vaccine is the origin of human immunodeficiency virus in Africa. In this regard, Dr. Alcabes (2) agrees with me (3) that the theory is not proven. He also agrees that the book is overlong and difficult to follow. Our basic difference concerns the obligation of the reader. Is it reasonable that various subtexts scattered through the volume will be picked up? Further, when it becomes clear that the author is not an independent investigator, but rather someone who is out to prove a point, how much credibility do these points have? Personally, I think not much, but since Dr. Alcabes has brought up a number of points, I will give my views on some of his comments. First, I do not believe that the scientific community should be immune from scrutiny by the press. However, I feel that we do not have to applaud the kind of personality dissection that goes on in this book, especially when people carrying it out are pursuing their own agendas, which happen to be wrong. In addition, I think it is incorrect to apply the standards of the 21st century to work done in the middle of the 20th. This does not mean that what was done in Africa during that period was right. Neither was research done in vulnerable populations in the developed world. However, this was not the main topic of the book. Such value judgments are unrelated to my conclusion that it might not be worth the time of busy readers to work their way through this diffuse text.

References

1. Hooper E. The river: a journey to the source of HIV and AIDS. New York, NY: Little, Brown & Co, 2000.
2. Alcabes P. Re: "The river: a journey to the source of HIV and AIDS." (Letter). Am J Epidemiol 2002;155:1143.
3. Monto AS. The river: a journey to the source of HIV and AIDS. (Review). Am J Epidemiol 2001;154:484–5.

Arnold S. Monto
Department of Epidemiology
School of Public Health
University of Michigan
Ann Arbor, MI 48109–2029

RE: "TRACKING OF CARDIOVASCULAR RISK FACTORS: THE TROMSØ STUDY, 1979–1995"

Wilsgaard et al.'s recent paper (1) addressed an important issue, the tracking of cardiovascular risk factors. Their study also reminds us about the current methodological difficulties in studying "tracking" and the related controversies (2–4). The authors followed 17,710 men and women for 16 years and found high or moderate tracking for a number of risk factors (1). They defined tracking as "either the stability of a certain variable over time (e.g., maintenance of a relative position . . .) or the predictability of later values from earlier measurements . . . " (1, p. 418). Several issues need to be addressed.

First, Wilsgaard et al. (1) used "standardized regression coefficients" (called "generalized estimating equation (GEE) tracking coefficients") to compare the tracking of different outcomes. It is not clear how they calculated these coefficients. Neither they nor Twisk et al. (5), cited in their paper as reference 19, provided clarification. This calculation can be confusing. To our knowledge, some researchers previously used this term to mean regression coefficients adjusted for covariates. We suspect that before fitting GEE models, Wilsgaard et al. "standardized" their variables (e.g., $Y_{\text{std}} = (Y_{it} - \mu)/\sigma$), although they claimed "Y_{it} is the dependent variable (which . . . may be blood pressure, BMI [body mass index] . . .) for subject i at time t_2 or t_3 . . . " (1, p. 420). Note that μ (mean) and σ (variance) should be examination specific, because they might change when subjects become older. In addition, these authors should clarify how statistical tests were performed to test the differences, because a separate model was fit for each outcome by sex and age (1, p. 422 (table 3)).

Second, they presented the "proportion of subjects who remained in the same sextile throughout the different examinations" (1, p. 420). It is not clear whether these were age- and sex-specific sextiles. This point is important because most outcomes examined are likely to increase with age. Subjects' age at baseline was 20–55 years. If age-specific sextile was not used, older subjects would have been more likely to be in the upper sextile. This difference would cause misclassification and influence the authors' findings.

Third, Wilsgaard et al. (1) discussed sources of bias. However, an important issue, the influence of "survivor effects," was not addressed. The tracking patterns observed might have been stronger if such influence could have been corrected; those subjects who tracked the risk factors were more likely than their counterparts to have suffered from diseases and to have died during follow-up.

Finally, the authors made an interesting point when they examined the linkage between odds ratios and percentage of tracking. They suggested using 50 percent to classify moderate tracking (1, p. 424). This cutpoint is obviously arbitrary, and it seems too conservative. Because the expected tracking rate by chance (i.e., no tracking) is 16.7 percent (i.e., 6/36 for sextile and two examinations), 50 percent is three times greater. Considering the health consequences of tracking and possible approaches to preventing tracking, we think that a lower cutpoint would be more appropriate (e.g., 1.5–2 times the expected rate).

In addition, although odds ratio and percentage of tracking are closely related, they are not the same. Tracking percentage indicates how many of those subjects in the upper sextile remain in that category in later examinations; odds ratio tells how much more likely it is for subjects to remain in the upper sextile than to move into the upper sextile from other categories. To calculate the odds ratio, Wilsgaard et al. (1) fit logistic models by using GEE. Since data on subjects for whom values were missing were included in GEE whereas sextile was examination specific, the linkage between odds ratio and percentage of tracking is not as straightforward as it appears, especially when there are considerable missing values. Thus, one should be cautious when interpreting the related results.

In short, Wilsgaard et al.'s findings (1) are interesting and useful. Further studies on tracking, particularly those that address the related methodological issues, are needed to fill gaps in the literature.

