Upsala J Med Sci 80: 135–140, 1975

Prediction of Birth Weight

Results of a Multiple Regression Analysis

FOLKE PETTERSSON and STIG MELANDER

From the Department of Social Medicine (Head: Ragnar Berfenstam) and the Department of Gynaecology and Obstetrics (Head: Carl Gemzell), University Hospital, Uppsala, Sweden

ABSTRACT

The present study is an attempt to isolate by a multiple regression analysis some factors probably influencing the birth weight and to utilize the results of the analysis for statistical prediction of birth weight at different gestational ages for various combinations of factors.

INTRODUCTION

From a scientific as well as from a practical point of view it is of interest to obtain figures on birth weight of children born at various stages of gestation and to find out what factors influence birth weight. Such data can give us a picture of the fetal growth curve during pregnancy and make possible a definition of the limits for classification purposes, e.g. classification of a baby as 'small for date'.

The figures reported on mean birth weight vary from country to country. Within several countries secular trends in birth weight have been registered in a number of studies during recent decades (Abolins, 1961; Ashford et al., 1969).

Factors responsible for differences and changes are not easily detected. Besides the fact that increases in birth weight run parallel to more favourable living conditions and probably improved nutritional conditions, some single factors influencing birth weight have been isolated. Male newborns, for example, have a higher average birth weight than female newborns (Karn & Penrose, 1954). Length of pregnancy has a significant influence on the birth weight, as also have smoking habits—babies of smokers weigh less than babies of non-smokers (Lowe, 1959; Järvinen & Österlund, 1963; Zabriskie, 1963; Mulcahy, 1968; Pettersson, 1969). Birth order exerts an influence on birth weight, the weight increasing with increasing values of birth order.

Mean birth weight is generally found to increase with increasing maternal age (Kontsek, 1940; Karn, 1954; Selvin & Janerich, 1971). Some authors (McKeown & Gibson, 1951; Fraccaro, 1956; Abolins, 1961) did not find any relationship between birth weight and age of the mother. Others (Karn & Penrose, 1951) found a slight decrease in birth weight with increasing maternal age. Pettersson (1970) found that babies of primiparae 40 years of age or older had a significantly lower birth weight than babies of primiparae of other ages. Selvin & Janerich (1971) showed that especially in primiparous women, birth weight decreased in inverse proportion to increasing age of the mother.

There are significant differences between birth weight of infants from different ethnic groups (Hytten & Leith, 1964). Reports have been published on a positive relationship between the mother's height and the baby's birth weight (Cawley et al., 1954), on a positive relationship between the mother's weight and the birth weight of her baby (McKeown & Record, 1957), on a positive correlation between maternal diabetes and birth weight (Miller et al., 1954) and on a negative relationship between non-toxemic hypertension in pregnancy and birth weight (Johnson, 1958). Within sibship, birth weights show a strong positive correlation (Tanner et al., 1972).

Present study

The present study is an attempt to isolate by a multiple regression analysis some factors probably having an influence on birth weight and to utilize the results of the analysis for statistical prediction of birth weight at different gestational ages for various combinations of factors.

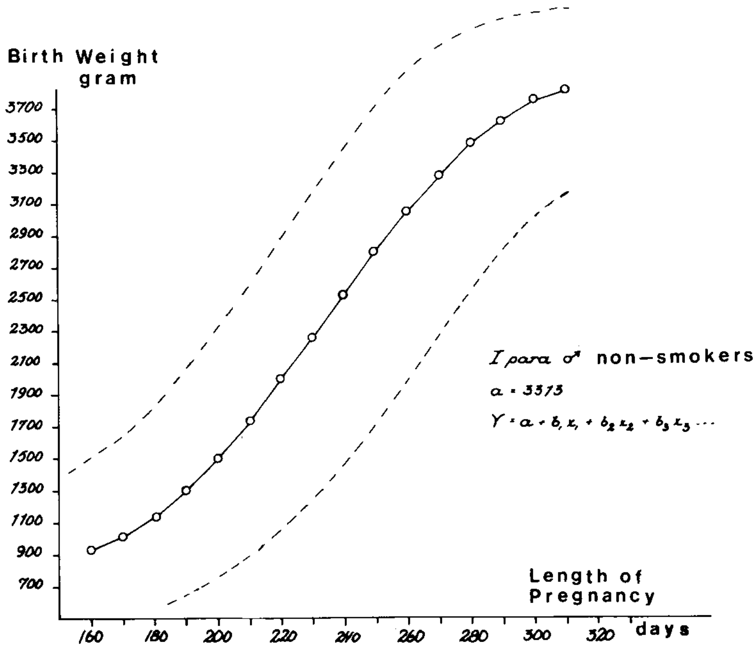

Fig. 1. Predicted values of birth weight for babies born to non-smoking primiparous women by length of pregnancy.

