The Relationship between Insulin Resistance and Related Metabolic Variables to Coronary Artery Disease: A Mathematical Analysis

David Eddy, MD, PhD, Archimedes, Inc.
Len Schlessinger, PhD, Archimedes, Inc.
Richard Kahn, PhD, American Diabetes Association
Barbara Peskin, PhD, Archimedes, Inc.
Rick Schiebinger, MD, Eli Lilly and Company

Corresponding author:
David Eddy, MD, PhD
author@archimedesmodel.com


This is an uncopyedited electronic version of an article accepted for publication in Diabetes Care. The American Diabetes Association, publisher of Diabetes Care, is not responsible for any errors or omissions in this version of the manuscript or any version derived from it by third parties. The definitive publisher-authenticated version will be available in a future issue of Diabetes Care in print and online at http://care.diabetesjournals.org.
**Objective:** People with diabetes have an increased risk for coronary artery disease (CAD). An unanswered question is what portion of CAD can be attributed to insulin resistance, related metabolic variables, and other known CAD risk factors.

**Research Design and Methods:** The Archimedes model was used to estimate the proportion of myocardial infarctions (MIs) that would be prevented by maintaining insulin resistance and other risk factors at healthy levels. Person-specific data from NHANES 1998 – 2004 were used to create a simulated population representative of young adults in the U.S. This population was then entered into a series of simulated clinical trials designed to explore the effects of each risk factor. Each trial had a control arm (all risk factors were allowed to progress without interventions) and a treatment arm (a risk factor was held to its value in young healthy adults). The trials continued for 60 years. The effects of these hypothetical “cures” of each risk factor provide estimates of their impact on CAD.

**Results:** In young adults, preventing insulin resistance would prevent about 42% of MIs. The next most important determinant of CAD is systolic hypertension (SBP), prevention of which would reduce MIs by about 36%. Following SBP are HDL-C (31%); BMI (21%); LDL-C (16%); triglycerides (10%); FPG and smoking (both ~9%), and family history (4%).

**Conclusion:** Insulin resistance is likely the most important single cause of CAD. A better understanding of its pathogenesis and how it might be prevented or cured could have a profound effect on CAD.
Considerable research has been done to understand the effects of insulin resistance on metabolism in different tissues (e.g. liver muscle, fat), inflammation, and other important biological processes. Downstream from these adverse effects of insulin resistance are clinically measured abnormalities such as hypertension and dyslipidemia (e.g. high triglyceride (TG) levels and low HDL cholesterol (HDL-C) levels). Thus, through its effects on these and other variables, insulin resistance could be the underlying cause of much of CAD (1, 2).

To derive the greatest clinical value from the research outlined above it is important to understand the relative role of insulin resistance in CAD compared to other well-known cardiovascular risk factors that may exist independent of insulin resistance. For example, if insulin resistance could be prevented, how much CAD would be prevented? What is the effect of obesity on CAD, as mediated through insulin resistance? What proportion of CAD is caused by other risk factors such as LDL, CRP, and blood pressure, and by non-metabolic risk factors such as age, sex, and race/ethnicity? Answers to these questions are important for identifying targets and priorities for treatments, for understanding the potential effects of interventions that specifically reduce insulin resistance such as exercise and weight loss, as well as for simply understanding of the etiology of CAD.

Ideally, to determine the effects of any particular variable on CAD, one would conduct a clinical trial in which the variable was controlled to its normal level. No such trials exist or are possible until treatments are developed that specifically and only target the variable in question. In the absence of such trials, insights can be gained from combining the results of the research that already exists. This can be done mathematically; mathematical models are the only method available for integrating the results of the research conducted to date to help understand the biological pathways and the relative causes of adverse outcomes.

