Metabolic Syndrome?

A critical look from the viewpoints of causal diagrams and statistics

Eyal Shahar, MD, MPH

Address:

Eyal Shahar, MD, MPH
Professor
Division of Epidemiology and Biostatistics
Mel and Enid Zuckerman College of Public Health
The University of Arizona
1295 N. Martin Ave.
Tucson, AZ 85724

Email: Shahar@email.arizona.edu
Phone: 520-626-8025
Fax: 520-626-2767

Word count: 4,536

Competing interest: none
Abstract

The debate on the metabolic syndrome has attracted great interest among physicians and researchers. This article sheds a new light on the term by using a tool called causal diagrams (also known as directed acyclic graphs). Formal analysis according to causal and statistical principles reveals little substance behind the new syndrome, as well as numerous false claims. From a research viewpoint, continued use of a variable called “metabolic syndrome status” should be discouraged.

Keywords: metabolic syndrome; causal diagrams; prediction
Introduction

PubMed search for the words “metabolic syndrome” in the title of articles and letters has found 175 publications in 2002, 870 in 2005, and 1,431 in 2007. At the time of this writing, the trend might have reached a plateau, counting about 700 titles by mid 2008. Undoubtedly, the term “metabolic syndrome” has found a place of honor on the pages of scientific and medical journals, but has it also survived numerous attacks by critical minds? Moreover, it is difficult to recall another example of a newly discovered, prevalent syndrome whose very existence had to be defended, repeatedly.

In this article, the term "metabolic syndrome" is analyzed from two related viewpoints: causal and statistical. To shed a new light on the debate, a simple tool called causal diagrams is used. Causal diagrams, formally known as directed acyclic graphs (DAG), encode causal assertions unambiguously; mercilessly expose foggy causal thinking; and create a bridge between statistical associations and causal reality. In epidemiology, for example, this tool proved to be a unified method to explain the main categories of bias: confounding, selection bias, and information bias.

The article is divided into two parts: The first part lays essential theoretical foundation. In the second part, various aspects of the new syndrome are critically examined.
Part I: Theoretical Foundation

Causal diagrams

The essence is simple. We write down the names of variables and draw arrows to connect them such that each arrow emanates from a cause and points to its postulated effect. For example, “smoking status→lung cancer status” encodes the idea that smoking causes lung cancer. The sequence “weight→insulin resistance→vital status” encodes the theory that weight affects survival through an intermediary variable called insulin resistance.

“HDL cholesterol←gender→hemoglobin” encodes the statement gender affects both HDL-cholesterol and hemoglobin. The variables in question may be binary, nominal, ordinal, or continuous, but they must be variables and not values of variables. For example, formally we should not write “smoking→lung cancer” because “smoking” and “lung cancer” are not variables.

Causal diagrams assume an underlying causal structure, which percolates up to create the familiar statistical associations between variables.\textsuperscript{13} For instance, we observe a statistical association between pack-years of smoking and survival because “pack-years of smoking→vital status”. Many statistical associations, however, do not reflect the cause-and-effect of interest. One key explanation for observing an association between two variables is their sharing of at least one common cause. For example, fasting blood glucose and resting blood pressure are associated, at least in part, because weight affects both. And in general: a crude association between two variables contains both the effect of one on the other (if any) and the contribution of their common causes (if any). In causal inquiry, these common causes are called confounders. Their contribution to the crude association is called confounding.
Natural variables and derived variables

Some variables may be called natural in the sense that “nature has created their values through various causal mechanisms, and we just try to measure those values.” Fasting glucose level and weight are examples of natural variables (although their measured version already contains the influence of human measurement.) Trisomy 21 (present, absent) is another example. At the other extreme we find human-made variables in the sense that “we, rather than nature, are the ultimate reason for their existence.” Body mass index (BMI), for instance, is not a natural variable because we create the content (values) of that variable from the measured version of two natural variables: weight and height. Stated differently, natural variables are measured, whereas their human-made counterparts are derived from natural variables (and sometimes from other derived variables.) The derivation could be carried out by an arithmetic expression (BMI=weight/height²) or by conditional statements (If fasting glucose<C, then diabetes status is “no diabetes”; otherwise diabetes is present). There are intermediate kinds of variables as well: pack-years of smoking is a natural variable, quantifying lifetime smoking exposure, but we typically derive a proxy from the average number of cigarettes smoked per day and the number of years smoked.

