An approach was recently proposed for grouping similar sufficient causes into classes and estimating the proportion of disease that is due to these classes (1). However, the interpretation of the classes formed was partly misleading. We want to clarify the meaning of the derived classes of sufficient causes and to illustrate the former misinterpretation.

The recently published method (1) provides a possible approach to estimate the proportion of disease attributable to different classes of sufficient causes or, equivalently, to different combinations of risk factors. However, utmost caution is needed in interpreting the classes considered. The classes defined do not consist of sufficient causes of which each has the same combination of known risk factors as component causes, but rather they consist of those sufficient causes that can be additionally completed under the same, given combination of risk factors. Although that classification and the alternative one indicated above seem to be similar, a subtle distinction in grouping of sufficient causes exists.

We consider a simple example of two dichotomous risk factors, \( X_1 \) and \( X_2 \). For simplicity, assume that only two types of sufficient causes exist: sufficient causes that are completed if \( X_1 \) and an unknown component \( U_1 \) are present \((S_1)\) and sufficient causes that are completed if \( X_1, X_2 \), and an unknown component \( U_{12} \) are present \((S_{12})\). These two types or classes of sufficient causes are formally represented by the familiar pie charts in the top part of figure 1 and are characterized by their component causes.

In the paper of Hoffmann et al. (1), another classification scheme was defined by grouping sufficient causes that can be completed under the different combinations of present and absent risk factors (formula 4 (1)). This classification is illustrated by pie charts with overwritten text indicating the combination of present factors (bottom of figure 1). Note

**FIGURE 1.** The relation between the two classification schemes for sufficient causes in a simple example of two classes \((S_1\) and \(S_{12}\)) in the case of two known risk factors \((X_1\) and \(X_2\)). The \( U \) symbols represent one or more unknown component causes. \( S_{E(1,0)} \), class of sufficient causes of only type \( S_1 \); \( S_{E(1,1)} \), class of both sufficient causes \( S_1 \) and \( S_{12} \).
that the risk factors are not assigned to circular sectors of the pie chart because they may not be necessary for completion of the sufficient causes. In this simple example, two classes, $S_{E(1,0)}$ and $S_{E(1,1)}$, exist that are different from $S_1$ and $S_{12}$. Whereas the class $S_{E(1,0)}$ consists of sufficient causes of only type $S_1$, the class $S_{E(1,1)}$ formally consists of both sufficient causes $S_1$ and $S_{12}$. However, if people are classified according to their combination of risk factors, the resulting groups are mutually exclusive because the associated combinations of risk factors are defined that way.

Hoffmann et al. (1) applied their method to evaluate causes of myocardial infarction. The proportions of disease due to four classes are given in figure 2 of their article (1) and can be interpreted as follows. Assuming causality, each number estimates the percentage of cases that would not have occurred if all individuals with the indicated multifactorial exposure profile had had the low risk combination of all risk factors. For example, if those with all four risk factors had instead none of the risk factors considered, 24.1 percent of the cases could be avoided. Some of the cases may be due to sufficient causes that do not require all of the risk factors, even if present, as component causes.

Identification of which sufficient causes are responsible for most of the disease burden is an important and attractive task that can be helpful for improving and directing prevention strategies in public health. As illustrated, classification of sufficient causes based on combinations of causal components requires careful interpretation. However, classification of risk factors by the combinations present readily allows an evaluation with public health relevance.

**ACKNOWLEDGMENTS**

The authors thank Dr. Kenneth Rothman for his helpful comments.

Conflict of interest: none declared.

**REFERENCE**


Kurt Hoffmann$^1$ and W. Dana Flanders$^2$ (e-mail: khoff@dife.de)

$^1$ Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke, 14558 Nuthetal, Germany

$^2$ Department of Epidemiology, Rollins School of Public Health, Atlanta, GA 30322

**Editor’s note:** In accordance with Journal policy, Hoffmann et al. were asked whether they wished to respond to this letter, but they chose not to do so.

DOI: 10.1093/aje/kwk082; Advance Access publication November 3, 2006