J. Perinat. Med.
2 (1974) 37

Multifactorial study of the risk of prematurity at 32 weeks of gestation

II. A comparison between an empirical prediction and a discriminant analysis

Monique Kaminski, Emile Papiernik

Received January 8, 1973. Accepted October 18, 1973.

We have proposed an empirical method for estimating the risk of premature delivery at 32 weeks of gestation [5, 6]. The present study is a statistical validation of this method, having compared the results to those of a discriminant analysis.

1. Material and method

We chose **35 maternal characteristics which were known to favor a premature delivery or the birth of an infant weighing less than 2500 g.** Some of these characteristics are often cited in the literature, others are of our own selection [7].

These characteristics are grouped in series of comparable signs (social data, weight, height . . ., unfavorable obstetrical or gynecological antecedents, factors of fatigue, signs of danger at examination, signs of threat of imminent delivery). We assigned to each of the characteristics **a number of points varying from 1 to 5 according to their assumed importance.** The sum of the points corresponding to the characteristics present in a pregnant woman provides the **Empirical Coefficient of Risk of Premature delivery: E. C. R. P.** [1] at the time of the examination.

The study was made on a **risk group** (153 mothers of all babies weighing less than 2500g and born at "Maternité de Port-Royal" in 1969: Group R) and a **control group** (222 cases of mothers of infants weighing more than 2500 g born during the same year at the same hospital and chosen at random: Group C). We wanted to assess the risk by a single examination, and we chose as an optimum date 32 weeks of gestation from the first day of the last menstruation.

The information was taken from the routine prenatal care records filled in during the examination and transcribed after the delivery onto our special forms. Our study deals with the statistical results of the prediction of risk by the empirical method and then attempts an improvement using a discriminant analysis.

We decided to register **the signs occurring at the examination closest to the 32nd week.** As many women were sent to "Maternité de Port-Royal" from the whole Paris area when they showed signs of risk of premature delivery, the time of the examinations has a high variability **around the 32nd week,** and signs of imminent delivery were abnormally frequent.

In the present study, we have excluded mothers whose recorded examination was followed by delivery less than two days later. We thought that the information concerning these women who consulted late, might not have been recorded as carefully as for the others. These women differ only in their lower social class and a higher value for the danger of imminent delivery.

Finally, the calculations for discriminating between the two Groups R and C defined above were made on 107 mothers of low birth-weight infants (Group R) and 207 mothers of infants weighing more than 2500 g (Group C).

The function "empirical coefficient of risk (E C R P)" is a linear combination of the 35 registered variables, the coefficients of which were determined empirically according to what was known about their relative importance.

Discriminant analysis is a multivariate method of analysis (i. e. considering all the data together), which, theoretically, leads to the best linear combination separating the two groups of the sample, i. e. leads to the choice of the best coefficients. This is true if the variables are normally distributed, and if the variance and covariance matrices are equal in the two groups. In this study, our variables are qualitative and the variances and covariances are not equal. Hence the method will not give the best function, but it generally gives good results, as several authors have

Curriculum vitae

MONIQUE KAMINSKI *Ph. D. mathematics and statistics, is attachée de Recherches at the French Institute for Health and Medical Research at the Unit of Medical Statistics in Villejuif (Pr. D.* SCHWARTZ*).*

pointed out [8, 1] especially when the number of variables is great enough for the discriminant function to be distributed normally.

2. Results

2.1 Empirical coefficient of risk of prematurity

a) **The empirical coefficient of risk** is different in the Groups R and C:

mean of the ECRP in Group C: 6.2
mean of the ECRP in Group R: 12.2.
This difference is significant ($p < 0.01$).

b) **The distributions of the empirical risk coefficient** have been studied for Groups R and C (Fig. 1).
The distributions are different; in Group R we tend to find more of the higher ECRP values, but there is a large overlap between the two Groups; this shows that it is not possible to separate the two Groups completely using this function.

c) **We divided the Groups R and C** into three decision classes by choosing two limits for the ECRP values:

Class 1: no risk
Class 2: no decision
Class 3: high risk.