Medicine is rich in human-made, derived variables, many of which originate in continuous variables. Take a measurement of a continuous trait, such as blood pressure, convert the result to a binary or an ordinal variable on the basis of some cutoff point(s), and you have created a human-made variable, perhaps “hypertension status”. Reporting the so-called upper limit of normal is another example.
Deriving a variable usually carries some penalty, but it is sometimes essential. Much of medical practice consists of categorical decisions—to act one way or another, or not to act—and physicians try to make those decisions on the basis of external information, which is often inherently continuous. They have to derive categorical variables because there is no other practical way to import continuous information into the realm of categorical decisions.

Consider a simple, familiar example: To prescribe an oral hypoglycemic drug to an asymptomatic patient, a line must be drawn between levels that “need treatment” and levels that “do not need treatment”. In other words, a binary variable (diabetes status) must be derived from a continuous trait. Blood pressure and hypertension treatment make up another well-known example, and there are many more. As a side note, it may be interesting to recall countless debates about the right way to categorize a continuous trait. Categorization is sometimes unnecessary and other times—a necessary evil. But it is almost never “right” for at least one reason: no matter where the line is drawn, adjacent points on opposite sides of the line are forced to be very different, and that is rarely true, if ever.\(^{18}\)

**Derived variables and causal diagrams**

Thinking about cause-and-effect usually invokes the idea of a relation between two natural variables where the values of one affect the values of the other. Set weight to be 300lb, rather than 150lb, and chances are that fasting blood glucose will rise. But there is no reason to exclude derived variables from the domain of causal connections. Their creation is a form of causation, just like the “creation” of fasting glucose by weight. Set the weight of a 5-foot person to be 300lb, rather than 150lb, and BMI will rise. The rules of causal diagrams, therefore, apply.
expression “BMI=weight/height^2” is encoded just as any other causal relation between two causes and their common effect: “weight→BMI←height”. Similarly, “fasting glucose level→diabetes status” encodes the derivation of a variable called “diabetes status” according to conditionals about fasting glucose and cutoff points.

**Predicting effects from their causes**

There is one important empirical difference, however, between causal relations among natural variables and causal relations that involve derived variables. No set of causal variables will precisely predict the fasting glucose level (a natural variable) of any patient, either due to unknown causes or because causation is inherently indeterministic. In contrast, the patient’s diabetes status (a derived variable) is fully determined by his or her level of fasting blood glucose because someone set up a causal mechanism—the derivation rule—to link the two. Likewise, no set of causal variables will precisely tell us anybody’s weight, but weight and height will precisely determine the value of BMI.

Which leads to the following conclusion: the information that is contained in a derived variable is usually present in the variables from which it was derived. Only in special causal circumstances, derived variables might carry new information and thereby predict something beyond their makers (e.g., U=V^2 for a quadratic dose-response function or U=V_1*V_2 for an interaction). For many kinds of derivations, it is difficult to find theoretical arguments for incremental prediction—much less so for substituting a derived variable for the original information.19,20
Part II: Analysis

Deriving “metabolic syndrome status”

For some writers the metabolic syndrome was discovered; for others it was defined; and for others it was made up. Technically, however, “metabolic syndrome status“ is a derived variable. Actually, there are dozens of derived variables that claim the title—as many as there are proposed definitions, or more correctly, as many as there are rules of derivation. ²¹⁻²⁸

Almost every proposed derivation of metabolic syndrome status follows the same format.³ Let \( V_1, V_2, \ldots, V_n \) denote a set of \( n \) continuous variables, either natural or derived. For each variable, decide on a cutoff point and derive a binary variable (0, 1) on the basis of that cutoff point and a conditional. Next, add up the values of these binary variables to derive a summation variable, say, SUM. Finally, derive “metabolic syndrome status” from SUM using a cutoff point and a conditional: if SUM<\( k \), then the metabolic syndrome is absent; otherwise, the metabolic syndrome is present.

