A MULTIPLE REGRESSION ANALYSIS OF ADMISSION OF INFANTS TO HOSPITAL

A REPORT FROM THE JERUSALEM PERINATAL STUDY

S. HARLAP, N. N. STENHOUSE AND A. M. DAVIES

Department of Medical Ecology, Hebrew University Hadassah Medical School, Jerusalem, Israel, and Medical Statistics Unit, University of Western Australia, Perth

In Jerusalem, the rate of admission of infants to hospital is probably the highest in the world. More than 18% of babies enter hospital at least once before their first birthday, and some are admitted several times, so that the rate of hospitalization episodes is over 26% of live births (Bentwich, 1970). The peak risk is at 3-6 months of age when 88% of episodes are for gastroenteritis or respiratory disease. Towards the end of the first year, however, and in the second and third years of life, congenital malformations are the most important primary cause of admission (unpublished data).

Admission to hospital is the most expensive form of medical care and although it may be instrumental in reducing infant mortality it is pertinent to ask whether such high rates are necessary. Earlier studies have shown that birth weight, birth order, parental education, and country of origin are the most important predictors of hospitalization in the first year of life (Bentwich, 1970). A multiple regression analysis (Harlap *et al.*, 1971) using the method of Feldstein (1966) confirmed that these are independent predictors and identified a number of other significant variables, notably sex, season of birth, area of residence, and delivery by Caesarean section.

The present study uses a more sophisticated multiple regression technique to search for interactions between variables such as sex, birth weight, and ethnic group, to define more accurately the effect of non-linear variables such as season of birth, birth weight and social class, and to explore further the relative importance of the age, education, and ethnic characteristics of both parents.

CHILD HEALTH SERVICES IN JERUSALEM

Preventive services are provided by a network of municipal mother-and-child stations which are widely used by all sections of the population

(Grushka, 1968). The average Jerusalem infant visits the centre 16 times in the first year of life, is seen by the doctor two to three times, and is visited at home by the nurse three times (Halevi and Handelsman, 1971). Curative services are provided by the sick funds, to which more than 90% of the population belong (Davies, 1972). Uninsured poor are provided for by the Ministry of Social Welfare in the framework of the sick funds. In addition, there is extensive use of private paediatricians by insured and uninsured alike.

Infants are admitted to hospital either by referral from their paediatrician or directly from the hospital emergency room; there is no distinction between private and other patients. Recent pilot studies have shown that many parents use the hospital emergency room as the first source of medical care. Other parents may consult several doctors for their child's illness.

MATERIALS AND METHODS

The data used are those of the Jerusalem Perinatal Study, whose design, methodology, and scope have been described in detail in previous publications (Davies, Tzur, Prywes, and Weiskopf, 1968; Davies, *et al.*, 1969). Since 1964 a file has been opened for every baby born to a resident of West Jerusalem (before 1967, the Israeli sector). The file is based on the birth notifications and contains additional information on the mother and delivery abstracted from delivery room records. A computerized record linkage technique continuously updates the file with details of death or of malformations reported from maternity wards, mother-and-child health stations, hospital paediatric departments, and necropsies.

For a defined period (1966-69) information on all admissions to paediatric wards was also recorded and linked to the file, and these data form the basis of

the present study. Each of the town's three paediatric wards was visited weekly by specially trained public health nurses who recorded the name, date of birth, identity number, and other demographic information on each child admitted. After discharge the three main diagnoses were recorded, together with details of malformations, operations, and other events taking place during the inpatient stay, as well as selected measurements of weight and haemoglobin.

There were 18,035 live births during 1966-68, of which 791 were rejected from the study. Among the rejects were 365 infants who were born in hospitals outside the Jerusalem area, 278 who died before the 28th day of life, 153 with birth weight unknown, and 15 whose mothers' ages were unknown.