MATERIAL AND METHODS

The primary material consisted of a consecutive series of deliveries from January 1963 to April 1964 at the University Hospital of Uppsala, Sweden. Altogether 3 782 deliveries were examined. Due to incompletemenss in registered data the material was reduced to a final size of 3 307 single births.

Table I.

Dependent variable is birth weight

The studied independent variables are:

1. Pregnancy order—its mean value
2. (Pregnancy order—its mean value)^2
3. Age—its mean value
4. (Age—its mean value)^2
5. Socio-economic group I=1, others=0
6. Socio-economic groups I or II=1, others=0
7. Married=1, unmarried=0
8. Smoking habits I=1, others=0 (I=non-smokers)
9. Smoking habits I or II=1, others=0 (I=non-smokers, II=women with less than 10 cigarettes)
10. Alcoholic group I=1, others=0 (I=abstainers)
11. Women with previous legal abortions=1, others=0
12. Women with previous spontaneous abortions=1, others=0
12. Sex of the child, male=1, female=0
14. Gestational age in days—its mean value
15. (Gestational age in days—its mean value)^2
16. Women from Uppsala town=1, others=0
17. (Gestational age in days—its mean value)^3

The following factors assumed to exert an influence on birth weight were studied (see Table I). Birth weight was used as dependent variable (Y) and, as independent variables, pregnancy order, maternal age, sex of the baby, social class, civil status, smoking habits, drinking habits, previous history of legal spontaneous abortion, length of pregnancy, place of domicile (urban or rural). In order to account for curvilinear regressions, pregnancy order, age of the mother and length of pregnancy were also taken in square and length of pregnancy moreover in cubic.

The variables were coded and transferred from the ordinary delivery records to IBM cards. The cards were thereafter analysed by the standard computer program BMDO3R.

RESULTS

The analysis gave a coefficient of multiple determination (R^2-value) of 0.31. Table II shows the regression coefficients obtained. * ** *** denote significance at confidence levels 0.05; 0.01 and 0.001 respectively.

As can be seen from Table II the sex of the child, the pregnancy order, a history of reproductive failure, and smoking during pregnancy significantly influenced birth weight. The male child weighed, on average, 152 g more than the female child. A history of previous spontaneous abortion was associated with a lower birth weight than the birth weight without a history of previous spontaneous abortion. The same trend was found with regard to a history

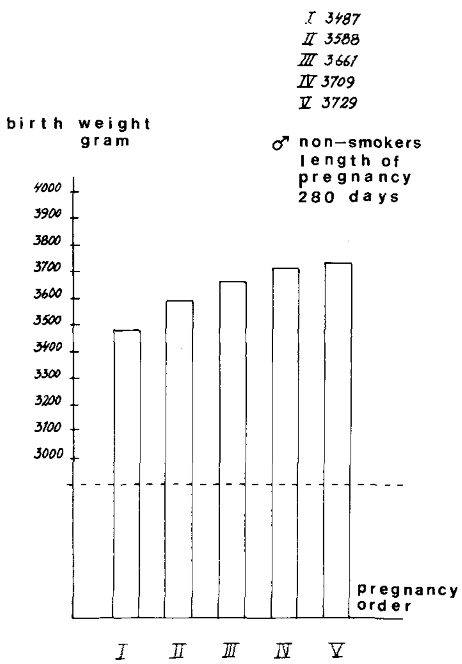

Fig. 2. Predicted values of birth weight for babies born to non-smokers on the 280th day of pregnancy by pregnancy order.

of legal abortion, although the regression was not statistically significant. Babies of smoking mothers weighed, on average, 110 g less than babies of non-smoking mothers.