Figure 1 shows the causal diagram of the process for \( n=5 \), which is a common number of input variables for writers about the new syndrome. Moving from left to right along the axis of time, we find four generations of variables. Almost all of the variables in the first generation are natural, but all subsequent generations are derived. As shown in the figure, the immediate cause of “metabolic syndrome status” is SUM, whose causes are five derived binary variables.
That someone derived a variable indeed makes it exist, but existence *per se* is not a big
achievement in this case. Derived variables exist in the trivial sense that "we created them from
some other variables". No special insight is needed to follow the process shown in Figure 1: it
requires no more than a group of variables, perhaps a group that has something in common, and
a derivation algorithm.

One matter may, therefore, be settled at this point. Regardless of whether *the one and only*
metabolic syndrome does exist (in some yet unclear sense), what surely exists are many derived
variables that carry the title. Rather than naming them after endorsing organizations, it is
better to use numerical subscripts to indicate the chronology of the proposed rules: “metabolic
syndrome status₁”, “metabolic syndrome status₂”, “metabolic syndrome status₃”, and so on. The
sequence has no meaningful order other than chronology, and may continue indefinitely.

**Clustering of risk factors**

Almost every writer about the metabolic syndrome, whether a proponent or an opponent,
mentions the clustering of risk factors as a key feature of the syndrome. For example, a group of
proponents writes: "Five risk factors of metabolic origin (atherogenic dyslipidemia, elevated
blood pressure, elevated glucose, a prothrombotic state, and a proinflammatory state) commonly
cluster together". Likewise, a group of opponents writes: "The term 'metabolic syndrome'
refers to a clustering of specific cardiovascular disease (CVD) risk factors..." What is
clustering, however? What statistical idea underlies that powerful word?

A patient with high blood pressure is more likely to have a high level of blood glucose than a
patient with low blood pressure, and a patient with low blood pressure is more likely to have a low level of blood glucose than a patient with high blood pressure. But no one would say that blood pressure and blood glucose cluster or "cluster together". We would say that these traits are correlated or associated. Even if a third variable is added, say plasma triglycerides, which correlates with both, we would still not use the word "cluster" because it is not used in the context of continuous variables. The word is reserved for categorical variables, selectively pointing to one aspect of a well-known statistical idea: association.

Let Binary $V_1$, Binary $V_2$, ..., Binary $V_n$ be a group of binary variables each taking the values of 1 ("bad, high risk") or 0 ("good, low risk"). Clustering is said to exist if patients with a value of 1 on any one variable are more likely to have a value of 1 on all others (than patients with a value of 0 on that variable). If that is the case, however, zero values cluster, too: patients with a value of 0 on any one variable are more likely to have a value of 0 on all others (than patients with a value of 1 on that variable). For example, and regardless of mechanism, patients who don't smoke are more likely to not drink than patients who do and vice versa (clustering of no smoking and no drinking). To sum up, clustering is a word to describe a group of categorical variables, usually binary, where each variable is associated with all others.

The phrase "clustering of risk" (implying “clustering of high risk”) may be rhetorically helpful, but it is nonetheless poor scientific terminology for several reasons: First, why talk about clustering of one value of variables (“high risk”) when the underlying statistical phenomenon is an association between variables? Second, the complement, favorable clustering of the other value ("low risk") is conveniently ignored—hardly an objective representation of statistical reality. Third, there is a better, core description of the phenomenon behind the so-called
metabolic syndrome: several *natural, continuous* variables are associated with each other (for reasons that will be discussed later).

Indeed, opponents of the syndrome have already reduced the "clustering" into common statistical jargon: "...certain 'metabolic' factors tend to associate with each other..." \(^5\) Similar, though less clear, expression may also be found in the writing of a proponent: "multiple risk factors that are metabolically interrelated". \(^11\) Surprisingly, however, numerous writers from both camps have also adopted a pseudo-statistical idea—that the observed clustering exceeds the clustering that would be expected by chance alone. Although we can estimate the magnitude of an association between variables and perhaps gather evidence against the claim of “no association”, no statistical computation can tell us how strong of an association is expected by chance alone, because chance alone could account for *any* association. That erroneous idea can probably be traced to prevailing misinterpretation of a *P*-value as "the probability of observing this result by chance".

