institutions and request ever-increasing sums of money to compete with one another as the family practice bidding war heats up.

Donald R. Frey, M.D.
United Hospital Center
Clarksburg, WV

References

1. Laufenburg HF. Recruitment to family practice residencies. J Am Board Fam Pract 1992; 5:109.
2. Page L. Oklahoma luring FP residents with 65% raise. American Medical News 1992 January 20.

Physical Activity and Coronary Heart Disease

To the Editor: I was pleased to see the meta-analysis conducted by Dr. Eaton on the relation of physical activity to coronary heart disease.[1] It is exciting to see meta-analyses appearing in the family medicine literature; however, I wish to raise four issues that, I believe, are relevant to the study.

First, with the exception of comparing cohort and case-control studies, there is no assessment of study quality and its relation to effect size. Sacks, et al.[2] emphasize the importance of such analysis.

Second, the use of historical cohorts in two of the studies is problematic. With the probable exception of Glass, most meta-analysts would recommend excluding studies that did not use equivalent controls.[3]

Third, of more than 75 articles reviewed, only 14 studies were used. Optimally, we should be given a list of rejected studies with the reason for rejection. Minimally, a list of rejected studies should be available upon request.[2]

Perhaps my greatest concern is the lack of recognition of possible publication bias. That the unpublished literature frequently differs from the published literature in its findings is well documented.[4] Although some meta-analysts have rationalized ignoring the unpublished literature,[5] the appropriateness of this approach has not been proved.[6] When conducting a meta-analysis, every attempt should be made to locate unpublished studies. If this search is not done or is unfruitful, other techniques should be employed. In Dr. Eaton's study, he reports the overall weighted relative risk (RR) = 1.37 using 12 studies. Using the natural logarithmic transformation and Rosenthal's formula[7] for fail-safe N — the number of unpublished studies with null results needed to overturn an overall significant effect size — I calculated that less than one study would be needed. Using Orwin's formula,[8] I calculated that 26 studies having RR ≤ 1.1 would be needed. Although Orwin's formula produces somewhat reassuring results, if we remember that an estimated 2 to 10 unpublished studies are needed for every published study, then needing 25 unpublished null studies to overturn the results of 12 published studies is of concern. Using a funnel graph plot[9] — effect size versus sample size — also raises concern because the only published study with

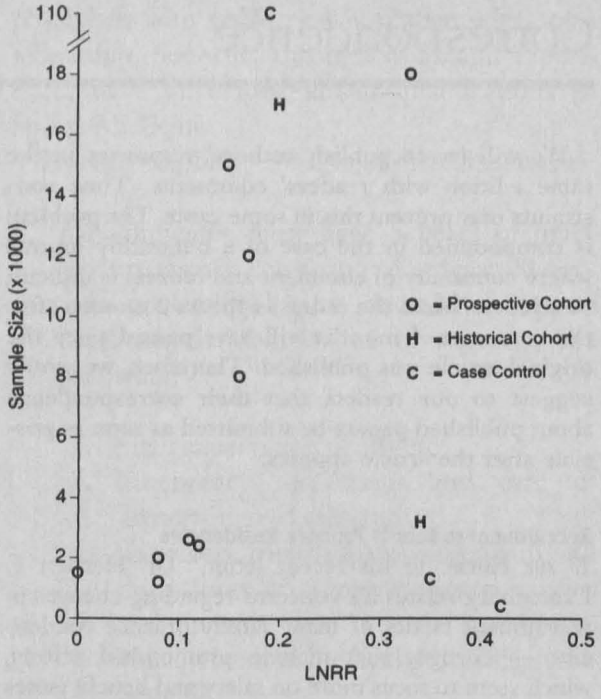

Figure 1. Funnel graph.

sample size < 1000 had the greatest relative risk. In fact, as a rule, the cohort studies had the lowest RRs (Figure 1).

Although I enjoyed Dr. Eaton's article and believe that the conclusions are probably valid, I think the concerns that I have raised do have important implications.

David A. Katerndahl, M.D.
University of Texas Health Science Center
at San Antonio

References

1. Eaton CB. Relation of physical activity and cardiovascular fitness to coronary heart disease, Part I: a meta-analysis of the independent relation of physical activity and coronary heart disease. J Am Board Fam Pract 1992; 5:31-42.
2. Sacks HS, Berrier J, Reitman D, Ancora-Berk VA, Chalmers TC. Meta-analysis of randomized controlled trials. N Engl J Med 1987; 316:450-5.
3. Bangert-Drowns RL. Review of developments in meta-analytic method. Psychol Bull 1986; 99:388-99.
4. Katerndahl DA. Believing everything you read. Fam Pract Res J 1991; 11:343-7.
5. Steinberg KK, Thacker SB, Smith SJ, Stroup DF, Zach MM, Flanders WD, et al. Meta-analysis of the effect of estrogen replacement therapy on the risk of breast cancer. JAMA 1991; 265:1985-90.
6. Smith ML. Publication bias and meta-analysis. Eval Educ 1980; 4:22-4.
7. Rosenthal R. The file drawer problem and tolerance for null results. Psychol Bull 1979; 86:638-41.
8. Orwin RG. A fail-safe N for effect size in meta-analysis. J Educ Stat 1983; 8:157-9.

