CHAPTER 3.6

MULTIVARIATE ANALYSIS OF MAJOR RISK FACTORS AND CORONARY HEART DISEASE RISK IN THE SEVEN COUNTRIES STUDY

Alessandro Menotti
Multivariate analytical techniques in the Seven Countries Study were first made using cross- classifications compacted by the Mantel-Haenszel method (1). This used multiple cells in a cross-classification to take into account several risk factors simultaneously (each subdivided into a few classes) in relation to an outcome variable, such as mortality from CHD. Observed and expected cases were computed for a summary chi-square with one degree of freedom. Variances were calculated of the number of cases (events) in specified classes from sums of the expected and observed cases. However, cross-classification of risk factors became increasingly difficult with a greater number of risk factors, so a new approach was needed in multivariate analysis.

The Framingham Study led in development of the multiple logistic function for predicting morbid or fatal events. A milestone first paper on this technique was published in 1967 (2), and other relevant reports came later (3,4). The multiple logistic function has the advantage of treating the dependent variable (the event) as a probability. Mathematically speaking, the multiple logistic function represents the extension to many risk factors of the computation of odds ratios derived from simple 2 x 2 contingency tables.

The Seven Countries Study has had a unique opportunity to compare risk functions derived from populations differing in geography, culture, risk factor levels and burden of morbid and fatal events that started with the 5-year follow-up data (5). Moreover, it was the first study to produce the back application of risk functions from one population to risk factor distributions of another, and to evaluate the precision and force of different risk estimates – absolute vs. relative.

With the prolongation of the Seven Countries Study follow-up period beyond 10-15 years, it became necessary to use the Cox proportional hazards model (6), which is based on life table principles. In fact, the multiple logistic function does not take into account the time elapsed between measurement of risk factors and events and the same weight is given to persons who die early or late after the measurement of the risk factors. When the follow-up period is prolonged, the number of subjects with events not related to the one explored in the analysis increases largely. The problem is what to do with these events. If these cases are excluded from the computation, the structure of the population is crippled, and if they are left in the model, the predictive power of risk factors can be distorted (usually diluted) if they are also associated with events of primary interest. The Cox proportional hazard model requires the date for each event, and assumes that the risk is proportionate for different levels of risk factors. Tests are performed to show that this assumption is met, and the shape of the function is non-parametric, since no parameters are included to describe the function itself.

Most recently, attempts have been made to use the accelerated failure-time model, that is, a log-linear model incorporating the Weibull distribution (7), which includes parameters defining the shape of the risk function. This model allows us to follow with time the possibly changing shape and acceleration (or deceleration), of the hazard. An example of this model is given in the chapter on mortality. In this chapter an overview will be given of the results of multivariate analyses in predicting CHD risk performed during the course of the follow-up.