**Possible causal mechanisms behind "clustering"**

Having translated the "clustering phenomenon" to “multiple associations among derived, binary variables”, we may now turn to the scientific questions of interest: Why Binary V\(_1\), Binary V\(_2\),...,Binary V\(_5\) are all associated with each other? Which causal mechanisms have created associations among these derived variables? Why are they "interrelated" or plainly related?

As explained earlier, two main mechanisms contribute to an association between two variables: 1) one variable causes the other; 2) they both share at least one common cause. (A
third mechanism will be mentioned later.) As shown in Figure 1, the first mechanism does not operate in that diagram: no causal arrow emanates from any binary variable and points to another—and rightly so. The only immediate causes of a derived variable are the variables from which it was derived. We may therefore conclude that the observed associations among the binary variables in Figure 1 must be attributed to their sharing of at least one common cause, which is missing from the figure. The causal diagram in Figure 1 must be incomplete.

Figures 2-4 show minimal causal structures that would create an association between each of the five binary variables behind metabolic syndrome status and the other four. To check the claim, we just need to verify that each pair shares at least one common cause. Indeed, if we pick any two binary variables and follow their arrows “upstream” to their causes, we will always end up in a common cause. In Figure 2, the common cause is U; in Figure 3 it is U, too (as well as $V_4$ for the pair Binary $V_4$ and Binary $V_3$); and in Figure 4 it is $V_1$. Notice that in each case, the explanation for the associations among the binary variables has nothing to do with these variables per se; everything happened between natural variables at earlier stages of causation.

According to Figure 2 or Figure 3, one U will explain the “clustering of metabolic risk factors”, and it is not too difficult to name at least two candidates: age and maybe the amount of abdominal fat. As for Figure 4, one of the five continuous variables should assume the role of a common cause of all other. That variable may also be abdominal fat, whenever it is part of the derivation of metabolic syndrome status. The clustering mystery is finally and trivially solved.
Causal structures that cannot cause "clustering"

Figures 5-8 show several examples of causal diagrams with no shared cause of all five binary variables. For reasons that are well-established in the theorems of causal diagrams,\textsuperscript{13,14} these causal structures will not create an association between every pair of the five binary variables. For example, Binary V\textsubscript{3} and Binary V\textsubscript{4} would not be associated in Figure 5; Binary V\textsubscript{2} and Binary V\textsubscript{4} would not be associated in Figure 6; Binary V\textsubscript{3} and Binary V\textsubscript{5} would not be associated in Figure 7; and Binary V\textsubscript{1} and Binary V\textsubscript{5} would not be associated in Figure 8.

Proponents of the metabolic syndrome state that no common cause is needed. Needed for what? No common cause is needed to derive any variable, including one called “metabolic syndrome status”, but a causal structure with no common cause could not have created that “clustering”, which was the motivation for deriving metabolic syndrome status in the first place. Moreover, at least one common cause does exist (age) and maybe there are others.

Spurious “clustering”

Two variables that do not cause each other, nor share a common cause, may still be associated due to a third, less well-known mechanism called selection bias.\textsuperscript{15} In brief, selection bias arises from unnecessary manipulation of a common effect of the variables of interest. For example, two variables that are not associated at all will be associated within at least one stratum of a third variable, if that variable is their common effect.\textsuperscript{14} The observed association has no interesting causal meaning: it reflects neither cause-and-effect nor a common cause (confounding).
As shown in all figures, "metabolic syndrome status" is a common effect of all preceding variables. Therefore, stratifying on this variable might create associations between components of the syndrome among patients who are classified as having the syndrome. Stated in the "clustering" jargon, part of the observed clustering of metabolic risk factors among patients who carry the label "metabolic syndrome" is likely spurious, attributable to stratification on the derived variable. It is misleading to examine the clustering in patients who received the label, or in patients who did not.

"The combined effect is more than the sum"

This loose idea shows up occasionally, usually referring to "summation" versus "multiplication" of the effects of the natural variables from which metabolic syndrome status is derived. The underlying concept is called interaction by statisticians, or effect modification by epidemiologists, and like other methodological topics, it is much deeper and more subtle than is usually appreciated.\textsuperscript{29-32} For example, the phenomenon depends on the scale on which associations are measured and almost always exists on some scale.\textsuperscript{33,34}

Regardless, the so-called metabolic syndrome is neither the sum nor "more than the sum" of its components. There is no theoretical basis for the claim that deriving a binary variable from five continuous variables will somehow capture their combined effect, or a complex structure of multiplicative or additive interactions among them. Interactions among variables are modeled by interaction terms, not by reducing five continuous variables to one binary variable through cutoff points and derivation rules.
A predictor or a risk factor? Both or neither?

