Multivariate Analysis of the Insulin Resistance Syndrome in Women

Karen L. Edwards, Melissa A. Austin, Beth Newman, Elizabeth Mayer, Ronald M. Krauss, Joe V. Selby

Abstract The insulin resistance syndrome (IRS), or syndrome X, is characterized by a constellation of interrelated coronary heart disease (CHD) risk factors, including dyslipidemia, obesity, central obesity, elevated systolic blood pressure, and hyperinsulinemia. Factor analysis was used to investigate the clustering of these risk factors in individuals by examining the correlational structure among these variables. Data from 281 genetically unrelated nondiabetic women who participated in exam 2 (1979 to 1980) of the Kaiser Permanente Women Twins Study were used. Factor analysis reduced 10 correlated risk factors to 3 uncorrelated factors, each reflecting a different aspect of the IRS: factor 1 (increased body weight, waist circumference, fasting insulin, and glucose), factor 2 (increased postload and fasting glucose and insulin and systolic blood pressure), and factor 3 (larger low-density lipoprotein particles, decreased plasma triglycerides, and increased high-density lipoprotein). Together, the factors explained nearly 66% of the total variance in the data. Thus, factor analysis defined three distinct aspects of the IRS in this sample of nondiabetic women. These factors may reflect separate underlying mechanisms of the syndrome, each of which may also be involved in CHD risk. (Arterioscler Thromb. 1994;14:1940-1945.)

Key Words • insulin resistance syndrome • multivariate analysis • lipoproteins • obesity • women

The insulin resistance syndrome (IRS), or syndrome X, is characterized by a clustering of hemodynamic, metabolic, and anthropometric disorders, including hypertension, dyslipidemia, glucose intolerance, and central obesity. Although these disorders cluster within individuals more often than predicted by chance, the metabolic and physiological basis for this clustering remains to be elucidated. As the name implies, insulin resistance itself may play a key role in the syndrome. This concept is supported by epidemiological and clinical studies showing that hyperinsulinemia, a marker for insulin resistance, and insulin resistance itself are both associated with the characteristic disorders of the IRS.

Furthermore, these disorders, including hypertension, dyslipidemia, body fat mass and distribution, elevated plasma glucose levels, and non-insulin-dependent diabetes mellitus (NIDDM), are themselves considered risk factors for coronary heart disease (CHD). Several studies suggest that hyperinsulinemia as well is associated with an increased risk of CHD. Other studies also show that combinations of these risk factors are associated with CHD. Traditional multivariate approaches to studying risk factors for CHD have involved assessment of the independent effects of several risk factors for the IRS. However, the metabolic and physiological relationships among these multiple risk factors may mask important biological associations between aspects of the IRS and CHD. Thus, it may be appropriate to consider the risk factors of the IRS in aggregate, rather than independently, when evaluating CHD risk in an epidemiological study.

Factor analysis provides a method for investigating interrelated variables. On the basis of the correlational structure of quantitatively measured variables, this method can be used to empirically describe the clustering of these variables. Simplifying the characterization of the IRS in this way may lead to new insights into the underlying mechanisms of the IRS and the association of the syndrome with risk of CHD.

The purpose of this study was to use factor analysis to reduce the set of interrelated disorders of the IRS to a smaller number of uncorrelated composite variables, using data from a large sample of genetically unrelated, nondiabetic women.

Methods

Subjects

Study subjects were participants in the second examination of the Kaiser Permanente Women Twins Study in Oakland, Calif. Examination 2 was conducted between 1989 and 1990, and included 704 subjects, 81% of the original cohort examined between 1979 and 1980. At the time of the second exam, each woman completed a health history questionnaire and underwent a physical examination that included anthropometric and laboratory measurements.