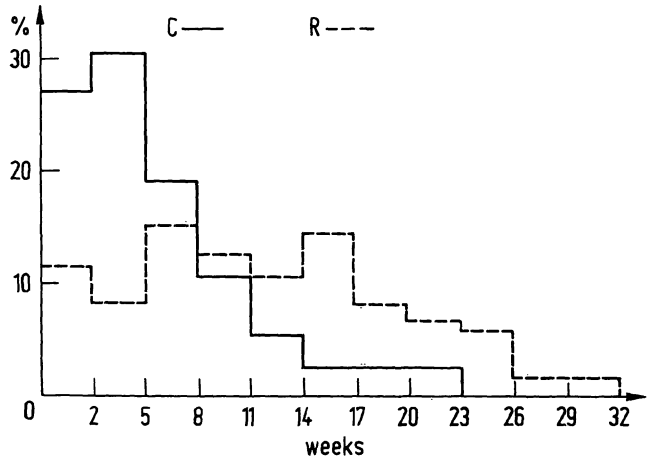

Fig. 1. Distribution of the E C R P in Groups C (—) and R (- - -).

This classification (Tab. I) is interesting. In the risk Group R, 54% are correctly classified as being at high risk, 19% fall into a moderate risk class and 27% are wrongly classified as low risk. On the other hand, in Group C there are 17% "false alarms", 19% in the "no decision" area and 64% are correctly classified as "low risk".

Tab. I. Distribution of Groups C and R in the ECRP risk classes.

ECRP	C	R
0—6	64%	27%
7—10	19%	19%
11 and more	17%	54%

2.2 Discriminant function

2.2.1 First analysis

A first discriminant analysis was made with all the 35 characteristics composing the ECRP, and resulted in a function, D 1, the coefficients of which (Tab. II) have the following properties:

a) Four coefficients are negative. They would be **"protective" factors or neutral ones, as their absolute value is very low.** These are unwed mother, unfavorable age, cylindrical uterus, work outside the home (too frequent in our sample to be discriminant).

b) Others are zero, i. e. "excessive gain in weight" (which seems **to act as a protection,** as it is more frequent in Group C than in Group R (7)).

c) Some factors were assessed **with difficulty in our retrospecitve sample.** These are the group of the 4 "accidental factors": unusual effort, long travel, unusual fatigue, more than 10 cigarettes a day. The discriminant analysis yields coefficients which are very different from those of the ECRP, but which we cannot accept with confidence, as these characteristics are poorly known in this sample.

2.2.2 Second analysis

We decided to do another analysis after having excluded the 4 characteristics with negative coefficients in D 1, and the 4 accidental factors. We obtained a second discriminant function, D 2, calculated from the remaining 27 characteteristics, the coefficients of which (Tab. II) are very similar to those in D 1. Two factors have a

Tab. II. 35 characteristics composing the E C R P.

Characteristics	Significance of the difference of frequencies in Group C and R	Coefficient in E C R P	Coefficient in D 1	Coefficient in D 2
Low social class	*	1 or 3	1	1
Unwed mother		2	1	.
Unfavorable age $\leqq 20$ or $\geqq 40$		2 or 4	1	.
Height < 150 cm		3	6	6
Weight < 45 kg		3	2	2
More than 2 children without domestic help		1	2	1
More than 2 D and C		2 or 3	2	2
Uterine malformation		5	4	4
Cylindrical uterus		3	1	.
Late abortion	*	5	1	1
Short interval since last pregnancy		1	2	1
Premature delivery	**	5	2	3
Work outside the home		1	1	.
Strenuous work	**	3	3	2
Apartment above 3rd floor without elevator		2	0	0
Long daily commuting	**	5	2	3
Unusual effort		3	7	.
Long travel		3	7	.
Unusual fatigue		1	3	.
More than 10 cigarettes a day		2	11	.
Excessive weight gain		1	0	0
Less weigth than the month before		3	4	3
Gain in weight lower than 5 kg	**	2	8	7
Proteinuria		2	3	2
Hypertension		2	2	2
Metrorrhagia	*	4	1	1
Placenta praevia	*	5	8	8
Hydramnios		5	6	6
Twins	***	5	16	15
Breech		3	5	5
Presenting part + 1 or 0 station	**	3	6	6
Thinned lower uterine segment	*	3	1	1
Shortened cervix	***	4	3	3
Patency of internal os	**	4	1	1
Contractile uterus	*	5	1	1

Significance * $p \leqq .05$ ** $p \leqq .01$ *** $p \leqq .001$

coefficient of zero: "excessive gain in weight", and "apartment above 3rd floor without elevator".

