RE: "(MIS)USE OF FACTOR ANALYSIS IN THE STUDY OF INSULIN RESISTANCE SYNDROME"

We read with interest the recent *Journal* commentary by Lawlor et al. (1) regarding the utility of factor analysis in understanding the metabolic syndrome (MetS). We agree that the results of studies using exploratory factor analysis should be interpreted with some caution given that it is not, strictly speaking, a hypothesis-testing technique and that certain arbitrary decisions are necessarily made in using this approach. These issues have been reviewed in detail previously in this journal (2). We also agree that, in analyzing the clinical importance of the MetS, it would be of value to assess the ability of factor scores, in comparison with standard definitions, to predict outcomes such as diabetes and cardiovascular disease. In fact, two recent studies compared factor scores against impaired glucose tolerance (3, 4) and standard MetS definitions (4) in the prediction of diabetes.

However, we disagree with a number of Lawlor et al.'s (1) criticisms of the literature regarding MetS factor analysis. First, the authors suggest that factor analysis has been used to determine whether the MetS actually exists. To our knowledge, the goal of the more than 20 studies cited by Lawlor et al. has been to understand the underlying correlation structure of the interrelated metabolic variables thought to constitute the syndrome. It is the other body of research mentioned by Lawlor et al., tests of whether MetS variables coexist to a greater degree than expected by chance (5–7), that might provide evidence supporting the existence of a distinctive syndrome. Even these methods can only demonstrate the lack of independence among the risk factors and the plausibility of a syndrome. In any case, factor analysis is *not* a method for proving or disproving the existence of the MetS.

Second, in presenting their rationale for the suggestion that confirmatory factor analysis is preferable to exploratory factor analysis, Lawlor et al. propose that "knowledge from biologic studies could be used to determine a priori the number of unmeasured (latent) variables that represent the syndrome and thus the number of factors that one would expect to extract" (1, p. 1015). We believe that scientific understanding of the primary underlying pathophysiology

of the MetS is still not developed well enough to allow generation of firm hypotheses regarding the number of latent variables underlying the MetS. This issue becomes especially apparent when one considers the rapidly evolving literature on nontraditional components of the MetS, including inflammation, oxidative stress, endothelial dysfunction, and disorders in adipokine biology. The discovery of adiponectin is an excellent example of this issue (8). Adiponectin, a collagen-like protein secreted exclusively by adipocytes, has a growing list of pleiotropic effects and has been associated (inversely) with many aspects of the MetS, including insulin resistance, dyslipidemia, endothelial dysfunction, inflammation, diabetes, and cardiovascular disease. However, it remains unclear to which of the core MetS disorders adiponectin is primarily linked. While exploratory factor analysis would not provide a definitive, final answer to this question, in our opinion it offers a valuable, unbiased, complementary analytical option to shed light on the complex pathobiology of adiponectin.

Lawlor et al. further suggest that ''in the absence of such biologic evidence, the findings of current (exploratory) factor analysis studies should be formally tested in confirmatory studies using independent data sets, but including the same component variables and applying the same factor analysis procedures'' (1. p. 1015). In this context, it is important to highlight the remarkable uniformity of the more than 20 exploratory factor analysis studies on the MetS published thus far. In epidemiology, consistency of results across diverse studies is usually taken as confirming the validity of population-based observations. That so many ''exploratory'' factor analyses of core MetS variables arrive at the same answer provides convincing ''confirmation'' of the main findings. These studies all strongly support the contention that the MetS comprises two to three distinct dimensions (hyperglycemia, dyslipidemia, hypertension) that share a fourth, obesity-hyperinsulinemia, as a common, central component. Results of empirically derived factor analyses lend substantial credibility to current, widely applied, but more arbitrary definitions of the MetS that require the presence of a few of a broader set of traits to be present to make the diagnosis (9, 10).

Lawlor et al. further state the following: ''None of the studies using factor analysis to explore the [MetS] has stated whether it considers its approach to be exploratory or confirmatory . . . [or used] knowledge from biologic studies . . . to determine a priori the number of unmeasured (latent) variables that represent the syndrome and thus the number of factors one would expect to extract'' (1, p. 1015). Many authors did in fact indicate clearly their rationale for using factor analysis. Furthermore, in one of our own papers in this field (11), we clearly stated that if insulin resistance were the only physiologic domain underlying the MetS, then we would expect to identify just one factor. We found more than one factor by using a principal components model and then used a confirmatory, hypothesis-testing factor analysis to demonstrate that an analytic solution yielding more than one factor was a significantly better fit to the data than a solution yielding only one factor. Although it would perhaps have strengthened this body of research if more authors had also taken this second step, we stand by our contention that the remarkable consistency of results across the broadest array of human populations stands as a powerful epidemiologic confirmation of the findings of these more rigorous early studies.

Next, the authors suggest that the conclusion of many factor analysis papers that ''hypertension may not be linked to the insulin resistance syndrome to the same extent as other components . . . reflects a misunderstanding of factor analysis'' (1, p. 1015). We suggest that there are several biologic reasons why hypertension would not fit closely with the syndrome, including the vasodilative effects of insulin (12, 13) and elements of blood pressure control by the renin-angiotensin, neurohormonal, and arterial vascular systems that are largely independent of insulin resistance and compensatory hyperinsulinemia (14). Furthermore, Lawlor et al. argue that it may be inappropriate to include systolic and diastolic blood pressure together in a factor analysis. In fact, several other pairs of highly correlated variables, including fasting and 2-hour plasma glucose and insulin and different measures of body mass, are very commonly included in the same factor analysis model, and they not only load together in the same factor but also load with other features of the metabolic syndrome (2). As one of the authors of this letter (R. B. D.) has recommended in a widely cited methodological text on factor analysis (15), more than one highly correlated variable (or measures of similar domains) should be included in the factor analytic models to ensure that each domain is contributing enough variance to the model to stand apart if in fact a separate latent trait is present. This is an important element in defining factors using Thurstone's simple structure concept (16). In addition, systolic and diastolic blood pressure should not be considered merely parallel or redundant hemodynamic measures. Rather, they are often only moderately correlated ($r < 0.7$), and they are known to have different physiologic determinants; therefore, they encompass enough biologic uniqueness, analogous to waist circumference and body mass index or high density lipoprotein cholesterol and triglyceride, to separately indicate an underlying MetS domain itself rather than being two overlapping measures of a single MetS variable.