Because data from twins in the same pair are not independent observations, the results presented here are based on one randomly selected twin from each pair, leaving the other half of the sample for validation. The average age of the women at...
Data Collection

Height and weight were measured with the women wearing lightweight clothing and no shoes. Body mass index (BMI) was defined as weight (kg)/height (m²). Waist and hip circumference were measured using a standardized format: the waist was measured at the natural indentation or, if that was not readily apparent, at a point midway between the iliac crest and the lowermost portion of the rib cage. Hip circumference was measured at the point of maximal protrusion of the buttocks. The waist-to-hip ratio (WHR) was calculated as waist circumference divided by hip circumference. After each subject had been seated for 5 minutes, a mercury sphygmomanometer was used to measure systolic and diastolic blood pressure. Two measurements, 1 minute apart, were taken on the right arm. The average of these two measurements was used for all analyses.

Plasma glucose and insulin were measured after an overnight fast and again 2 hours after a 75-g oral glucose load (Glutol, Paddock Laboratories). Insulin concentrations were measured by radioimmunoassay using commercial kits (RSL kit for 57% and Pharmacia for 43% of the women in the present study) at SmithKline Laboratories. The correlation coefficient between the two assays was .97. Compared with the Pharmacia kit, however, the RSL kit tended to overestimate insulin levels.* The calibration formula relating the two assays was \( Y = 0.72X - 3.4 \), where \( Y \) = Pharmacia value and \( X \) = RSL value (\( \mu U/mL \)). The formula was provided by SmithKline Laboratories and applied to the present data (Dr B. Scales, SmithKline Laboratories, personal communication, April 1993).

After the overnight fast, 30 mL of whole blood was collected into EDTA-containing tubes for lipid determination. Plasma was separated by centrifugation within 2 hours and stored under refrigeration. Nondenaturing gradient gel electrophoresis was performed on the plasma using 2% to 16% polyacrylamide gradient gels (Pharmacia), as previously described.29,30 The estimated diameter of the major low-density lipoprotein (LDL) subclass was calculated based on a calibration curve constructed from high-molecular weight standards run on the same gel.29 The diameter, denoted LDL peak particle diameter (DLDL-PPD), is a continuous variable and was used in the factor analysis as a measure of LDL heterogeneity. Total high-density lipoprotein cholesterol (HDL-C)31 and triglycerides32 were determined by standardized methods at the Donnor Laboratory (University of California, Berkeley), a participating laboratory in the Centers for Disease Control and Prevention lipid standardization program.

Subjects reporting a physician's diagnosis of diabetes and current use of insulin or oral hypoglycemic agents were classified as diabetic (n = 19) and excluded from all analyses. For 24 subjects postload glucose values, insulin values, or both were missing, and data from these subjects were also excluded from all analyses. Data from an additional 28 women were excluded from analysis for the following reasons: missing data on LDL-PPD (n = 3) or body weight (n = 1); pregnancy (n = 2); currently taking medications for high blood pressure (n = 19); or plasma triglyceride level greater than 400 mg/dL (n = 3). After these exclusions, there were 281 individual women with complete data on all variables used in the analysis. On the basis of World Health Organization criteria,33 242 of these subjects were classified as having normal glucose tolerance and 39 were classified as having impaired glucose tolerance.

Statistical Analysis

All variables included in the factor analysis were first adjusted for age using linear regression. Relationships between individual pairs of risk factors were then examined using Pearson's correlation coefficients. Because of the large number of comparisons, a significance level of \( P < .01 \) was used to identify significant correlations.