Age of the mother, social class grouping, place of domicile, civil status, and drinking habits, were all found to exert no significant influence upon the birth weight.

Fig. 1 shows the regression of the length of pregnancy on the birth weights. The curve is a constructed one obtained in the following way.

The multiple regression equation or, according to Croxton's terminology (Croxton, 1959), the estimating equation $Y=a+b_1x_1+b_2x_2+b_3x_3+ \ldots +b_nx_n$ makes it possible to calculate the expected birth weight for each individual combination of variables. In this calculation the regression coefficients obtained (b_1, b_2 . . . see Table II) are used as predictors and x_1, x_2 . . . are the variables, each with a certain numerical value for each individual case. Thus for an individual primipara on her 280th day of pregnancy when delivered, non-smoker, living in the town of Uppsala, giving birth to a boy, and so on, it is possible to calculate the expected weight of her baby according to the estimating equation, because we know the regression coefficients (see Table II); a is here 3313 and we can also obtain the various x_n-values for this particular patient.

In this way the expected individual birth weights were calculated for all primiparae in our material. Fig. 1 shows the arithmetic mean values of all those calculated birth weights for various lengths of pregnancy. It can be noted that the regression curve is a sigmoid one.

The predicted values of birth weight for babies born on the 280th day of pregnancy in women who do not smoke is shown in Fig. 2. The figure shows the predicted values of birth weight for birth order I–V.

The construction of correct statistical limits around the regression curve shown in Fig. 1 involves, however, a complex statisticaly theory with heavy computations. We have therefore chosen a simplified method for fitting confidence intervals which means a loss of statistical precision and stringency but which, in our opinion, does not mean a serious misrepresentation of the data. Thus, while our calculated limits and ranges shown should be taken as approximate on theoretical grounds, we nevertheless consider that they are quite good estimates of the true confidence intervals. For a description and detailed discussion of the method which we have adopted, see Croxton (1959) and Hyrenius (1962).

Table II. *Multiple regression analysis*

Variable no.	Reg. coeff.	S.E. of reg. coeff.	Computed *T*-value
1	83.676	11.108	7.5***
2	−13.251	2.841	4.7**
3	0.531	2.127	0.2
4	0.088	0.225	0.4
5	22.014	24.196	0.9
6	25.915	19.057	1.4
7	45.218	28.948	1.6
8	107.836	18.713	5.7***
9	7.563	33.762	0.2
10	16.238	17.188	0.9
11	−45.601	67.842	0.7
12	−61.580	27.440	2.2*
13	151.848	16.302	9.3***
14	17.913	0.742	24.1***
15	− 0.194	0.018	10.7***
16	0.437	17.586	0.0
17	−0.001	0.000	5.9***

Sample size 3 307. Intercept (a value) 3 313. S.E. of the estimate 466.6.

The method utilized is based on the formula

$$S_{1.1234\ldots n}=\sqrt{S_1^2\,(1-R_{1.234\ldots n}^2)}$$

or with simpler symbols

$$S_{Y.X}=\sqrt{S_1^2\,(1-R^2)}$$

where $S_{1.234\ldots n}=S_{Y.X}$ is the standard error of prediction, S_1^2 is the variance of the dependent variable (here the birth weight) and $R_{1.234\ldots n}^2=R^2$ is the coefficient of multiple determination. We can expect that, *on the average*, about 68% of the predicted birth weights will fall within $\pm 1.S_{Y.X}$ (vertically) of the predicting curve, about 95% within $\pm 2.S_{Y.X}$ and 99.7% within $\pm 3.S_{Y.X}$. Thus $S_{Y.X}$ is a measure of the general reliability of prediction rather than a specific measure applicable to a particular prediction.

In the situation of Fig. 2 we have $S_1^2 = 560560 = 313600$ and $R^2=0.31$. Then

$$S_{Y.X}=\sqrt{313600\,(1-0.31)}$$
$$S_{Y.X}\approx 465$$

The value of $S_{Y.X}\approx 465$ means that if the predicted birth weight according to Fig. 1 is 3 400 g at 280 days of pregnancy, this predicted birth weight falls, on average, within 3 400±930 g in 95% of all predicted values, in other words within 2 470–4 330 g.