This paper describes such an approach. Specifically we used existing research to develop a mathematical representation of current theories of the physiological pathways that relate diabetes and other metabolic and non-metabolic variables to CAD. We then used the model to simulate the trials that would ideally be conducted to examine the effects of each variable, if that were possible. We used this approach to estimate the effects on CAD events (fatal and non-fatal myocardial infarctions) of the following variables: insulin resistance, obesity, HDL-C, LDL-C, triglycerides (TG), systolic blood pressure (SBP), smoking, age, sex, and race/ethnicity. We also examined the possible causal roles of free fatty acid (FFA), apolipoprotein B (ApoB), lipoprotein a (Lp(a)), C-reactive protein (CRP), and homocysteine.

RESEARCH DESIGN AND METHODS
Physiology. For this analysis we constructed a model of diabetic dyslipidemia based on pathways described by Ginsberg (3) (Figure 1). Briefly excess energy intake is stored in fat and liver causing central adiposity. This decreases the effects of insulin on uptake of glucose in muscle and fat, and production of glucose by the liver (insulin resistance). This in turn increases plasma glucose levels. Insulin resistance in adipocytes results in greater release of FFA from fat into the circulation. The resulting increased FFA flux to liver increases synthesis of VLDL (4), resulting in increased levels of TG and ApoB, smaller and denser LDL, as well as decreased availability of HDL. FFA also increases insulin resistance in liver (5, 6). While there is some evidence that
FFAs may influence CAD risk through additional mechanisms, such as endothelial dysfunction (7), hypercoagulation, impaired fibrinolysis (8, 9), and increased blood pressure, these effects are not quantifiable at this time. Figure 1 also shows the effects of other variables and risk factors that contribute to CAD.

**General.** The analysis was conducted using the Archimedes model (10, 11). Briefly, the model is a person-by-person, object-by-object simulation written at a relatively high level of biological, clinical, and administrative detail using object-oriented programming. The core of the model is a set of continuous equations that represent the physiological pathways pertinent to diseases, such as those illustrated in Figure 1. Currently the model includes CAD, diabetes and its complications, congestive heart failure, stroke, obesity, smoking, and metabolic disorders in a single integrated model, enabling it to address co-morbidities and syndromes in a realistic way. Variables in the model pertinent to this analysis are shown in Figure 1. For CAD the model includes both gradual and sudden occlusion of coronary and cerebral arteries. The use of differential equations preserves the continuous nature of biological variables.

To conduct simulations, Archimedes creates virtual people, each of whom has his or her own simulated physiology and can get diseases, develop symptoms, seek care, and so forth. To ensure that the virtual people are representative of real people, Archimedes creates copies of real people using person-specific data from datasets such as the National Health and Nutrition Evaluation Survey (NHANES), health risk appraisals, personal health records, and electronic medical records. It does this at the level of detail captured in the dataset including, if available, demographic characteristics, physical examination results, behaviors, family history, current medical conditions, past medical history, biological variables (lab results), symptoms, and current medications. The methods for creating copies ensure that the distributions and correlations of all the important variables are the same in the simulated population as the real population.

In the model, when patients seek care providers apply protocols and follow guidelines for tests and treatments. Test results are functions of the underlying variables being measured, and can have systematic and random errors. Interventions are modeled through their effects on the underlying biological variables. Simulated providers have behaviors that affect their performance and practice patterns. Simulated patients have behaviors relating to seeking care and adhering to treatment recommendations. Because the model is continuous in time, symptoms and the ensuing clinical events can occur at any time and are different for every patient. Care processes pertinent to this analysis were based on guidelines of the American Diabetes Association and American Heart Association.

We validate the model using methods described elsewhere (12, 13). Briefly we use the model to simulate real clinical trials and compare the simulated and real results. To date, this has been done for 48 clinical trials relating to diabetes and CAD. Results of the first 18 trials have been published (12).

**Specific Methods.** The variables and pathways in the model that are most pertinent to diabetes and CAD are shown in Figure 1, with the arrows indicating equations that relate the variables to each other and to the progression of atherosclerosis. In the model the insulin resistance variable represents not only the resistance of fat, muscle, and liver to the effects of insulin, but also the change in production of insulin by pancreatic beta cells (initial beta cell compensation and eventual beta cell fatigue). In the model, insulin resistance affects not only glucose, but other risk factors for CAD, such as SBP, HDL-C,
TG, and ApoB. Equations relating insulin resistance to these variables were estimated from data derived from the UKPDS trial (14, 15) for glucose, from NHANES (1998-2004) for TG and HDL-C, and from data on blood pressures in various populations of people with and without diabetes (16). Other variables and sources are described elsewhere (17).