STATISTICAL TECHNIQUES

The regression technique is a linear additive model, as described by Armitage (1971), using a process of matrix inversion to solve the normal equations. A Fortran program, MR40TG, developed by and in use in the Department of Medical Statistics of the University of Western Australia was adapted to cope with up to 64 variables and to be used on the Hebrew University CDC6400 computer where it is named MR64. The program uses a stepdown procedure, starting with a regression on all variables to be tested, successively rejecting the variable with the smallest *t*-value, and repeating the procedure. Options exist whereby only a summary of each step is obtained until each of the variables has a *t*-value exceeding a preset limit. Full analyses of variance are then given until all *t*-values exceed another, usually higher, limit at which stage the program proceeds to the next regression problem.

The 'best' regression (equation) may be taken as that which has the least residual (deviations) variance. This corresponds to a maximum value of the A statistic proposed by Fisher (1924). If V_t is the variance of the dependent variable and V_d the deviations mean square then

$$A = \frac{V_t - V_d}{V_t}$$

More usually, that regression is chosen in which all the regression coefficients have a *t*-value which is significant. When so many significance tests are conducted simultaneously caution must be exercised in the interpretation of the results.

The program also calculates sums, means, standard deviations, standard errors, maxima, and

minima for all variables as well as a simple correlation matrix. Transformations are available, and second or third degree polynomials of those variables whose effect on hospitalization was suspected of being curvilinear were fitted. Seasonal effects were examined by using the sine and cosine of the day of admission expressed in radian measure.

For the classificatory variables a set of dichotomous variables is created in the manner outlined by Feldstein (1966). If there are *k* classes then there will be *k*-1 variables created, each taking the value 1 or 0, depending on whether the individual has or has not the characteristic. Obviously the last (or dummy) class will be represented by *k*-1 zeros. The advantage of using multiple regression techniques is that measure variables, functions of measure variables, ranked variables, and classificatory variables, together with their interactions, may be tested as determining variables simultaneously even though statistical and functional relationships exist between them.

DESCRIPTION OF VARIABLES AND HYPOTHESES TO BE TESTED

The dependent variable in this study is the number of episodes of admission before the infant's first birthday.

The characteristics of the infant are birth weight, twinning, sex of birth, and congenital malformations. Birth weight is coded in kilogrammes, to the nearest 100g, and is expressed in deviations from its population mean of 3.27 kg. Since the effect of birth weight is expected to be curvilinear the expressions birth weight squared and birth weight cubed are also tested. Twinning and sex (maleness) are coded as dichotomies, being 1 if the characteristic is present, 0 if absent. Season of birth is tested as sine and cosine functions of time. Major and minor malformations are treated as dichotomies: the ICDA codes for their definition have been listed in full elsewhere (Harlap *et al.*, 1971).

Maternal and paternal ages are coded in years and expressed as deviations from their means. The 187 fathers with age unknown have been assigned an age of 4.3 years greater than that of the mother. Age squared is also tested for each parent.

'Ethnic' origin groups are coded in a series of dichotomies, depending on the birthplace of the parent's father, i.e., the baby's grandfather. There are five groups for each parent, involving four dichotomies of 1 or 0: Jews of Asian, North African, and Western origin, and Arabs. The fifth group,

parents who are themselves second generation Israelis, is the dummy variable, being coded 0 on each of the four dichotomies. A sixth variable, a dichotomy for immigration, defines whether the parent was born abroad (1) or in Israel (0).

There are five education groups for each parent, depending on the years of schooling. These are 0-4, 5-8, 9-12 (high school), and 13+ (university or other higher education), each of which is coded as a dichotomy of 1 or 0, while the unknowns form the fifth group. Because there are large numbers of unknowns, an attempt is made to construct a further variable from the father's occupation. The 100 occupation groups (classified by the Central Bureau of Statistics) are ranked according to the mean years of education of the fathers and treated as a continuous variable, called here 'social class'. There is no accepted ranking of occupations in use in Israel, such as the social class ranks of Britain.