In the manner now described, limits around the predicting curve were constructed.

DISCUSSION

In routine clinical practice the comparison of the weight of a newborn fetus or child with a standard reference is one of the methods used for judgement of its degree of development. It follows that this method requires availability of standard references.

Several studies have tried to present standard reference materials. In Sweden, Engström & Sterky (1966) studied the birth weight of a primary material of 92 348 children born at various gestational lengths during 1956-07-01 – 1957-06-30, which means the majority of all children or liveborn fetuses in Sweden during this period (110 000). From the material were excluded all cases of stillbirth, malformation, multiple pregnancy, maternal diabetes, and toxemia of pregnancy. Furthermore, only mothers with regular menstrual intervals (21–35 days) were accepted for further evaluation. In this way the material was reduced to 58 984 fetus or children of which 54.5% were boys and 45.6% girls. This reduced material showed a mean pregnancy length of 281.2 days (boys) and 281.2 days (girls) with a range of gestational age from about 33 weeks to 45 weeks. The boys had a full-term birth weight of 3 596±542 g and the girls 3 408±503 g. Based on this material, regression lines were constructed for the relationship between the length of pregnancy and the corresponding weight of males and females respectively. The regression lines have since been used as nomograms assumed to show the normal intra-uterine rate of growth of Swedish fetus or children.

Timonen and collaborators (1969) studied hospital deliveries in Finland during 1957-07-01 – 1958-06-30, altogether 57 089 deliveries and moreover deliveries at the University of Helsinki, the Central Hospital II, during 1951–60, altogether 27 522 deliveries. Twinbirths, deliveries of diabetic mothers, stillborn babies and babies dead during parturition and malformed children were excluded from the material. On the basis of this material, nomograms were made relating weight of the baby, length of pregnancy, sex of the baby, and parity of the mother, to each other.

In the USA, Hendricks (1964) studied patterns of fetal and placental growth during the second half of normal pregnancy. The study was based on a material of 11 000 births at the University Hospitals of Cleveland, 1956–62. From the material were excluded, e.g., twin pregnancies, stillbirths, preeclamptics, and diabetics. Nomograms of fetal growth by weeks of gestation were constructed and also nomograms of mean daily fetal growth. Relations of sex of the child to fetal weight were also studied and likewise of parity to fetal weight.

It is characteristic of these three studies that they consist of very large materials and that the primary materials have been reduced by omitting abnormal pregnancies. The motivation for omitting complicated pregnancies has been to get a picture of the *normal* growth curve of the fetus, a motivation which has been explicitly expressed by Hendricks (1964) in a sentence.

No doubt huge materials give a lot of information, but much is involved in collecting them. The omission of complicated pregnancies may be a questionable principle. It can be argued, for example, that one must not compare the growth curve of a fetus of a diabetic mother with a superimposed hypertension of pregnancy, with the growth curve of a fetus of a mother with a pregnancy considered to be normal. In such a situation the comparison must be made with a diabetic pregnancy not superimposed by hypertension, especially if one wants to evaluate, for example, the effect of treatment of the hypertension. Moreover, the selection of normal pregnancies requires criteria of normality and such criteria can vary from place to place and from time to time. It is therefore doubtful if one can compare without reservation various assumed normal growth curves with each other.

One alternative to selected, reduced materials is unselected, total materials, an alternative which has been used in this study. In order to obtain information on factors and combinations of factors that may influence the growth of the fetus when working with unselected, total materials, several procedures may be undertaken. One method is a far-reaching division of the material into subgroups. A new problem will then arise, as the subgroups will tend to become too small to permit reliable statistical conclusions. Methods also exist for multi-factorial analysis, e.g. various standardizing procedures, but these methods will also in several instances prove impractical when many factors are involved.

In this study we have chosen a multiple regression analysis. It seems that with this method, small materials give the same amount of information and of the same reliability as large materials. It also seems evident to us that the powerful statistical method of multiple regression analysis has given, in spite of the limited size of the material, substantial information on the factors probably influencing the fetal growth curve during pregnancy without any need to omit certain cases.