For this analysis our objective was to estimate the effects of the variables in Figure 1. Our approach was to simulate the clinical trials that would ideally be performed if they were possible: treat each variable one-by-one to its normal value and measure the change in CAD events over a long period of time. The results provide an estimate of the proportion of CAD “caused” by each variable, taking into account its effects on other variables downstream in the physiological pathways (Figure 1). For example, for the simulated trial for insulin resistance we created a hypothetical treatment that maintained the effects of insulin on liver, fat and muscle at their normal levels. The normalization of the effects of insulin then affected downstream variables such as TG and HDL, which in turn affected the development of atherosclerotic plaque and myocardial infarctions.

For the simulated trials we used person-specific data from the 1998-2004 NHANES survey to create a simulated population representative of young adults age 20 – 30 in the US today (from NHANES), with one exception. The average BMI in people age 20-30 in the US is 26, which is generally considered to be overweight. Rather than use this as the value to represent good health, we arbitrarily chose a value of 22.5.

The values for young adults obtained by this method are shown in the first data column of Table 1. We call these “normal” values and use the term “abnormal” to describe values above (or in the case of HDL-C, below) these levels. We use the term “normalize” to describe a protocol that tests everyone annually, identifies people whose values of a variable exceed (or in the case of HDL-C, fall below) the normal value, and then “treat” those people to the normal values. Treatments were hypothetical, designed to control a variable precisely to its normal value. In this sense the treatments were analogous to clamp studies or “knockout” mice. Treatment of any particular variable would affect any downstream variables illustrated in Figure 1. People who have values below (or in the case of HDL-C, above) the normal values were not treated. For the five new variables added to the model, we conducted “what-if” trials in which we calculated their possible effects on CAD on an assumption they are causal, just to determine the possible magnitudes of their effects.

Each simulated trial was conducted using 10,000 individuals, with the same 10,000 simulated people being used for both the treatment and control arms for a given trial. Subjects were followed for 60 years or until they died. All the pertinent variables and outcomes were measured annually. Although a large number of outcomes were recorded, for this analysis we used the cumulative probability of fatal and non-fatal MI (including repeat MIs) as the primary end point.

RESULTS
As calculated by the model, young adults age 20 to 30 in the U.S. today have about a 43% lifetime rate of fatal or non-fatal MIs (95% confidence interval, 42% to 44%). The effects of normalizing insulin resistance on MIs are shown in the right column of Table 1; it prevents approximately 42% of MIs. Figure 2 shows the rates of MIs in people who are destined to get insulin resistance (solid lines) and those who are not (dashed lines). Approximately 50% of young adults are destined to get some degree of insulin resistance, although insulin resistance progresses to the point at which diabetes develops in less than one fifth of them. Those who are destined to develop some degree of insulin resistance face more than three times greater risk of CAD than those who are not. In those people who are destined to develop insulin resistance, curing insulin resistance reduces the risk by approximately 55%.

The effects of insulin resistance are also affected by sex. Today’s young men face a higher rate of MIs than today’s young women, 55% versus 32%. However, insulin resistance plays a larger relative role in women than in men, with normalization of insulin resistance reducing the MI rate about 57% (from 32% to 14%), compared to about 29% (from 55% to 39%) for men.

The effects on MIs of other variables are shown in Table 1. As causes of CAD they range from SBP causing about 36% of CAD, to family history being responsible for about 4%. The five new variables for which causality was assumed are shown in the bottom half of the table. If they are eventually established to be causal, normalizing them should decrease CAD rates by the amounts shown in the table. Otherwise, the values in the table for these variables indicate the proportions of CAD risk for which they are markers.

DISCUSSION

In this study we estimate the proportion of CAD due to insulin resistance, other metabolic variables and other cardiovascular risk factors. Our analysis takes into account the large number of people who develop some degree of insulin resistance, the long time course of developing insulin resistance, the pathological effects of a low degree of insulin resistance, and the effects of insulin resistance on other metabolic variables.