9. Light R, Pillemer D. Summing up. Cambridge, MA: Harvard University Press, 1984.

The above letter was referred to the author in question, who offers the following reply.

To the Editor: Dr. Katerndahl's letter discusses some important issues related to the utilization of meta-analysis as a useful analytic tool in medicine. In my article the meta-analysis used for estimating a summary "measure of effect" was not that of physical inactivity and coronary heart disease (CHD), which had been performed in 1987 by Powell, et al.[1] and more formally in 1990 by Berlin and Colditz,[2] but rather was the study of the independent or multivariate relation of physical inactivity and CHD. Both of the former meta-analyses performed quality estimates using a scoring system of 0 to 6 in Powell, et al. and 0 to 32 in Berlin and Colditz based upon the measure of activity, the measure of outcome, and the epidemiologic methods. The 12 studies included in my meta-analysis received scores from 4 to 6 on Powell, et al.'s scale and 18 to 26 using Berlin and Colditz's scale. These were not included in the description of the studies for the sake of brevity, as the reasons for inclusion were outlined in the Methods section in paragraphs 2 and 3. Only studies that used easily standardized hard endpoints and adjusted for confounders of age, smoking, lipids, and blood pressure were utilized in my analysis. Additional exclusions were studies that assessed cardiovascular fitness (attribute) and not physical inactivity (behavior). It is for this reason that only 12 out of the 75 studies were included, and all had quality scores near the highest categories established by Powell, et al. and Berlin and Colditz.

The term *historical cohort* used to define the studies by Dr. Ralph Paffenbarger, including the San Francisco longshoremen study[3] and Harvard alumni study,[4] refers to the fact that these two studies used job classification and mail surveys, respectively, to define physical inactivity in a previously established cohort. They certainly included a control group of unexposed individuals and used the same criteria for evaluating dependent and independent variables in both groups. Perhaps Dr. Katerndahl is confusing case series with historical cohort studies. These two studies[3,4] are considered critical studies in our understanding the relation of physical inactivity and CHD, because the San Francisco longshoremen showed an apparent threshold effect, and the Harvard alumni study showed an effect not only on CHD mortality, but also on total mortality. It seems unwise to have had left out these important studies from my analysis.

Of more than 75 articles reviewed, 14 studies met the standards of evidence described above and evaluated the *independent relation* of physical inactivity and CHD. Two studies (Health Insurance Plan[5] and British civil servants[6]) listed in Table 1 used stratification to analyze confounding and thus could not be included in the weighted summary measure used in my meta-analysis and were therefore not included in Figure 1. Lists of the other 61 references are available in the bibliographies of Powell, et al. and Berlin and Colditz for 53 references and in Part II of this review regarding cardiovascular fitness and secondary prevention trials. I will be glad to make these available upon request to the interested reader.

Dr. Katerndahl's concern about the possibility of publication bias needs further comment. While this is an important concern of meta-analysis in general, it is unlikely that unpublished studies related to the multivariate relation of physical inactivity and CHD were as problematic as he has suggested. First, this review follows that of Berlin, et al., who attempted to find unpublished data. Second, I reviewed the manuscript with several "content experts" in both cardiovascular epidemiology and exercise science-sports medicine who were unaware of any other unpublished materials. Since the submission of meta-analysis at least four additional publications related to physical inactivity and CHD have been published. The Adventist mortality study[7] showed a decreased mortality with physical activity but did not take into account lipids in its multiple variable analysis. The Alameda County study[8] showed benefit to physical activity in all age groups but only accounted for smoking and body mass index in its adjustments. The MRFIT 10.5-year mortality follow-up study[9] showed essentially identical results as those it had published previously and were included in my original meta-analysis. The Dubbo study,[10] while showing a beneficial effect of physical activity using the "soft endpoint" of any CHD, did not show a protective effect using a definite hard endpoint of myocardial infarction.

The funnel graph depicted by Dr. Katerndahl shows nicely the fact alluded to in my manuscript that the more precise studies with larger sample sizes generally show measures of effect closer to 1 (null hypothesis) than those studies that were less precise and therefore with smaller sample sizes. The relation between sample size and the width of the confidence interval comes from the fact that the standard error of the relative risk is inversely related to the sample size. This fact can be seen visually by reviewing Figure 1 of my original manuscript and recognizing that those studies with the smallest 95 percent confidence intervals are closer to 1, whereas those with the greatest relative risks have large confidence intervals. It is for this reason that the often quoted median relative risk of 1.9 for physical inactivity based upon the review of Powell, et al. is spurious and that the relative risk of 1.37 derived from my weighted estimate accounting for sample size is a better estimate of the true independent relative risk of physical inactivity.

Charles B. Eaton, M.D.
Memorial Hospital of Rhode Island
Pawtucket