The mother's occupation at the time of the birth is coded as a dichotomy, being 1 if she reports herself as a 'housewife' and 0 if she states that she has an occupation. Women who work usually do so until the onset of labour and commonly return to work three to six months after the birth, but can take up to 12 months' unpaid leave without losing work tenure. Three months' full salary is provided by national insurance.

Obstetric characteristics of the mother include parity, the number of previous still births or child deaths, and eight dichotomies for induction of labour, forceps or vacuum delivery, breech presentation, Caesarean section, suspected toxæmia, premature rupture of membranes, fetal distress or asphyxia and other obstetric complications.

Other variables are linked to standard of living and to the use of mother-and-child health centres. There are three dichotomies for the four hospitals of birth (all births take place in hospital) and five dichotomies for six areas of residence. Of these, the first three are classifications by the municipal tax level of the area of residence, the fourth is the town commercial centre, the fifth the rural hinterland of Jerusalem, and the sixth, the 'dummy', represents unknown and new areas as yet unclassified for tax purposes.

Interaction variables test the hypothesis that the effects of certain variables change according to the value of other variables. They are constructed from the products of the pairs of variables being tested. Possible interactions tested in this study are between social class and father's age, sex and birth weight, sex and ethnic group, and sex and minor malformations.

RESULTS

For the 17,244 infants in the study, Table I shows totals, means, standard deviations, maxima, and minima for each of the variables and the crude correlation coefficients between episodes of hospital admission and each of the possible predictors.

The best regression equation is shown in Table II. Significant predictors of hospital admission episodes were the presence of major and minor malformations; parity; linear and quadratic functions of birth weight, parental ages, and social class; a constellation of variables describing parental education, ethnic group, and area of residence; Caesarean section; and season of birth.

Male infants whose fathers or grandfathers were born in Asian or North African countries were more at risk for hospital admission than females. Among other ethnic groups there was no significant effect of sex. Another significant interaction was between social class and father's age, the effect of class increasing with the age of the father.

The 30 predictor variables in the regression equation together explain 13% of the variation in hospital admission episodes.

DISCUSSION

The investigation of a large number of variables and the interaction between them is extremely cumbersome, and a multiple regression approach is, at this time, the only possible one. Its value in this study is clearly demonstrated. The work involved in computing rates and cross-tabulation to define 30 of the variables as significant and reject the remainder would be an insurmountable task by any other method.

Birth weight, parity, and the presence of a major congenital malformation are the three most important predictors of admission to hospital in the first year of life. Unpublished portions of the Jerusalem Perinatal Study show, in addition, that infants with major malformations are more likely than others to be hospitalized with gastroenteritis and respiratory infections as well as being more likely to have surgical operations. The excess risk for babies born with minor malformations may merely reflect the increased chance of detecting them when infants are admitted to hospital.

Of major interest are the findings in this study in relation to the infant's sex. There is no significant sex difference in hospital admissions except within the groups of infants whose fathers were born in Asia or North Africa, where boys are valued more than girls. The data suggest that sex differences are due to parental concern for males rather than to a