We studied the prognostic value of D 1 and D 2, for it was not evident which was the better; **D 2 is certainly more stable, being less sensitive to sample particularities, D 1 provides more information.**

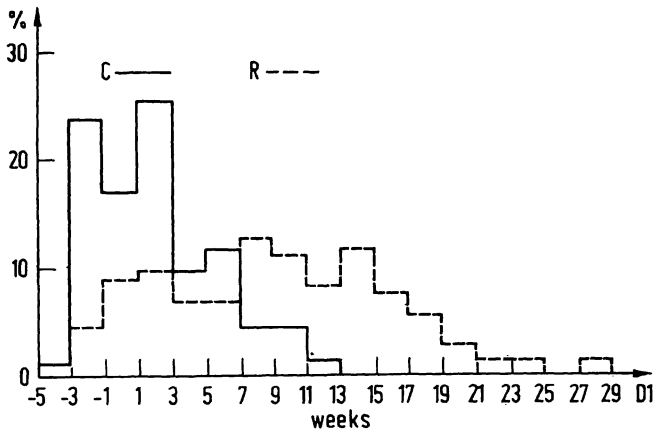

Fig. 2. Distribution of D 1 in Groups C (—) and R (- - -).

2.2.3 Study of D 1

The mean of D 1 in Groups C and R are different. This difference is significant ($p < 0.001$):

mean of D 1 in Group C: 2.5
mean of D 1 in Group R: 9.2.

Here again there is a large overlap between the two distributions; and for the higher values of the function there is a very high risk of belonging to Group R (Fig. 2).

If we construct **three classes of values for D 1, corresponding to three risk groups,** as was done previously for the ECRP; by choosing two limits such that the distribution of Group C is similar to that of the three ECRP classes, we obtain:

no risk: D 1 less than or equal to 3
no decision: D 1 between 4 and 6
high risk: D 1 greater than or equal to 7.

The distributions of Groups C and R are given in Tab. III.

Tab. III. Distribution of Groups C and R in D 1 risk classes.

D 1	C	R
$\leqq 3$	68%	25%
4—6	19%	10%
$\geqq 7$	13%	65%

We see that only 13% are "false alarms" (instead of 17% with the ECRP) and if **25% misclassifications** still remain in R, 65% cases in this group are now in the high risk class.

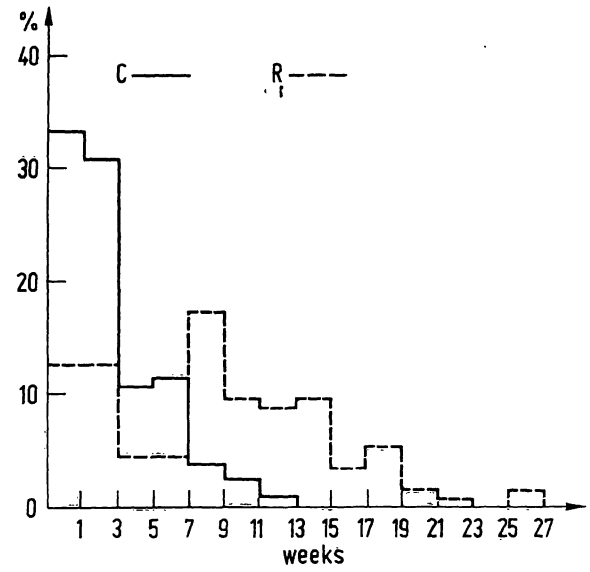

Fig. 3. Distribution of D 2 in Groups C (—) and R (- - -).

Tab. IV. Distribution of Groups C and R in D 2 risk classes.

D 2	C	R
$\leqq 3$	67%	26%
4—6	19%	8%
$\geqq 7$	14%	66%

2.2.4 Study of D 2

The mean of D 2 in Groups C and R are significantly different at $p < 0.01$:

mean of D 2 in Group C: 3.0
mean of D 2 in Group R: 9.0.

The distribution of D 2 in Groups C and R is similar to that of D 1, i. e. existence of a common area, and for the higher values of the function, a great chance of belonging to Group R (Fig. 3). As for D 1, we constructed three risk classes for D 2 (Tab. IV), with similar results.