Factor Analysis

Factor analysis is used to investigate relationships among several correlated variables by identifying presumed underlying "factors." As described below, factor analysis is essentially a three-step process: (1) extraction of the initial components, using principal component analysis, (2) rotation of the components, resulting in elucidation of factors, and (3) interpretation of the factors.34,35

Principal component analysis. Principal component analysis was used to extract the initial components. By finding linear combinations of the variables that account for the maximum proportion of total variance in the set of variables, principal component analysis transforms the original variables into a new set of uncorrelated components.34 These components are defined as linear combinations of the variables included in the analysis and are determined on the basis of the correlations among them. The first principal component is the linear combination of the variables accounting for the maximum amount of variance in the data, the second principal component accounts for the next largest amount of the remaining variance, and so on. Although there may be as many principal components as original variables, generally the first few components account for most of the total variation in the data set.35 Thus, in concurrence with the literature,34,36 only components with eigenvalues greater than 1.0 were retained in the analysis (the eigenvalue is the sum of the squared factor loadings and represents the amount of variance attributable to each component).34 An important aspect of the principal components is that each component is designed to be uncorrelated with the others; that is, the Pearson correlation between the components is designed to be zero.34

Rotation of principal components. Once defined, the components were rotated to facilitate their interpretation. After rotation the components are referred to as factors. Varimax rotation, an orthogonal rotation,34,36 was used in this analysis to obtain the factors. The advantage of an orthogonal rotation is that it maintains the independence between the factors; that is, the correlation between factors remains zero.34 Because rotation reallocates the proportion of total variance explained by each factor, the amount of variance accounted for by each factor is recalculated after rotation.

Interpretation of factors. Factor loadings, equivalent to a Pearson's correlation coefficient between each variable and each factor, are used to interpret the factors. Interpretation of the factors involves examining which variables load high on a particular factor and then naming the factor accordingly. Only variables sharing at least 15% of the variance with the factor are used for interpretation.34 This corresponds to a factor loading with an absolute value greater than or equal to .40. Although only variables with a factor loading greater than or equal to .40 are used for interpretation, significant correlations \( (P < .01) \) between other variables and the factors, corresponding to a factor loading with an absolute value greater than or equal to .30 as recommended by Stevens,34 are also noted in the analysis. SAS was used for all analyses.37

Results

The characteristics of the study sample are presented in Table 1. Mean triglyceride and HDL levels were well within desirable ranges,38 as were mean systolic blood pressure39 and mean fasting and postload glucose levels.38 Mean BMI in this sample of women was 24.8 kg/m², which is below the level indicating obesity (27.8 kg/m²).40 Although there are no current recommendations regarding WHR, the mean WHR was less than 0.85, the value that has been interpreted as reflecting central obesity in women.41 In general, mean levels of
Table 1. Characteristics of Study Subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value (mean±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>49.7±12.6</td>
</tr>
<tr>
<td>Triglyceride, mmol/L</td>
<td>2.4±1.4</td>
</tr>
<tr>
<td>HDL cholesterol, mmol/L</td>
<td>1.7±0.4</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>114.8±17.9</td>
</tr>
<tr>
<td>Fasting glucose, mmol/L</td>
<td>2.3±0.3</td>
</tr>
<tr>
<td>LDL peak particle diameter, Å</td>
<td>269.8±7.2</td>
</tr>
</tbody>
</table>

All subjects were female (n=281).

HDL indicates high-density lipoprotein; LDL, low-density lipoprotein.

* Corrected for insulin assay.

Table 2. Age-Adjusted Pearson Product Moment Correlation Coefficients of Variables of the Insulin Resistance Syndrome in Individual Women (n=281)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight, kg</td>
<td>1.00</td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td>1.00</td>
</tr>
<tr>
<td>F. Ins</td>
<td>0.88*</td>
</tr>
<tr>
<td>F. Gluc</td>
<td>0.57*</td>
</tr>
<tr>
<td>P. Gluc</td>
<td>0.38*</td>
</tr>
<tr>
<td>P. Ins</td>
<td>0.27*</td>
</tr>
<tr>
<td>LDL-PPD</td>
<td>0.28*</td>
</tr>
<tr>
<td>HDL</td>
<td>0.38*</td>
</tr>
<tr>
<td>Trig</td>
<td>0.38*</td>
</tr>
<tr>
<td>SBP</td>
<td>0.20*</td>
</tr>
</tbody>
</table>

Weight indicates body weight (kg); Waist, waist circumference; F., fasting; Ins, insulin; Gluc, glucose; P., postload; LDL-PPD, low-density lipoprotein-peak particle diameter; HDL, high-density-lipoprotein cholesterol; Trig, triglyceride; and SBP, systolic blood pressure.