TABLE I
BASIC DATA FOR VARIABLES

		Sum	Mean	Standard Deviation	Maximum	Minimum	Crude Correlation Co-efficient with Dependent Variable
Dependent variable	Variable						
	Hospital admission episodes	46,49·000	0·270	0·705	8·0	0·0	1·000
Characteristics of infant	Birth weight, deviation from mean (3·27 kg.)	-0·004	-0·000	0·500	2·4	-2·6	-0·087***
	Birth weight deviation squared	4,302·523	0·250	0·439	6·9	0·0	-0·082***
	Birth weight deviation cubed	-396·334	-0·023	0·741	14·4	-18·0	-0·075***
	Twins	382·000	0·022	0·147	1·0	0·0	0·038***
	Sex	8,796·000	0·510	0·500	1·0	0·0	0·052***
	Season 1 year cycle, sine	-161·684	-0·009	0·703	1·0	-1·0	0·016*
	cosine	-251·359	-0·015	0·711	1·0	-1·0	-0·018*
	Minor malformations	749·000	0·043	0·204	1·0	0·0	0·130***
Major malformations	595·000	0·035	0·183	1·0	0·0	0·235***	
Demographic characteristics of mother	Age, deviation from mean (27·9)	0·077	0·000	5·899	26·1	-13·9	0·024**
	Age deviation squared	600,118·000	34·802	46·628	683·7	0·0	0·046***
	Birth place of mother's father						
	Asia	5,773·000	0·335	0·472	1·0	0·0	0·087***
	North Africa	4,173·000	0·242	0·428	1·0	0·0	0·067***
	West	4,702·000	0·273	0·445	1·0	0·0	-0·136***
	Arabic	419·000	0·024	0·154	1·0	0·0	0·072***
	Immigrant (mother born abroad)	10,474·000	0·607	0·488	1·0	0·0	0·084***
	Years of education 0-4	2,792·000	0·162	0·368	1·0	0·0	0·131***
	5-8	4,972·000	0·288	0·453	1·0	0·0	0·038***
	9-12	4,774·000	0·277	0·447	1·0	0·0	-0·080***
	13+	3,014·000	0·175	0·380	1·0	0·0	-0·117***
	Housewife	11,381·000	0·660	0·474	1·0	0·0	0·095***
Demographic characteristics of father	Age, deviation from mean (32·1)	-702·329	0·041	7·031	43·9	-16·1	0·029***
	Age deviation squared	852,378·000	49·430	93·618	1,923·8	0·0	0·023**
	Birthplace of father's father						
	Asia	5,717·000	0·332	0·471	1·0	0·0	0·080***
	North Africa	3,866·000	0·224	0·417	1·0	0·0	0·071***
	West	4,794·000	0·278	0·448	1·0	0·0	-0·133***
	Arabic	410·000	0·024	0·152	1·0	0·0	0·074***
	Immigrant (father born abroad)	10,517·000	0·610	0·488	1·0	0·0	0·085***
	Years of education 0-4	1,762·000	0·102	0·303	1·0	0·0	0·116***
	5-8	4,442·000	0·258	0·437	1·0	0·0	0·057***
	9-12	6,324·000	0·367	0·482	1·0	0·0	-0·074***
	13+	3,088·000	0·179	0·383	1·0	0·0	-0·118***
	Social class deviation from mean (10·3)	0·216	0·000	3·187	8·0	-6·5	-0·147***
Social class deviation squared	175,129·700	10·156	11·736	64·5	0·0	-0·062***	
Obstetric characteristics of mother	Parity	40,287·000	2·336	2·626	20·0	0·0	0·134***
	Previous still birth or child death	1,672·000	0·097	0·408	6·0	0·0	0·079***
	Induction of labour at birth	1,643·000	0·095	0·294	1·0	0·0	-0·023**
	Forceps or vacuum delivery	936·000	0·054	0·227	1·0	0·0	-0·027***
	Breech presentation	545·000	0·032	0·175	1·0	0·0	0·025**
	Caesarean section	674·000	0·039	0·194	1·0	0·0	0·016*
	Suspected toxemia during pregnancy	229·000	0·013	0·114	1·0	0·0	0·021**
	Premature rupture of membranes	382·000	0·022	0·147	1·0	0·0	0·007ns
	Fetal distress and/or asphyxia	439·000	0·054	0·227	1·0	0·0	-0·005ns
	Persistent occipitoposterior presentation	430·000	0·025	0·156	1·0	0·0	-0·003ns
	Other obstetric complications	2,901·000	0·168	0·374	1·0	0·0	0·007ns
Other variables	Hospital of birth number 2	5,480·000	0·318	0·466	1·0	0·0	0·069***
	3	4,706·000	0·273	0·445	1·0	0·0	0·009ns
	4						-0·036***
	Municipal tax						
	High	4,316·000	0·250	0·433	1·0	0·0	-0·090***
	Medium	5,570·000	0·323	0·468	1·0	0·0	-0·009ns
	Low	3,961·000	0·230	0·421	1·0	0·0	0·037***
Town centre	465·000	0·027	0·162	1·0	0·0	0·005ns	
Rural hinterland	2,647·000	0·154	0·360	1·0	0·0	0·083***	
Interactions	Father's age × social class	-9,404·100	-0·545	21·695	239·9	-168·3	-0·012ns
	Birth weight × sex	533·511	0·031	0·362	2·1	-2·6	-0·063***
	Sex × Asian father	2,862·000	0·166	0·372	1·0	0·0	0·081***
	Sex × North African father	2,005·000	0·116	0·321	1·0	0·0	0·080***
	Sex × Western father	2,490·000	0·144	0·352	1·0	0·0	-0·076***
	Sex × Arabic father	221·000	0·013	0·112	1·0	0·0	0·060***
	Sex × minor malformations	548·000	0·032	0·175	1·0	0·0	0·113***