2.3 Comparison of the ECRP with the discriminant functions

First looking at the tables giving the distribution of the risk classes defined by the ECRP, by D 1, and by D 2 in Groups C and R, we see that, in our sample, with a similar percentage of false alarms, the **discriminant functions give a**

better prediction of prematurity (10% improvement).

To test the higher predictive values of D 1 and D 2, as compared to that of the ECRP, we performed a covariance analysis on each pair of prediction functions in turn. This analysis tests whether, when one function is fixed, the other still has a discriminative value between the Groups C and R. In the sample:

for constant D 1 or D 2, the ECRP is no longer predictive,

when the ECRP is kept constant, D 1 and D 2 remain predictive,

for D 1 constant, D 2, which is practically "contained" in D 1, is no longer predictive, and for D 2 constant, D 1 remains predictive. This last result is very logical, as, for our sample, D 1 is theoretically closer to the "best" discriminant function than D 2 which can be considered as having zero coefficient for several characteristics (those excluded from the second analysis). **The best value for these coefficients is given by D 1,** if reality is not too far from the theoretical model.

Thus, in our sample, we have improved on the ECRP prediction of prematurity by the use of discrimination.

3. Discussion

The empirical function of risk of prematurity which we propose provides better results than those which have been presented to date, particularly by RANTAKALLIO [8].

RANTAKALLIO classifies into two groups: low risk and high risk (i. e. there is no area of "no decision"). In the sample on which she bases her function, her decision rule gives the following results: **41% in Group R well-classified with 59% omissions, 88% in Group C well-classified with 12% false alarms.** This lower proportion of successes does not appear to be due to the method but to choice of the recorded characteristics. In particular, RANTAKALLIO's remarkable analysis did not use any characteristic relating to the course of the pregnancy.

In other studies [such as 2], **a similar linear multiple regression analysis is a method of either explaining, or predicting birth weight.** Often not only characteristics of the mother and of the course of her pregnancy are used, but also characteristics of the delivery, and even of the infant (sex, gestational age). It is obvious that now the authors are no longer concerned only with prediction.

As we had access to a sample of 750 women from a **prospective study** being analysed at present, we used it to test which was the most predictive function when applied to this sample: the ECRP, D 1 or D 2. We performed the analysis of covariance between each pair of variables with the following results:

1. when the ECRP is constant, both D 1 and D 2 remain predictive,
2. when D 1 is constant, the ECRP and D 2 remain predictive, but D 2 is, as expected, more predictive. Although D 1 is more accurate than D 2 when applied to the sample from which it was calculated, it is however, too dependent on the particularities of the first sample and generally less stable.
3. D 1 and the ECRP have no value when D 2 is constant.

Hence, on the test sample, **it is the discriminant function D 2 which is most predictive,** even though it does not include several characteristics. But, when we look at the distribution of Group C and Group R of this sample in the risk classes defined by the ECRP, D 1 or D 2 **the difference is minimal.**

When we began this study, we thought that, by approximating a theoretical model of discrimination and, in particular, by using variables which were nearly normal, we could improve our prediction, even though this entailed the risk of losing some of our information. **Thus we performed a discriminant analysis with the characteristics in quantitative form** whenever possible (e. g. mother's age) and we constructed pseudo-quantitative variables representing families of similar signs in the following way: each of these variables was equal to the number of signs in the family present for the patient (number of obstetrical antecedents, number of signs of fatigue, number of signs of threat of imminent delivery . . .). This amounts to giving the same weight to all signs in a family, as compared with

our previous functions ECRP, D 1 or D 2, where within the same family, the point values vary from 1 to 5.

This function was less predictive than the ECRP even for the first sample. We conclude that what gives a better predictive value to the function in the individual weighting of each sign, and not the relative importance of a family of signs considered "equal", in comparison with another family.

In Tab. II, we can see which signs differ between the ECRP, D 1 and D 2. We also selected those signs which had a different distribution in Group C and in Group R: the first column of Tab. I gives the level of significance of these differences. We observe that the significant differences are associated with signs having values higher than two except for one characteristic "gain in weight lower than 5 kg", which is of importance in D 1 and D 2.

Signs having a high coefficient but for which the difference of the distribution in R and C is not significant are not frequent: uterine malformation, breech, hydramnios . . . A few signs disappear completely in D 2: unwed mother, unfavorable age, cylindrical uterus, work outside the home, apartment above 3rd floor without elevator, excessive weight gain.

