIA participants, the mean (SD) asleep SBP was 120 (14.7) mm Hg and 111 (15.1) mm Hg, respectively (**Table 1**; p <0.001). There was no evidence of a difference in the prevalence of LVH and albuminuria between JHS and CARDIA participants (p=0.29 and 0.30, respectively).

*Evaluation of blood pressure sampling variations*

Kappa statistics for BP sampling variations ranged from 0.68 to 0.85 in CARDIA and 0.66 to 0.83 in the JHS (**Table 2** [top 12 variations] and **Table S3** [all 74 variations]). Among all 74 variations, mean absolute error for SBP (DBP) ranged from 2.95 (2.41) to 5.94 (4.79) mm Hg in CARDIA and 3.12 (2.53) to 6.45 (5.30) in the JHS. The highest Kappa statistic overall and for JHS participants was obtained from sampling BP at 1, 2, 4, and 5 hours after sleep (CARDIA: 0.84, JHS: 0.83). For CARDIA participants, the highest Kappa statistic was obtained from sampling BP at 1, 2, 4 and 5 hours after midnight (CARDIA: 0.85, JHS: 0.78).

*Comparisons of Kappa statistics among blood pressure sampling variants*

Among all pairwise comparisons conducted in the current analysis, there was no evidence that a consecutive BP sampling variation obtained a higher Kappa statistic than a distributed BP sampling variation using the same number of measurements. Sampling BP at 1, 2, 4 and 5 hours after falling asleep increased the Kappa statistic by at least 0.0362 (95% CI -0.0028 – 0.0779) and 0.0045 (95% CI -0.0398 – 0.0494) among JHS and CARDIA participants, respectively, compared to other BP sampling variations that measured time relative to falling asleep (**Figure 2**). Sampling BP at 1, 2, 4 and 5 hours after midnight yielded a higher Kappa statistic by at least 0.0099 (95% CI -0.0289 – 0.0491) and 0.0562 (95% CI 0.0139 – 0.102) among JHS and CARDIA participants, respectively, compared to other BP sampling variations that measured time relative to midnight (**Figure S1).**

*Consistency of results across studies*

The overall correlations between the JHS and CARDIA study rankings of BP sampling variants according to the mean absolute difference in SBP, mean absolute difference in DBP, and Kappa statistics were 0.92, 0.93, and 0.78, respectively.

*Asleep BP, left-ventricular hypertrophy, and albuminuria*

The overall prevalence ratio for LVH corresponding with 10 mm Hg higher asleep SBP was 1.22 (95% CI 1.02 – 1.46) when BP was measured throughout sleep versus 1.24 (95% CI 1.04 – 1.48) when BP was measured at 1, 2, 4, and 5 hours after falling asleep (**Table 3**). Prevalence ratios for albuminuria using the same BP sampling variations were 1.27 (95% CI 1.07 – 1.52) versus 1.35 (95% CI 1.15 – 1.60) (**Table S4**). For multi-variable Poisson regression models, when BP was measured throughout sleep versus at 1, 2, 4, and 5 hours after falling asleep, the model’s C-statistic was 0.712 (95% CI 0.659 - 0.765) versus 0.705 (0.651, 0.760) for LVH (p-value for difference: 0.31; **Table 4**) and 0.774 (95% CI 0.719 - 0.829) versus 0.776 (0.720 - 0.832) for albuminuria (p-value for difference: 0.72; **Table S5**).

Consistency of findings

The overall correlations between the JHS and CARDIA study rankings of BP sampling variants according to the mean absolute difference in SBP, mean absolute difference in DBP, and Kappa statistics were 0.92, 0.94, and 0.78, respectively.

**Discussion**

In two independent cohorts, we investigated 74 BP sampling variations based on the number and timing of BP measurements. The largest overall Kappa statistic resulted from sampling BP at 1, 2, 4, and 5 hours after falling asleep. This BP sampling variation also provided a relatively low mean absolute error for SBP/DBP during sleep (SBP: 3.11 mm Hg, DBP: 2.65 mm Hg). Bootstrapped comparisons of Kappa statistics identified 16 instances where a distributed BP sampling obtained a higher Kappa statistic than a consecutive BP sampling variation and found no evidence of a consecutive BP sampling variation obtaining a higher Kappa statistic than a distributed BP sampling variation. The prevalence ratios for LVH and albuminuria based on measurements at 1, 2, 4, and 5 hours after falling asleep were within 1 standard error of the prevalence ratios based on measuring BP throughout sleep, and there was no evidence of a difference in model discrimination (i.e., C-statistic) based on these two BP sampling variations. The high correlation of Kappa statistic and mean absolute error rankings for BP sampling variations in CARDIA and the JHS indicated that results were consistent among participants in the two cohorts, suggesting that findings from the current study are not overly influenced by results from a single cohort.

