p constraints have been applied If the maximum likelihood estimates of the set of parameters are β and β^* respectively, then -2 ln $[L(\beta^*)/L(\beta)]$ is asymptotically distributed as χ^2 on p degrees of freedom, where L(.) denoted the likelihood function When investigating the regression coefficients themselves, one can use the fact that they asymptotically have a multivariate normal distribution with a variance-covariance $[I(\beta)]^{-1}$ where

$$I(\beta) = \left[\frac{d^2L(\beta)}{d\beta_i d\beta_j}\right] m \times m$$

This enables one to obtain estimates of the standard deviation of the estimates of the regression parameters.

Table A.5 gives a list of the variables that were considered in the analysis of the UGDP data and Table A.6.1 summarizes the findings on them. The estimates of the regression coefficients when all of these variables have been included are given in Table A.6.2. As often happens when one does multiple regression with many parameters, there are redundancies, so that a relatively small subset of the variables gives nearly as good a prediction as the entire set. In looking for an appropriate subset of variables, one still includes the variables that are of greatest interest, in this case, the treatment effects. Sex was also included because of an interest in the treatment effects for each sex. It has already been noted that after adjusting for treatment variables, demographic variables, and time of follow-up, the clinic differences were not significant, and so the variables for clinics were dropped. Other demographic variables were added to the regression until the maximum of the likelihood did not differ significantly from the maximum of the likelihood when all the variables were included. The order in which the variables were entered into the regression depended on the absolute value of T (see Table A.6.2), the large values being entered first. The subset of variables thus identified (age, sex, systolic blood pressure, electrocardiographic abnormality, cholesterol level, and arterial calcification) is indicated in Table A.5 and these were used in the further analysis. A regression analysis with this subset of variables other than sex was also done separately for each sex. The results of the analysis using the subset of variables are summarized in Table 9.6.1, as well as in Table A.6.3. This analysis indicated that the treatment effects may be different from the two sexes. The harmful effect of tolbutamide treatment is most apparent for women, although the effect for men is not significantly different from that for women. It is not clear whether the results for tolbutamide apply only to women or whether the effect on women is more obvious because of the larger numbers involved.

Table A.6.4 gives the number of deaths observed in each treatment group, broken down by age and sex, along with the number expected on the basis of the model using the subset of variables just mentioned. It appears that the model does reasonably well in predicting the number of deaths in each group.

The variables used in the analysis of the Bedford study data are shown in Table A.9.1. The analysis was done for all causes and cardiovascular causes of death, and the results are summarized in Table A.9.2. The regression coefficients obtained when all the variables are included are shown in Table A.9.3.

APPENDIX B

Relative allocation method

In this section we outline the rudiments of the relative allocation (RA) method of analysis. Define

S_α = Total follow-up time for the α^{th} individual ($\alpha = 1,2,\ldots, n$);

$S_{i\alpha}$ = Total follow-up time for the α^{th} individual on the i^{th} treatment ($i = 0,1,2,3,4$);

$w_{i\alpha} = S_{i\alpha}/S_\alpha$ = Relative time on i^{th} treatment for α^{th} individual;

$\delta_\alpha = \begin{cases} 1 \text{ if } \alpha^{th} \text{ individual is dead (or if cardiovascular death)} \\ 0 \text{ otherwise.} \end{cases}$

We shall denote the treatments placebo, tolbutamide, standard-dose insulin, and variable-dose insulin by the subscripts $i = 1,2,3,4$, and $i = 0$ will refer to no treatment. Therefore, the α^{th} patient in the study supplies the vector of information $(S_\alpha, w_{i\alpha}, \delta_\alpha)i = 0,1,2,3,4$.

The relative allocation number of deaths for the i^{th} treatment is defined by

$$d_i' = \sum_{\alpha=1}^{n} w_{i\alpha}\delta_\alpha.$$