* P<.01.

Table 3. Results of Factor Analysis: Factors and Factor Loadings

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body weight, kg</td>
<td>0.92†</td>
<td>0.14</td>
<td>−0.16</td>
</tr>
<tr>
<td>Waist circumference, cm</td>
<td>0.87†</td>
<td>0.24</td>
<td>−0.27</td>
</tr>
<tr>
<td>Fasting insulin, µU/mL</td>
<td>0.61†</td>
<td>0.45†</td>
<td>−0.21</td>
</tr>
<tr>
<td>Fasting glucose, mmol/L</td>
<td>0.41†</td>
<td>0.59†</td>
<td>0.01</td>
</tr>
<tr>
<td>Postload insulin, µU/mL</td>
<td>0.15</td>
<td>0.79†</td>
<td>−0.25</td>
</tr>
<tr>
<td>Postload glucose, mmol/L</td>
<td>0.04</td>
<td>0.86†</td>
<td>−0.13</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>0.17</td>
<td>0.35†</td>
<td>−0.05</td>
</tr>
<tr>
<td>Triglyceride, mmol/L</td>
<td>0.07</td>
<td>0.22</td>
<td>−0.79†</td>
</tr>
<tr>
<td>HDL cholesterol, mmol/L</td>
<td>−0.27</td>
<td>−0.03</td>
<td>0.70†</td>
</tr>
<tr>
<td>LDL peak particle diameter, Å</td>
<td>−0.14</td>
<td>−0.13</td>
<td>0.88†</td>
</tr>
</tbody>
</table>

% Total variance: 22.9, 21.7, 21.2

% Cumulative variance: 22.9, 44.6, 65.8

HDL indicates high-density lipoprotein; LDL, low-density lipoprotein. Components are rotated using Varimax rotation to obtain factors.

* Factor loadings represent the correlation between the individual variable and each factor.

† P<.01.

The results of the factor analysis proved to be significant.

Factor Analysis

The results of the factor analysis are provided in Table 3. The first, the principal component analysis extracted three components with eigenvalues greater than 1.0, and these components were retained for rotation.

Factor 1 is characterized by positive factor loadings for body weight, waist circumference, fasting insulin, and fasting glucose. This factor was named the "Obesity Factor." Factor 2 was characterized by positive loadings for systolic blood pressure, triglyceride, and HDL cholesterol. This factor was named the "Hypertension Factor." Factor 3 was characterized by positive loadings for postload insulin and glucose. This factor was named the "Insulin Resistance Factor."
fasting glucose, postload insulin, postload glucose, systolic blood pressure, and triglyceride. LDL-PPD and HDL were both negatively correlated with factor 1, although neither negative correlation was significant. Only body weight, waist circumference, fasting insulin, and fasting glucose were significantly correlated with factor 1.

Factor 2 is also characterized by positive factor loadings for body weight, waist circumference, fasting insulin, fasting glucose, postload insulin, postload glucose, systolic blood pressure, and triglyceride. In contrast to the case with factor 1, systolic blood pressure, postload insulin and postload glucose were significantly associated with factor 2, while body weight and waist circumference were not. Fasting insulin, fasting glucose, postload insulin, postload glucose, and systolic blood pressure were all significantly correlated with factor 2.

In contrast to the other two factors, factor 3 was characterized by negative factor loadings for body weight, waist circumference, fasting insulin, fasting glucose, postload insulin, postload glucose, systolic blood pressure, and triglyceride. HDL and LDL-PPD were positively and significantly correlated with this factor, suggesting an association with higher HDL levels and larger size LDL particles, respectively. Triglyceride was significantly negatively correlated with factor 3.

