Appendix V

THE EXTENDED EJIGOU-McHUGH RELATIVE RISK ESTIMATOR

Let R_k, k=1,2, ..., K, denote the distinct numbers of comparisons matched to the cases and let n_k denote the number of matched sets with exactly R_k comparisons. A matched set is defined as the case and his matched comparisons. Let $n=n_1+n_2+ \ldots +n_K$ denote the total number of matched sets.

Define $Z_{k,i,T}$, k=1,2, ...,K, i=0,1, by

$Z_{k,0,T}$ = the number of matched sets, among those having exactly R_k comparisons, in which the case is alive and exactly T of the R_k comparisons have died, T=1,2, ..., R_k

$Z_{k,1,T}$ = the number of matched sets, among those having exactly R_k comparisons in which the case has died and exactly T of the R_k comparisons have died, T=0,1,2, ..., R_k-1.

The extended estimate, ψ, is given by

$$\psi = \frac{\sum_{k=1}^{K} \sum_{T=1}^{R_k} Z_{k,0,T} Z_{k,1,T-1} / (Z_{k,0,T} + Z_{k,1,T-1})}{\sum_{k=1}^{K} \sum_{T=1}^{R_k} T Z^2_{k,0,T} / (R_k - T + 1)(Z_{k,0,T} + Z_{k,1,T-1})}$$

and its variance is estimated by

$$\sigma^2 = \frac{\psi}{\sum_{k=1}^{K} \sum_{T=1}^{R_k} Z_{k,0,T} / [\psi + (R_k - T + 1)/T]}.$$

Since K is finite, this estimate has the same distributional properties as the Ejigou-McHugh estimator; it is asymptotically efficient and unbiased. The underlying assumptions used in its derivation are that the disease under study is of low incidence and that relative risk is constant over the levels of the matching variables. The Ejigou-McHugh estimate and the above extension are equivalent in asymptotic efficiency to maximum likelihood estimation (12).