Of the risk factors that we believe are sufficiently well studied to permit quantitative analysis, insulin resistance is the most important single risk factor for CAD. Our results indicate that insulin resistance is responsible for approximately 42% of MIs. Its effect on CAD is indirect, mediated though its effects on other variables such as SBP, HDL-C, TG, glucose, and ApoB. Each of those variables in turn is affected by other variables such as age, sex and race/ethnicity. If each risk factor is considered by itself, the next most important cause of CAD is SBP, normalization of which would prevent about 36% of MIs. After SBP are HDL-C (31%); BMI (at least 21%); LDL-C (16%); TG (10%); FPG (9%), smoking (9%), and family history (4%).

Our analysis also highlights the role of obesity in the etiology of both diabetes and CAD. There is good evidence that obesity is a major cause of insulin resistance, and through insulin resistance obesity affects BP, TG, HDL-C, FPG and ApoB. Just by these effects it is a powerful risk factor for CAD; in our analysis normalizing BMI at 22.5 would prevent more than a fifth of MIs in the US (Table 1). Beyond this, it is possible that obesity has other “direct” effects on CAD that are not represented in our model.

Our results are not directly comparable to the effects seen in clinical trials, where the effects of glucose lowering on CAD were either much smaller (19, 20, 21) or null (22, 23). The reason is that in the
clinical trials the focus was on lowering blood glucose, not preventing or curing insulin resistance. The drugs used in the trials either lowered glucose without affecting insulin resistance (e.g. sulfonylureas, insulin) or lowered insulin resistance to some extent but did not eliminate it (e.g. metformin, rosiglitazone). Furthermore, we normalized insulin resistance over the entire lifetimes of the subjects, whereas the treatments in the trials were given only after people had developed diabetes and given only for the limited durations of the studies. Thus, the results of the trials do not represent the full effect of normalizing insulin resistance and are actually consistent with our results.

Our finding that insulin resistance is responsible for 42% of CAD suggests the possible value of a 100% effective treatment of insulin resistance, should there eventually be one. Although, insulin resistance can be ameliorated by weight loss, our data indicate that other interventions will be needed. Increased physical activity, diet modification and/or drug therapy are obvious approaches, although we could not model the effects of these interventions on insulin resistance because their quantitative effects are unknown or unclear. Also, the more we understand the underlying etiology and effects of insulin resistance, for example its relationship to free fatty acid flux (24) or in the inflammatory process (25), new interventions to prevent or treat insulin resistance or factors upstream from it will likely be developed.

Our results indicate that, because of its effects on IR, curing obesity could be expected to prevent at least 21% of MIs. Thus, interventions that prevent excess weight gain or maintain weight loss should have a major effect on CAD. We could not model the separate effects of visceral obesity, ectopic fat, or other measures of weight related metabolic abnormalities, because of insufficient data.

The main limitation of this analysis is that it is based on a mathematical model, rather than empirical studies. We have tried to make the model as realistic and accurate as possible by reproducing current theories of metabolic pathways, by ensuring that each equation is derived from and validated against empirical evidence, and by testing the accuracy of the full set of equations by calculating the occurrence of diabetes and CAD in a wide variety of clinical trials. Based on these validations it is reasonable to say that the model is entirely consistent with the best available published evidence. Nonetheless, our analysis is limited to variables for which there are sufficiently good data to write and validate equations. In the Archimedes model data have to not only establish a qualitative relationship between variables, but also have to enable writing and validating equations that describe that relationship quantitatively. It is possible that a relationship between variables exists but cannot yet be described quantitatively from the available data.