In a previous study, Yang et al concluded that randomly measuring BP four times during sleep versus measuring BP throughout sleep does not lead to a meaningful loss of information in hypertension categorization or risk stratification.13 In the current study, we found substantial variability in the information retained by BP sampling variations that used four measurements. Specifically, among CARDIA participants, the Kappa statistic ranged from 0.745 (4 consecutive BP measurements starting at 3am) to 0.854 (4 distributed BP measurements at 1, 2, 4, and 5 am). Among JHS participants, the Kappa statistic ranged from 0.674 (four consecutive measurements starting at 4am) to 0.832 (four distributed measurements at 1, 2, 4, and 5 hours after the onset of sleep). Our results are consistent with and extend findings from Yang et al by indicating that four BP measurements are sufficient for measuring BP during sleep and that the timing of BP measurements substantially impacts the accuracy of mean BP during sleep. Given that the median (interquartile range) number of successful BP readings during sleep for JHS and CARDIA participants in the current study was 24 (22 - 27) and 16 (14 – 18), respectively, BP monitoring may cause substantially less sleep disturbance if only four BP measurements are taken during sleep.

Among variants that used three or four BP measurements, several appear to be accurate. In total, five and nine BP sampling variations using three and four BP measurements during sleep, respectively, obtained an estimated Kappa statistic > 0.80, suggesting strong agreement with measuring BP throughout sleep. For example, sampling BP at 2, 3, and 4 hours after sleep (the sampling variant studied by Kario et al.) obtained an overall Kappa statistic of 0.81 (95% CI 0.78 – 0.85). However, the current study did not find any BP sampling variations using two measurements that obtained Kappa statistics exceeding 0.80. These results suggest that some flexibility is warranted for choosing the timing of three or four BP measurements during sleep, as many of these variations obtained excellent agreement with measuring BP throughout sleep.

The current study has several strengths. We analyzed data from two independent cohorts that collected ABPM data. We investigated a comprehensive set of variants for sampling BP during sleep, allowing us to identify several variants that exhibited high agreement with full ABPM. We conducted analyses separately by study, and the parallel assessment of each BP sampling variant reduced the likelihood of finding spurious results that would not generalize to broader settings. In addition, the current study is subject to some limitations. While sleep was monitored using actigraphy in the CARDIA study, the JHS relied on self-reported sleep diaries to identify awake and asleep times. Due to strict inclusion criteria, the current study excluded a substantial proportion of participants from each cohort. Results from the current study may not generalize to settings where participants sleep for <5 hours or miss any planned BP measurements.

In summary, measuring BP 3 or 4 times during sleep may provide mean asleep BP estimates that have high agreement with measuring BP throughout sleep. Additionally, measuring BP at 1, 2, 4, and 5 hours after sleep or 1, 2, 4, and 5 hours after midnight may obtain high agreement with measuring BP throughout sleep.

**References**

1. O’Brien E, Parati G, Stergiou G, Asmar R, Beilin L, Bilo G, Clement D, De La Sierra A, De Leeuw P, Dolan E. European Society of Hypertension position paper on ambulatory blood pressure monitoring. *Journal of hypertension*. 2013;31:1731–1768.

2. Parati G, Stergiou G, O’Brien E, Asmar R, Beilin L, Bilo G, Clement D, De La Sierra A, De Leeuw P, Dolan E. European Society of Hypertension practice guidelines for ambulatory blood pressure monitoring. *Journal of hypertension*. 2014;32:1359–1366.

3. Shimamoto K, Ando K, Fujita T, Hasebe N, Higaki J, Horiuchi M, Imai Y, Imaizumi T, Ishimitsu T, Ito M. The Japanese Society of Hypertension guidelines for the management of hypertension (JSH 2014). *Hypertension Research*. 2014;37:253–390.