One contentious topic has been the ability of the metabolic syndrome to predict outcomes, or more precisely, the ability of a derived variable called "metabolic syndrome status" to predict outcome status, above and beyond components of the syndrome. Discussing stroke as a possible outcome, one writer has summarized the issue in two questions: Is the metabolic syndrome a risk factor for stroke? Does metabolic syndrome status help to predict stroke?

Figure 9 is a revision of Figure 2, adding stroke status as an outcome variable. The arrows pointing from $V_1$, $V_2$, $V_3$, $V_4$, and $V_5$ to stroke status correspond to prevailing theories: components of the metabolic syndrome are causes of stroke (or risk factors, in another jargon). No arrow, however, emanates from metabolic syndrome status and points to stroke status, because causal ordering may place the stroke before a human mind decided to derive the variable (before its values were created.) Therefore, "metabolic syndrome status", in any of its versions, is not a risk factor for stroke. Almost every association between this variable and an outcome is due to confounding by their common causes—for example by $V_1$, $V_2$, $V_3$, $V_4$, $V_5$, and $U$ (Figure 9).

Moreover, even if someone entertains a causal arrow between the two variables, metabolic syndrome status would simply be an intermediary variable on five causal pathways from natural variables to stroke status. All of its "effect" on stroke is already contained in the effects of its causes. In fact, only part of their effects is captured by the metabolic syndrome because each of the five variables is also connected to stroke status through another causal pathway (Fig 9).
The terms "predictor" and "risk factor (cause)" are not synonyms, but the syndrome fails the prediction test, too. As explained earlier, the information carried by a derived variable is usually contained in the variables from which it was derived, except for special causal circumstances. Therefore, metabolic syndrome status should not predict anything beyond its makers, as found empirically. Components of the syndrome might predict stroke status better if we model interactions among them, but why would a single, derived binary variable reflect all those interactions?

**Circular causation**

The issue of circular causation is minor, but worth explaining, since it might be used to criticize the diagrams that were presented here. Causal diagrams are formally called directed *acyclic* graphs because a cycle of causation is not permissible. We may not draw a diagram in which we can make a full cycle along a causal chain, returning to the cause from which we started. Self-causation does not exist because a future effect cannot be a cause of its cause in the past.

Although controversial, some writers raise the possibility of vicious cycles among components of the metabolic syndrome. For example, abdominal fat ➔ insulin resistance ➔ abdominal fat ➔ insulin resistance, and so on. Nonetheless, the second showing of the variables "abdominal fat" and "insulin resistance" in that chain are new variables, and therefore the sequence requires subscripts to denote time-dependent variables: abdominal fat\(_1\) ➔ insulin resistance\(_2\) ➔ abdominal fat\(_3\) ➔ insulin resistance\(_4\), and so on. The so-called circular causation is not circular at all: abdominal fat at time 1 affects insulin resistance at time 2 which affects abdominal fat at time 3, not abdominal fat at time 1. Measurements of time-dependent variables are often taken in
longitudinal studies, allowing to estimate effects in such causal chains.\textsuperscript{37}

**What is a syndrome?**

"Thus, if the metabolic syndrome is defined as multiple risk factors that are metabolically interrelated, then the syndrome certainly exists."\textsuperscript{11}

This argument is linguistically clever: the syndrome exists because it is defined as "multiple risk factors that are metabolically interrelated", which empirically exist. In other words, if something exists and it was given a name, then the name it was given exists, too. But the question remains: what is a syndrome?

Online dictionaries offer numerous explanatory phrases, most of which seem to share one key idea: a syndrome is more than a collection of symptoms, signs, or physiological traits—more than "metabolically interrelated" variables. Both dictionaries and common medical usage require that a syndrome would "indicate", "characterize", or "be characteristic of" a disease, a medical condition, a particular underlying abnormality, and the like. It is not correlated components per se that make up a syndrome, as implied in the quote above, nor their sharing of a common effect. It is a meaningfully deeper abnormality which has caused them.