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$

TABLE II

MULTIPLE REGRESSION EQUATION FOR NUMBER OF HOSPITAL ADMISSION EPISODES IN FIRST YEAR OF LIFE

Variable				Regression Coefficient	t Value	Partial Correlation Coefficient
Characteristics of infant	Birth weight deviation	..	..	-0.1132	-10.920	-0.083
	Birth weight deviation squared ..	..	..	0.0748	6.447	0.049
	Season of birth (cosine)	..	..	-0.0154	-2.142	-0.016
	Minor malformations	..	..	0.3800	15.363	0.116
	Major malformations	..	..	0.8166	29.571	0.220
Demographic characteristics of mother	Age deviation	..	..	-0.0055	-3.621	-0.028
	Asian origin	..	..	0.0734	4.964	0.038
	North African origin	..	..	0.0714	4.247	0.032
	Education 9-12	..	..	-0.0423	-2.956	-0.025
	13+	..	..	-0.0458	-2.294	-0.017
Demographic characteristics of father	Age deviation	..	..	-0.0372	-3.181	0.024
	Arabic	..	..	0.2764	7.422	0.056
	Education 0-4	..	..	0.0510	2.111	0.016
	5-8	..	..	0.0442	-2.181	-0.017
	9-12	..	..	-0.0768	-3.331	-0.025
	13+	..	..	-0.0714	-2.888	-0.022
	Social class	..	..	-0.0106	-4.565	-0.035
	Social class squared	..	..	0.0017	2.965	0.023
Obstetric characteristics of mother	Parity	..	..	0.0320	10.148	0.077
	Previous stillbirth/death	..	..	0.0282	2.085	0.016
	Caesarean section	..	..	0.0767	2.929	0.022
Other variables	Hospital 2	..	..	0.0265	2.334	0.018
	Municipal tax	..	..			
	High	..	..	0.0868	2.149	0.016
	Medium	..	..	0.1049	2.620	0.020
	Low	..	..	0.1197	2.954	0.023
	Town centre	..	..	0.1346	2.711	0.021
	Rural area	..	..	0.0861	2.063	0.016
Interactions	Sex × Asian father	..	..	0.0858	5.672	0.043
	Sex × African father	..	..	0.1114	6.049	0.046
	Social class × father's age	..	..	0.0537	2.255	0.017

Variation due to:—		Degrees of Freedom	Sum of Squares	Mean Square	Variance Ratio
Regression	..	30	1,127.644	37.5881	86.8698
Deviations	..	17,213	7,447.981	0.4327	
Total	..	17,243	8,575.625		

R = 0.3626 A = 0.1300 Intercept on Y-axis = 1.0271

biological difference between sexes. It is of interest that, in the same cohort, the post-neonatal death rate is higher among females, further suggesting that females are less highly valued (unpublished studies).

The results are disappointing in that the proportions of variation in hospital admissions explained by the regression equation are small. There are clearly many unidentified factors at play; they probably involve the quality of extramural medical care as well as unmeasured factors such as crowding, sanitation, ability of the parents to cope, and other social and cultural variables.