The coefficient of determination (0.31) is rather low and indicates that we have not disclosed all variables of importance—for example, we have not been able to take into account data on the mother's height, her body weight, and increase in weight during pregnancy. These factors are known to influence the fetal weight. The ethnical composition of our material is rather homogeneous which means that racial differences probably have been of little importance. A genetic variation influencing the weight must be present in our material but we have not been able to disclose this factor in the present study. The uncertainty of menstrual data, with consequent uncertainty in the calculation of the expected date of confinement, also exists in our material of course and has probably given a larger variation around the variable used.

It is of interest to note that we could not show that sociomedical factors such as social grouping and civil status influenced the fetal weight. Such an influence has been noted by others with a lower fetal weight for children of single mothers and lower social classes. This problem has been discussed earlier (Pettersson, 1970). One explanation is that, *de facto*, no differences have existed in Sweden for many years regarding living conditions, nutritional conditions and antenatal maternal care between various groups of Swedish citizens.

REFERENCES

1. Abolins, J. A.: Acta Obstet Gynecol Scand *40:* 339, 1961.
2. Ashford, J. R., Fryer, J. G. & Brimblecombe, F. S. W.: Brit J Prev Soc Med *23:* 154, 1969.
3. Barron, S. L. & Vessey, M. P.: Brit J Prev Soc Med *20:* 127, 1966.
4. Cawley, R. H., McKeown, T. & Record, R. G.: Am J Hum Genet *6:* 448, 1954.
5. Croxton, F. E.: Elementary Statistics with Applications in Medicine and the Biological Sciences, Dover Publications, New York, 1959.
6. Engström, L. & Sterky, G.: Läkartidningen *63:* 4922, 1966.
7. Fraccaro, M.: Ann Hum Genet *20:* 282, 1956.
8. Hendricks, C. H.: Obstet Gynec *24:* 357, 1964.
9. Hyrenius, H.: Statistiska Metoder, 4th ed. Almqvist & Wiksell, Stockholm, 1969.
10. Hytten, F. E. & Leitch, I.: The Physiology of Human Pregnancy, 1st ed. Blackwell, Oxford, 1964.
11. Johnson, G.T.: *In* Symposium on Non-toxaemic Hypertension in Pregnancy (ed. Morris & McBrowne), p. 66. Churchill, London, 1958.
12. Järvinen, P. A. & Österlund, K.: Ann Paediat Fenniae 9: 18, 1963.
13. Kaminski, M., Goujard, J. E. & Rumeau-Rouquette, C.: Int J Epidemiol *2:* 195, 1973.
14. Karn, M. N.: Ann Eugen (London) *12:* 14, 1947.
15. Karn, M. N. & Penrose, L. S.: Ann Eugen (London) *16:* 147, 1951.
16. Lowe, C. R.: Brit Med J *2:* 673, 1959.
17. McKeown, T. & Gibson, J. R.: Brit J Soc Med *5:* 177, 1951.
18. McKeown, T. & Record, R. G.: Am J Hum Genet *6:* 457, 1954.
19. Miller, H. C., Hurwitz, D. & Kuder, K.: J Am Med Ass *124:* 271, 1954.
20. Mulcahy, R.: Am J Obstet Gynec *10:* 844, 1968.
21. Pettersson, F.: Acta Sociomed Scand *1:* 13, 1969.
22. Pettersson, F.: *In* Obstetrik & Gynekologi (ed. C. Gemzell) vol. 6, p. 263. Almqvist & Wiksell, Stockholm, 1970.
23. Pettersson, F.: Med Gyn Soc 5 (1): 10, 1970.
24. Selvin, S. & Janerich, D. T.: Brit J Prev Soc Med *25:* 1971.

25. Tanner, J. M., Lejarraga, H. & Turner, G.: The Lancet, pp. 193–197, July 1972.
26. Timonen, S., Uotila, U., Vara, P., Kunsiste, P. & Lokki, O.: Det rättsmedicinska bedömandet av konceptionstiden. Fostrets tillväxt under graviditeten. Medicinalstyrelsens undersökningar i Finland. Helsingfors, 1969.
27. Zabriskie, J. R.: Obstet Gynec *21:* 405, 1963.

Received April 10, 1975

Address for reprints:

Folke Pettersson, M.D.
Dept. of Gynaecology
Radiumhemmet
Karolinska sjukhuset
S-104 01 Stockholm 60
Sweden