Some factors have less value in D 1 and D 2 than in ECRP: obstetrical antecedents, metrorrhagia, and signs of threat of imminent delivery, except presenting part station + 1 or 0. But the grave signs, almost predictive by themselves, become very important in D 1 or D 2: hydramnios, placenta praevia, twin pregnancy, weight-gain lower than 5 kg, mother's height less than 150 cm.

We have seen that, if in the sample from which D 1 and D 2 have been calculated, these functions give a better prediction than the ECRP, on the test-sample, **D 2 is only slightly better than the ECRP.** Why is this improvement no more marked?

a) **The empirical function is composed of variables and weights based on the abundant information in the literature** and **on data resulting from wide medical experience** with cases which are not taken in consideration in the present study. The discriminant functions, however, are necessarily calculated from the information provided by limited samples. Here we see the difficulty of comparing the functions.

b) Two other difficulties result from the nature of the sample:

The ECRP was established as an assessment at 32 weeks of gestation, but D 1 and D 2 are calculated on a sample for which the moment of examination is more variable. Consequently, signs of threat of imminent delivery can be observed at 36 weeks of gestation for mothers in Group C (ANDERSON and TUNBULL [3] have the same results). This reduces both their predictive value and their weighting in D 1 or D 2 in comparison with that in the ECRP, and the predictive value of D 1 and D 2.

The ECRP was established independently of the treatments prescribed, but D 1 and D 2 were calculated on a sample of women, some of whom were high risk cases, but who after treatment, gave birth to infants of normal weight, which also reduces the predictive value of D 1 and D 2.

4. Conclusion

In conclusion, we can only insist on the importance of a **multifactorial prediction of prematurity.** The present study points out that the **ECRP provides an exact prediction in a large number of cases,** even if there is still an appreciable percentage of error, and in all cases it provides a better prediction than any individual characteristic. **Discriminant analysis provides a slightly more accurate prediction.** To improve the discriminant function, one needs **larger samples,** but in every case, we face the difficulty that women are treated presenting a high risk. It might be possible to improve the prediction by taking into account the different moments when signs assume a predictive value: socio-economic characteristics, mother's height, antecedents . . ., time lapse since conception, pathology at the beginning, in the middle or at the terminal stages of pregnancy. Such a study is at present being realized in the analysis of other surveys: prospective surveys, with larger numbers, (prospective survey at "Maternité de Port-Royal", surveys carried out by the "Institut National de la Santé et de la Recherche Médicale"), in which we have collected detailed information.

Summary

This study was performed to permit the comparison of the results of two methods **predicting the occurrence of premature delivery.** The first is an **empirical calculation** of the risk **(Fig. 1, Tab. I),** done at 32 ± 2 weeks of gestation, the second a function established **according to a discriminating analysis** which should permit us, using the best statistical method, to improve our possibilities of prediction and thus to prevent the risk **(Figs. 2—3, Tabs. II—IV).**

The discriminating analysis was established from clinical data gathered at 32 weeks of gestation for all the mothers having given birth to infants weighing less than 2500 g at the Obstetrical Department of Port Royal in 1969, compared with clinical data on 222 mothers of newborn infants more than 2500 g at birth, chosen at random.

We are able to show that **the function established through the discriminant analysis is,** as expected, **much more accurate when applied to the same sample as the empirical method.**

We have checked these results with a **prospective sampling** of 750 women visiting the prenatal clinic of Port-Royal. We established that in this second sample the calculated function is still better, but the **difference between the results of the calculated and the empirical evaluation of risk is considerably reduced.**

The discussion considers this last finding, the possibility of accurate prediction and the absence of important difference between the two functions.

Keywords: Prematurity (risk factors), prenatal care, statistics (discriminant analysis).

Zusammenfassung

Multifaktorielle Studie über das Frühgeburts-Risiko in der 32. Schwangerschaftswoche. II. Ein Vergleich zwischen empirischer Vorhersage und Diskriminanzanalyse.