The amount of variation in the data accounted for by each of the factors is listed in Table 3: factor 1 accounted for 22.9% of the variance, factor 2 accounted for 21.7%, and factor 3 accounted for 21.2% of the remaining variance. Together, these three factors account for nearly two thirds of the total variance in the data (65.8%).

The Figure is a graphical representation of the factor analysis and its interpretation. Principal component analysis first reduced a set of 10 correlated variables to a set of 3 uncorrelated components; varimax rotation was then used to obtain the factors. As noted in “Methods,” only factor loadings greater than .40 are used for interpretation. Factor 1 is dominated by large positive correlations with measures of body mass and fat distribution, suggesting that this factor reflects a body mass/fat distribution aspect of the IRS. Factor 2 is characterized by the clustering of fasting and postload plasma glucose and insulin, and is thus interpreted as an insulin/glucose factor. Factor 3 is interpreted as a lipid factor, and is characterized by larger diameter LDL, higher HDL, and lower triglyceride.

No substantial differences were found when the analysis was repeated in individual women stratified by smoking status or zygosity or with exclusions for β-blocker use (n=13). In addition, no substantial differences were observed in the factor analysis when BMI was substituted for body weight or when nonwhite women (n=20) were excluded.

### Discussion

To our knowledge, this is the first application of factor analysis to describe the clustering of the disorders characterizing the IRS in a sample of nondiabetic women. Factor analysis reduced 10 interrelated variables to 3 newly defined factors. Because the factors are uncorrelated, each one may represent a distinctly different aspect of the IRS. Factor 1, the body mass/fat distribution factor, was characterized by positive correlations for body weight, waist circumference, fasting insulin, and fasting glucose. Factor 2, the insulin/glucose factor, was characterized by positive correlations with postload and fasting glucose, postload and fasting insulin, and systolic blood pressure. Factor 3, the lipid factor, was characterized by larger LDL-PPD, higher HDL, and lower triglyceride. Because factors are linear combinations reflecting correlations, factor 3 could also be interpreted as smaller LDL-PPD, lower HDL, and increased triglyceride, corresponding to the more established findings.

This data set from twins provided a unique opportunity to validate the results of the factor analysis using a split-sample approach. The initial model was developed in one randomly selected twin and then repeated in the other twin of the pair. Three factors with very similar factor loadings were identified in the second half of the sample. The only differences were modest increases in the factor loadings for HDL, fasting and postload insulin, and a small decrease in the factor loading for fasting glucose on factor 1; and a modest increase in the factor loading for systolic blood pressure on factor 2. These changes did not affect the interpretation of either factor. There were no differences in the factor characterized by lipids (factor 3). The similarity of these results reflects the reliability of the factors presented in this study.

The factors defined in this study are consistent with several well-characterized aspects of the IRS, including elevated triglyceride and low HDL, excess body fat with central fat distribution, relative hyperinsulinemia and hyperglycemia, and elevated systolic blood pressure. Thus, it is tempting to speculate that these factors together describe the IRS, even in these nondiabetic women. The factors may also provide an alternative method of evaluating the association between interrelated features of the IRS and CHD. That is, factor scores could be used as composite risk factors to test for associations with disease in studies in which CHD event data are available.
Although the inclusion of hypertension in the IRS has been challenged,\textsuperscript{49-55} systolic blood pressure did load significantly ($P < .01$) on factor 2, clustering with those characteristics related to insulin/glucose metabolism. This finding appears to be consistent with the hypothesis that hypertension is associated with hyperinsulinemia and insulin resistance.\textsuperscript{23,24,48} However, systolic blood pressure was not highly correlated with the variables included in this analysis (Table 2). Similar results were obtained when diastolic blood pressure was substituted for systolic blood pressure, although the sample size was reduced ($n=277$) because of missing data. Furthermore, when both systolic and diastolic blood pressure were included in the factor analysis, a fourth factor explaining 9% of the total variance was identified. This factor was characterized by large factor loadings for systolic blood pressure (factor loading, .86) and diastolic blood pressure only (factor loading, .86), indicating that blood pressure could be considered as a separate factor.

