

**Building prediction models for coronary heart disease by synthesizing multiple longitudinal research findings.** *European Journal of Cardiovascular Prevention and Rehabilitation*, 12:459-464 (2005).

**ABSTRACT**

Background: No methodology is currently available to allow the combining of individual risk factor information derived from different longitudinal studies for a chronic disease in a multivariate fashion. This paper introduces such a methodology, named Synthesis Analysis, which is essentially a multivariate meta-analytic technique.

Design: The construction and validation of statistical models using available data sets.

Methods and Results: Two analyses are presented. (1) With the same data, Synthesis Analysis produced a similar prediction model to the conventional regression approach when using the same risk variables. Synthesis Analysis produced better prediction models when additional risk variables were added. (2) A four-variable empirical logistic model for death from coronary heart disease was developed with data from the Framingham Heart Study. A synthesized prediction model with five new variables added to this empirical model was developed using Synthesis Analysis and literature information. This model was then compared with the four-variable empirical model using the first National Health and Nutrition Examination Survey (NHANES I) Epidemiologic Follow-up Study data set. The synthesized model had significantly improved predictive power ( $v_2 = 43.8$ ,  $P < 0.00001$ ).

Conclusions: Synthesis Analysis provides a new means of developing complex disease predictive models from the medical literature.