References

1. Wilsgaard T, Jacobsen BK, Schirmer H, et al. Tracking of cardiovascular risk factors: The Tromsø Study, 1979–1995. Am J Epidemiol 2001;154:418–26.
2. Twisk JW, Kemper HC, Mellenbergh GJ. Mathematical and analytical aspects of tracking. Epidemiol Rev 1994;16:165–83.
3. Lindsay R, Hanson RL, Knowler W. Tracking of body mass index during childhood and adolescence. (Letter). Am J Clin Nutr 2001;74:149.
4. Wang Y, Popkin B, Wang X. Tracking of body mass index during childhood and adolescence. (Letter). Am J Clin Nutr 2001; 74:149–50.
5. Twisk JW, Kemper HC, Mellenbergh DJ, et al. Factors influencing tracking of cholesterol and high-density lipoprotein: the Amsterdam Growth and Health Study. Prev Med 1996;25: 355–64.

Youfa Wang
Department of Human Nutrition
Department of Epidemiology and Biostatistics
University of Illinois
Chicago, IL 60612

Xiaofei Wang
Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27599

THE AUTHORS REPLY

We appreciate the interest that Drs. Wang and Wang (1) demonstrate in our paper (2). Their comments highlight several methodological difficulties in studying tracking.

It was not clear to Wang and Wang (1) how we calculated our standardized regression coefficients (2). We apologize if this method was not stated clearly enough. We assumed a common definition of a standardized regression coefficient (3). Hence, if the estimated $\hat{\beta}$ is the regression coefficient for a given age and sex strata, and $\hat{\sigma}_x$ and $\hat{\sigma}_y$ are the sex- and age-specific sample standard deviations for the explanatory and response variables, respectively, our standardized regression coefficient was calculated as $\beta^* = \hat{\beta} \times \hat{\sigma}_x/\hat{\sigma}_y$. Equivalently, we could have standardized our variables before fitting the GEE models (e.g., $y_{\text{std}} = (y_{it} - \mu)/\sigma$). Wang and Wang think that we should have used examination-specific μ (mean) and σ (variance). We assumed that doing so would not influence the results in our adult population, but, in a separate set of analyses, we did standardize our variables by using examination-specific means and variances. As expected, the results remained virtually unchanged, and all conclusions were identical. The statistical tests of sex difference between the separate models were performed by using the test statistic $Z = (\beta_1^* - \beta_2^*)/\sqrt{(\text{SE}[\beta_1^*])^2 + (\text{SE}[\beta_2^*])^2}$, which approximately follows a standard normal distribution, where $\hat{\beta}_1^*$ and $\hat{\beta}_1^*$ are standardized regression coefficients for men and women, respectively.

We used sex-specific, but not age-specific, sextiles when presenting estimated probabilities of changing sextile group throughout the different examinations (2). It is correct, as stated by Wang and Wang (1), that older persons would be more likely to be in the upper sextile at baseline. However, we do not think that this caused significant misclassification. Our calculations were based on the fact that each subject used his or her initial rank (sextile) as the reference point; that is, given each subject's initial sextile group, we calculated the probability of his or her remaining at the same initial level. It is, of course, correct that tracking could be age dependent. The results in our table 4 (2) concerning the probability of changing sextile group could consequently differ between different age groups. However, to include age-specific results in the table would have made it very large indeed. In a separate set of analyses, we calculated age-specific results. The pattern of differences in tracking between the different age groups was in agreement with the age-specific results presented in table 3 (2); for the youngest persons, the degree of tracking was the lowest. However, we do agree that this information could have been included in our paper.

We did not address a possible survivor effect because our participants comprised a relatively young general population. The age of the oldest persons was 69 years at the last examination. Relatively few persons died during the follow-up period. For the oldest men, mortality during the 16 years of follow-up was not negligible, however. This finding if anything reduced the strength of the tracking in the older age groups.

Finally, classification of cutpoints of tracking in a high-risk group is not straightforward. We did not recommend that the cutpoint for the level of the predictive value be 50 percent, and we will not dispute Wang and Wang's suggestion (1) of a lower cutpoint. We used the value of 50 percent in our discussion because tracking of most of our considered risk factors was in the 40–60 percent range.

To our knowledge, the literature contains no common method of tracking assessment. We tried to assess a method that has several advantages (2). We would certainly like to see more studies on tracking by using different methods, making comparisons possible.

References

1. Wang Y, Wang X. Re: "Tracking of cardiovascular risk factors: The Tromsø Study, 1979–1995." (Letter). Am J Epidemiol 2002;155:1144–5.
2. Wilsgaard T, Jacobsen BK, Schirmer H, et al. Tracking of cardiovascular risk factors: The Tromsø Study, 1979–1995. Am J Epidemiol 2001;154:418–26.
3. Kleinbaum DG, Kupper LL, Muller KE, et al. Applied regression analysis and other multivariable methods. Pacific Grove, CA: Duxbury Press, 1998.

Tom Wilsgaard
Bjarne K. Jacobsen
Henrik Schirmer
Inger Thune
Maja-Lisa Løchen
Inger Njølstad
Egil Arnesen
Department of Epidemiology and Medical Statistics
Institute of Community Medicine
University of Tromsø
N-9037 Tromsø, Norway