In this large population-based study, fasting and postload insulin were used as markers for insulin resistance. Although these parameters reflect both insulin resistance and $\beta$-cell secretion of insulin,\textsuperscript{49} they have been shown to be reliable markers of insulin resistance in population studies.\textsuperscript{50} For example, in normal individuals, Laakso et al\textsuperscript{50} found the correlations of insulin resistance, measured by the euglycemic insulin clamp technique,\textsuperscript{51} with fasting and postload insulin levels to both be in the range of $-.58$ to $-.74$. These correlations were reduced in individuals with NIDDM and impaired glucose tolerance ($-.48$ and $-.47$, respectively), although the correlation between insulin resistance and fasting insulin remained significant ($P < .01$).\textsuperscript{50} Because the present study includes only nondiabetic women, insulin levels, particularly fasting levels, serve as reasonable markers of insulin resistance.

Although the underlying mechanisms of the IRS are unknown, several studies suggest that either insulin resistance itself\textsuperscript{34-47} or central fat distribution\textsuperscript{55,52-54} may be the underlying cause. The results of this analysis are consistent with these findings since two of the three factors appear to reflect these abnormalities. The inclusion of fasting insulin and glucose in the characterization of both factors 1 and 2 emphasizes the importance of these two variables in the IRS. One possible interpretation of this finding is that there are multiple pathways through which insulin mediates the clustering of risk factors. In fact, the presence of multiple uncorrelated factors describing different aspects of the syndrome, rather than one, suggests that the IRS may be heterogeneous. Alternatively, one of the factors may be a precursor to another.

Waist circumference and waist-to-hip ratio are both considered proxy measures of central obesity.\textsuperscript{55} When WHR rather than waist circumference was included in the analysis, three factors were again extracted. The only substantial difference was that WHR loaded significantly on all three factors, with positive factor loadings on factors 1 (factor loading, .40, $P < .01$) and 2 (factor loading, .39, $P < .01$), and a negative factor loading on factor 3 (factor loading, $-.45$, $P < .01$). In contrast, waist circumference loaded significantly only on factor 1. Although this finding strongly suggests an important underlying role of fat distribution in the IRS, the use of a ratio measure may create spurious correlations.\textsuperscript{42} In addition, as assessed by computed tomography (CT), waist circumference appears to be a slightly better marker of central obesity than WHR.\textsuperscript{55,56} Although generally not available in epidemiological studies, MRI or CT scan data could be used to more accurately characterize central obesity and its role in the IRS.

In conclusion, by empirically describing the clustering of risk factors using factor analysis, this study has simplified the complex set of interrelated variables characterizing the IRS into three uncorrelated composite variables with meaningful biological interpretations. These findings may provide important insights into the pathophysiology of the IRS, and suggest an alternative way to evaluate the combined effects of correlated risk factors on CHD risk.

Acknowledgments

This work was supported by program project grant HL-18574 and RO1 from the National Institutes of Health, and was partly conducted at the Lawrence Berkeley Laboratory through the US Department of Energy under contract No. DE-AC03-76SF00098. The authors wish to thank Laura Holl for performing gradient gel electrophoresis and Dr Gary Donaldson for providing his helpful comments on the manuscript. This work was conducted during Dr Austin's tenure as an Established Investigator of the American Heart Association.

References

Multivariate analysis of the insulin resistance syndrome in women.
K L Edwards, M A Austin, B Newman, E Mayer, R M Krauss and J V Selby

Arterioscler Thromb Vasc Biol. 1994;14:1940-1945
doi: 10.1161/01.ATV.14.12.1940
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1994 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/14/12/1940

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at:
http://atvb.ahajournals.org//subscriptions/