This study suggests that the majority of hospital admissions are not at present easily preventable. However, studies are continuing and the development of the multiple regression technique has facilitated the simultaneous testing of a number of hypotheses. Hopefully, we will be able in the near future to define other variables involved in the

unexplained variation in hospital admissions, and thus open up possibilities for the design of logical programs of prevention.

SUMMARY

A stepwise multiple regression technique has been used to select the combination of variables which together give the best estimate of the number of hospital admissions in the first year of life for the 17,244 infants born in West Jerusalem in 1966-68 who survived the first 28 days of life. From the 60 demographic and health predictors tested, 30 variables together explained 13% of the variation in hospital admissions. Significant predictors were the presence of major and minor congenital malformations, parity, linear and quadratic functions of birth weight, social class, and parental ages, Caesarean section, season of birth, parental education, ethnic group, and area of residence. There

was no significant effect of the mother's going out to work, nor of breech delivery, fetal distress or asphyxia, and a number of other obstetric complications.

Admissions decreased with increases in birth weight, social class, parental ages, and education. There was a significant interaction between social class and paternal age, the effect of class increasing with increasing paternal age. There were more hospital admissions with increasing parity, delivery by Caesarean section, the presence of malformations and among infants of Arab fathers and Jewish mothers of Asian and North African origin. There was also a significant interaction between sex and ethnic group, males being more at risk if the father or paternal grandfather had immigrated from Asian or North African countries. The data support the hypothesis that sex differences in hospital admissions of Jerusalem infants arise from a status difference rather than from a biological difference between males and females.

We should like to thank the staff and management of Jerusalem's paediatric and obstetric departments, without whose cooperation this study would not have been possible. We also gratefully acknowledge the help given by Mrs. R. Prywes, Mrs. H. Rossman, Mrs. N. Samueloff, and Mr. M. Haber.

REFERENCES

- ARMITAGE, P. (1971). *Statistical Methods in Medical Research*. Blackwell Scientific Publications, Oxford.
- BENTWICH, T. (1970). The background of infant hospitalization in the Jerusalem area with special reference to the social and medical factors involved. (in Hebrew) M.D. thesis. The Hebrew University of Jerusalem.
- DAVIES, A. M. (1972). Preventive medicine in Israel. *Prev. Med.*, **1**, 121.
- , PRYWES, R., TZUR, B., WEISKOPF, P., and STERK, V. V. (1969). The Jerusalem Perinatal Study I. Design and organization of a continuing community-based, record-linked survey. *Israel J. med. Sci.*, **5**, 1095.
- , TZUR, B., PRYWES, R., and WEISKOPF, P. (1968). *The Jerusalem Perinatal Study, Methodology Manual*, Department of Medical Ecology, Hebrew University of Jerusalem.
- FELDSTEIN, M. S. (1966). A binary variable multiple regression method of analysing factors affecting perinatal mortality and other outcomes of pregnancy. *J. roy. statist. Soc.*, **129**, series A, 61.
- FISHER, R. A. (1924). The influence of rainfall on the yield of wheat at Rothamsted. *Phil. Trans. B.*, **213**, 89.
- GRUSHKA, T. (1968). *Health Services in Israel*. Ministry of Health, Jerusalem.
- HALEVI, H. S., and HANDELSMAN, M. (1971). Mother and child health stations (1970). *Monthly Bulletin of Statistics*, **22**, Suppl. to No. 11, p. 108 (in English), p. 43 (in Hebrew). Central Bureau of Statistics, Jerusalem.
- HARLAP, S., DAVIES, A. M., HABER, M., SAMUELOFF, N., ROSSMAN, H., and PRYWES, R. (1971). Ethnic group, immigration and infant morbidity in West Jerusalem. Proc. 2nd int. symp. *Society Stress and Disease—Childhood and Adolescence*. Oxford University Press (to be published).