Finally, we would like to provide some clarification regarding definitions and methodological features of confirmatory factor analysis. Lawlor et al. state, ''In confirmatory factor analysis, one should be able to replicate findings from an exploratory study using the same rotation method together with analysis methods being identical'' (1, p. 1016). In this sense of confirmatory factor analysis, one can argue that the consistency across the various papers quoted by Lawlor et al. has established confirmation. Certainly the large body of published studies has demonstrated that more than one single factor (or underlying phenotype) is consistent with the data. However, a more appropriate use of the term confirmatory factor analysis would consist of an a priori statement of the number of factors, a statement of the relative sizes of the factor loadings for the variables within the factors, and then a formal statistical test to evaluate whether a new set of data is consistent with this hypothesized number of factors and the loadings of the variables (rotation is not used in confirmatory factor analysis).

In conclusion, we agree with Lawlor et al. (1) that the limitations of exploratory factor analysis need to be kept in mind when interpreting studies that have used this technique. However, given the rapidly growing literature on the prevalence, outcomes, and underlying pathobiology of the MetS and its relation to insulin resistance, we believe that exploratory factor analysis has made substantial contributions to our understanding of this common and powerful risk factor for type 2 diabetes and cardiovascular disease. Specifically, factor analyses have been a fundamental advance leading to the formulation of credible, clinically useful definitions of a high-risk metabolic state with profound public health implications (9, 10). Widespread application of these definitions has substantially heightened awareness of the importance and implications of the syndrome among the research, clinical, and public health communities.

References

1. Lawlor DA, Ebrahim S, May M, et al. (Mis)use of factor analysis in the study of insulin resistance syndrome. Am J Epidemiol 2004;159:1013–18.
2. Meigs JB. Invited commentary: insulin resistance syndrome? Syndrome X? Multiple metabolic syndrome? A syndrome at all? Factor analysis reveals patterns in the fabric of correlated metabolic risk factors. Am J Epidemiol 2000;152:908–11.
3. Hanley AJG, Festa A, D'Agostino RB Jr, et al. Metabolic and inflammation variable clusters and prediction of type 2 diabetes: factor analysis using directly measured insulin sensitivity. Diabetes 2004;53:1773–81.
4. Hanson RL, Imperatore G, Bennett PH, et al. Components of the ''metabolic syndrome'' and incidence of type 2 diabetes. Diabetes 2002;51:3120–7.
5. Schmidt MI, Watson RL, Duncan BB, et al. Clustering of dyslipidemia, hyperuricemia, diabetes, and hypertension and its association with fasting insulin and central and overall obesity in a general population. Metabolism 1996;45: 699–706.
6. Bonora E, Kiechl S, Willeit J, et al. Prevalence of insulin resistance in metabolic disorders. Diabetes 1998;47:1643–9.
7. Wilson PW, D'Agostino RB, Sullivan L, et al. Overweight and obesity as determinants of cardiovascular risk: the Framingham experience. Arch Intern Med 2002;162:1867–72.
8. Matsuzawa Y, Funahashi T, Kihara S, et al. Adiponectin and metabolic syndrome. Arterioscler Thromb Vasc Biol 2004; 24:29–33.
9. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). JAMA 2001;285:2486–97.
10. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med 1998;15:539–53.
11. Meigs JB, D'Agostino RB Sr, Wilson PW, et al. Risk variable clustering in the insulin resistance syndrome. The Framingham Offspring Study. Diabetes 1997;46:1594–600.
12. Haffner SM. Insulin and blood pressure: fact or fantasy? J Clin Endocrinol Metab 1993;76:541–3.
13. Anderson EA, Hoffman RP, Balon TW, et al. Hyperinsulinemia produces both sympathetic neural activation and vasodilation in normal humans. J Clin Invest 1991;87:2246–52.
14. Reaven GM, Lithell H, Landsberg L. Hypertension and associated metabolic abnormalities—the role of insulin resistance and the sympathoadrenal system. N Engl J Med 1996;334:374–81.
15. Cureton EE, D'Agostino RB. Factor analysis: an applied approach. Hillside, NJ: Lawrence Erlbaum Associates, 1983.
16. Thurstone LL. Multiple factor analysis. Chicago, IL: University of Chicago Press, 1947.

Anthony J. G. Hanley[1,2], James B. Meigs[3], Ken Williams[2], Steven M. Haffner[2], and Ralph B. D'Agostino, Sr.[4]

[1] *Leadership Sinai Centre for Diabetes, Mount Sinai Hospital, Toronto, Ontario, M5G 1X5 Canada*

[2] *Division of Clinical Epidemiology, University of Texas Health Sciences Center at San Antonio, San Antonio, TX 78207*

[3] *General Internal Medicine Unit, Massachusetts General Hospital, Boston, MA 02114*

[4] *Statistics and Consulting Unit, Department of Mathematics and Statistics, Boston University, Boston, MA 02215*