Sometimes, that deeper abnormality is a well-established cause of the syndrome (e.g., trisomy 21 behind Down syndrome; HIV infection behind AIDS). Other times the cause is a general pathological descriptor (e.g., acute myocardial ischemia behind acute coronary syndrome). In many instances no causal pathways are known yet, but even then we assume that some
"interesting" causal mechanism has generated the syndrome and we hope to discover it some day. For this reason, we don't call every set of age-related, correlated medical conditions, such as dementia, osteoporosis, and atherosclerosis—the "aging syndrome".

This cardinal feature of a syndrome is lacking in the so-called metabolic syndrome. Furthermore, an articulate proponent has stated that the term “does not commit to a particular pathogenesis” (and proposed an ambiguous causal distinction between “underlying causes and exacerbating factors”).\textsuperscript{11} Theoretically, he would not abandon the name even if age alone underlies the associations among the various components. A name is just a name, of course, but the so-called metabolic syndrome seems to propose a new meaning for the term "medical syndrome".

**The metabolic syndrome vis-à-vis the insulin resistance syndrome**

The origin of the new syndrome is often traced to the insulin resistance syndrome, which was postulated long ago.\textsuperscript{38, 39} The idea was both clever and simple: perhaps resistance to the action of insulin is one of the determinants (causes) of several continuous physiological traits, such as glucose tolerance, blood pressure, plasma triglycerides, and HDL-cholesterol. That does not mean, of course, that every patient with unfavorable levels of these variables suffers from insulin resistance, but it does raise the possibility that some patients do. Unlike the metabolic syndrome, the name was not merely a reduction of natural continuous variables to a derived binary variable, nor was it a means for labeling patients. It was a scientific hypothesis to be tested, corroborated, refuted, or perhaps revised.\textsuperscript{40} If true, we have enriched our understanding of the pathogenesis of several risk factors for cardiovascular disease.
In contrast, deriving the metabolic syndrome variable offers nothing scientifically new—neither fresh insight into pathogenesis nor a new daring hypothesis. Not surprisingly, much of the intellectual energy is spent on tangential matters: In what sense does the syndrome exist? What should we call it? What are the cutoff points? Whose definition will prevail? How do we promote another "worldwide definition"?21, 28, 41, 42

The merit of the derivation

What is left of the term? Are there any benefits to deriving that binary variable by one set of rules or another, and deciding whether a patient "has it"?

Proponents argue that the label would motivate patients to change risky behaviors and cause doctors to pay greater attention to risk factor modification, certainly a reasonable hypothesis. Opponents argue that patients who missed the labeling would have a sense of complacency and might do less to change their risk factor profile—another reasonable hypothesis. In the coding of causal diagrams both theories take the following general structure: “Metabolic syndrome status → knowledge of metabolic syndrome status → risk factor level → outcome”. Unfortunately, it is difficult to imagine a study that would estimate the net effect.

Two other merit-related questions might be asked: First, how many of those thousands of publications that contain the words "metabolic syndrome" in their titles would have been published if the term did not exist? Second, how much less we would have known today?
References

18. Greenland S. Dose-response and trend analysis in epidemiology: alternatives
Figure 1. A directed acyclic graph showing the causal structure behind the variable “metabolic syndrome status”
Figure 2. A directed acyclic graph showing a causal structure that would create an association between every pair of the five binary variables.
Figure 3. A directed acyclic graph showing a causal structure that would create an association between every pair of the five binary variables.
Figure 4. A directed acyclic graph showing a causal structure that would create an association between every pair of the five binary variables.
Figure 5. A directed acyclic graph showing a causal structure that would NOT create an association between every pair of the five binary variables.
Figure 6. A directed acyclic graph showing a causal structure that would NOT create an association between every pair of the five binary variables
Figure 7. A directed acyclic graph showing a causal structure that would NOT create an association between every pair of the five binary variables.
Figure 8. A directed acyclic graph showing a causal structure that would NOT create an association between every pair of the five binary variables.
Figure 9. A directed acyclic graph showing several causal pathways to stroke status.