4. Friedman O, Logan AG. Can nocturnal hypertension predict cardiovascular risk? *Integrated blood pressure control*. 2009;2:25.

5. Yano Y, Tanner RM, Sakhuja S, Jaeger BC, Booth JN, Abdalla M, Pugliese D, Seals SR, Ogedegbe G, Jones DW. Association of daytime and nighttime blood pressure with cardiovascular disease events among African American individuals. *JAMA cardiology*. 2019;4:910–917.

6. Kario K. Nocturnal hypertension: new technology and evidence. *Hypertension*. 2018;71:997–1009.

7. Pickering TG, Shimbo D, Haas D. Ambulatory blood-pressure monitoring. *New England Journal of Medicine*. 2006;354:2368–2374.

8. Ernst ME, Bergus GR. Favorable patient acceptance of ambulatory blood pressure monitoring in a primary care setting in the United States: a cross-sectional survey. *BMC family practice*. 2003;4:15.

9. Degaute JP, Kerkhofs M, Dramaix M, Linkowski P. Does non-invasive ambulatory blood pressure monitoring disturb sleep? *Journal of hypertension*. 1992;10:879–885.

10. Agarwal R, Light RP. The effect of measuring ambulatory blood pressure on nighttime sleep and daytime activity—implications for dipping. *Clinical Journal of the American Society of Nephrology*. 2010;5:281–285.

11. Yarows SA, Julius S, Pickering TG. Home blood pressure monitoring. *Archives of Internal Medicine*. 2000;160:1251–1257.

12. Kario K, Saito I, Kushiro T, Teramukai S, Ishikawa Y, Mori Y, Kobayashi F, Shimada K. Home blood pressure and cardiovascular outcomes in patients during antihypertensive therapy: primary results of HONEST, a large-scale prospective, real-world observational study. *Hypertension*. 2014;64:989–996.

13. Yang W-Y, Thijs L, Zhang Z-Y, Asayama K, Boggia J, Hansen TW, Ohkubo T, Jeppesen J, Stolarz-Skrzypek K, Malyutina S. Evidence-based proposal for the number of ambulatory readings required for assessing blood pressure level in research settings: an analysis of the IDACO database. *Blood pressure*. 2018;27:341–350.

14. Rinfret F, Ouattara F, Cloutier L, Larochelle P, Ilinca M, Lamarre-Cliche M. The impact of unrecorded readings on the precision and diagnostic performance of home blood pressure monitoring: a statistical study. *Journal of human hypertension*. 2018;32:197–202.

15. Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology*. 2010;73:1167–1179.

16. Taylor Jr HA, Wilson JG, Jones DW, Sarpong DF, Srinivasan A, Garrison RJ, Nelson C, Wyatt SB. Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis*. 2005;15:S6–4.

17. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, Liu K, Savage PJ. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *Journal of clinical epidemiology*. 1988;41:1105–1116.

18. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *European Heart Journal-Cardiovascular Imaging*. 2015;16:233–271.

19. O’Brien E, Mee F, Atkins N, O’Malley K. Accuracy of the SpaceLabs 90207 determined by the British Hypertension Society protocol. *J Hypertens*. 1991;9:S25–S31.

20. A. de Greef, Shannan AH. Validation of Spacelabs 90227 OnTrak upper arm blood pressure monitor, for clinical use, according to the European Society of Hypertension International Protocol 2010 and the British Hypertension Society Protocol. Available at http://www.dableducational.org/Publications/2014/ESH-IP%202010%20and%20BHS%20Validation%20of%20Spascelabs%2090227%20OnTrak.pdf.

21. Jaeger B. tibbleOne: Table One for “Latex”, “Word”, and “Html” “R Markdown” Documents. .

22. Buuren S van, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *Journal of statistical software*. 2010;:1–68.

23. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J. Welcome to the Tidyverse. *Journal of Open Source Software*. 2019;4:1686.

24. Halekoh U, Højsgaard S, Yan J. The R package geepack for generalized estimating equations. *Journal of Statistical Software*. 2006;15:1–11.

25. Efron B. Better bootstrap confidence intervals. *Journal of the American statistical Association*. 1987;82:171–185.

26. Zou G. A modified poisson regression approach to prospective studies with binary data. *American journal of epidemiology*. 2004;159:702–706.