Our results also depend on the targets chosen for treating the variables. For family history and smoking the targets are obvious -- eliminate the effects of family history and have people stop smoking. But other variables are continuously valued and there are no levels that can unequivocally be designated as "normal", "healthy", or "cured". We had to specify the values to which the variables would be controlled. Possible choices were the thresholds that organizations use to define diseases such as diabetes, or the treatment targets used in national guidelines or performance measures. We chose not to use any of these because in addition to being inconsistent with one another, they are all considerably higher than average values, often representing top quartiles or quintiles, and they typically represent people with moderately advanced disease. Instead we chose to treat the variables to the average values of people in the U.S. age 20-30 on the
assumption that this better represents a healthy, non-diseased state. This modeling exercise has important practical implications. It addresses the relative importance of well-known variables in the genesis of CAD, and suggests areas that should be the focus of research and treatment. More specifically, our results indicate that insulin resistance itself has a profound effect on CAD -- greater than previously realized. In fact it is likely to be most important single determinant of CAD. Additional research into the underlying pathogenesis of insulin resistance, and its downstream effects, prevention, and cure should receive high priority for the prevention of CAD.

Ideally the questions we address in this paper would be answered through empirical research. Unfortunately that is not possible. There is no way to normalize insulin resistance, get everyone to stop smoking, or implement most of the other interventions required. Even if the interventions existed, the empirical studies would be infeasible because of size, duration, cost, and speed of technological change. Yet the questions are undeniably important. A physiology-based model is the best available alternative. It is consistent with and works hand-in-hand with the available research. It converts the observations that have been made, and the theories that have been developed, into a form that can be used to estimate the approximate magnitudes of outcomes, stimulate debate and research, and begin a cycle that should gradually converge on a deeper and more accurate understanding of physiological pathways than would otherwise be possible.

ACKNOWLEDGEMENTS
The authors would like to thank Andrei Chtcheprov, PhD, Joel Scherer, MD, Christine Gathers, Jill Bowman, Julia Dziuba, and Jim Dudl, MD, for their contributions and guidance.
REFERENCES
5. Boden, G: Role of fatty acids in the pathogenesis of insulin resistance and NIDDM. Diabetes 46:3-10, 1997


Table 1. Treatment targets for variables and percent decrease in MIs from normalizing variables to treatment targets from age 20 – 30.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Target for normalization</th>
<th>% decrease in MIs (fatal and non-fatal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td>Average value at age 20-30</td>
<td>42%</td>
</tr>
<tr>
<td>SBP</td>
<td>114 mmHg</td>
<td>36%</td>
</tr>
<tr>
<td>HDL</td>
<td>46 mg/dL (men), 54 mg/dL (women)</td>
<td>31%</td>
</tr>
<tr>
<td>BMI</td>
<td>22.5</td>
<td>21%</td>
</tr>
<tr>
<td>LDL</td>
<td>108 mg/dL</td>
<td>16%</td>
</tr>
<tr>
<td>TG</td>
<td>108 mg/dL</td>
<td>10%</td>
</tr>
<tr>
<td>FPG</td>
<td>86 mg/dL</td>
<td>9%</td>
</tr>
<tr>
<td>Smoking</td>
<td>Never smoke</td>
<td>9%</td>
</tr>
<tr>
<td>Family history</td>
<td>No family history</td>
<td>4%</td>
</tr>
<tr>
<td>FFA</td>
<td>20 mg/dL</td>
<td>18%</td>
</tr>
<tr>
<td>CRP</td>
<td>0.32 mg/dL</td>
<td>10%</td>
</tr>
<tr>
<td>ApoB</td>
<td>85 mg/dL</td>
<td>9%</td>
</tr>
<tr>
<td>Lp(a)</td>
<td>57 mg/dL (blacks), 21.5 mg/dL (non-blacks)</td>
<td>9%</td>
</tr>
<tr>
<td>Homocysteine</td>
<td>7.0 µmol/L</td>
<td>5%</td>
</tr>
<tr>
<td>All variables</td>
<td>See above</td>
<td>94%</td>
</tr>
</tbody>
</table>

Causality not established
Figure 1. Variables and pathways in the Archimedes model pertinent to coronary artery disease. Variables in solid circles are calculated for this analysis.
Figure 2. Effect of “curing” insulin resistance on the expected rate of fatal and non-fatal MIs for people who are destined or not destined to get insulin resistance