Der Zweck dieser Studie ist ein Vergleich zweier Methoden, die eine **Vorhersage über das Auftreten einer Frühgeburt ermöglichen** sollen. Erstere besteht in einer **empirischen Abschätzung** des Risikos in der 32. (30.—34.) Gestationswoche **(Fig. 1, Tab. I).** Die zweite stellt eine **diskriminanzanalytische Funktion** dar, die aufgrund exakter statistischer Methoden eine Verbesserung der Vorhersage und damit eine Verminderung des Risikos ermöglichen soll **(Figs. 2—3, Tabs. II—IV).**

Die Diskriminanzanalyse wurde aus den klinischen Befunden der 32. Schwangerschaftswoche von allen Müttern, die 1969 in der Maternité de Port-Royal ein Kind unter einem Geburtsgewicht von 2500 g zur Welt brachten, berechnet und mit den klinischen Befunden von 222 willkürlich ausgewählten Müttern verglichen, deren Neugeborenes mehr als 2500 g wog.

Wir konnten zeigen, daß **die Funktion mittels der Diskriminanzanalyse,** wie erwartet, **beim gleichen Fall genauer als die empirische Methode** ist.

Wir haben diese Ergebnisse mit einer **prospektiven Studie** an 750 Frauen verglichen, die die Geburtsklinik von Port-Royal aufsuchten. Dabei stellten wir fest, daß bei dieser zweiten Untersuchung die berechnete Funktion noch besser ist, jedoch der **Unterschied zwischen den Ergebnissen der berechneten und der empirischen Abschätzungen des Risikos beträchtlich vermindert** ist.

Die Diskussion behandelt die letzten Ergebnisse, die Möglichkeit einer genauen Vorhersage und das Fehlen eines bedeutsamen Unterschiedes zwischen den beiden Funktionen.

Schlüsselwörter: Frühgeburt (Risikofaktor), Schwangerschaftsfürsorge, Statistik, Diskriminanzanalyse.

Résumé

Etude multifactorielle sur le risque de prematurité dans la 32ème semaine de grossesse. II. Comparaison entre une prévision empirique et une analyse discriminante.

Cette étude a pour but de comparer les résultats d'une **appréciation empirique du risque de survenue prématurée de l'accouchement** faite à 32 semaines ± 2 de grossesse **(Fig. 1, Tab. I)** et d'une **fonction établie par une analyse discriminante** dans le but d'améliorer, par la meilleure technique statistique, la prévision en vue de la prévention **(Figs. 2—3, Tabs. II—IV).**

L'analyse discriminante a été faite sur des renseignements cliniques enregistrés à 32 semaines de gestation pour les mères de tous les enfants nés en 1969 à la Maternité de Port-Royal et pesant moins de 2500 g en les comparant avec les renseignements cliniques necucillis pour 222 mères d'enfants témoins pesant plus de 2500 g à la naissance.

Nous montrons que la **fonction calculée par l'analyse discriminante est bien entendu meilleure que la fonetion empirique sur l'échantillon étudié.**

Nous verifious ces résultats sur un **échantillon prospectif** de 750 femmes suivies à la consultation prénatale de

la Maternité de Port-Royal. Nous montrons alors que sur le second échantillon, la **fonction calculée reste légèrement meilleure mais que les différences dans les résultats sont nettement plus faibles.** La discussion porte sur ce problème, sur les possibilités de prédiction du risque de prématurité et sur l'absence de grande différence entre les fonctions.

Mots-clés: Prématurité (facteurs de risque), surveillance prenatale, statistiques (analyse disciminante).

Acknowledgement

This study was supported by a grant from the Ministery of Public Health and the Department "Maternité-Pédiatrie" of the Institut National de la Santé et de la Recherche Médicale.

Bibliography

[1] Abernathy, J. R., B. G. Greenberg, J. F. Donnely: Application of discriminant functions in perinatal death and survival. Amer. J. Obstet. Gynec. 95 (1966) 860

[2] Abernathy, J. R., B. G. Greenberg, H. B. Wells, T. M. Frasier: Smoking as an independant variable in a multiple regression analysis upon birth weight and gestation. Amer. Public Health 50 (1966) 626

[3] Anderson, A. M., A. C. Turnbull: Relationship between length of gestation and cervical dilatation, uterine contractibility and other factors during pregnancy. Amer. J. Obstet Gynec. 105 (1969) 1207

[4] Greenberg, B. G., H. B. Wells: Linear discriminant analysis in perinatal mortality. Amer. J. Public Health 53 (1963) 594

[5] Papernik, E.: Coefficient de risque d'accouchement prématuré. La Presse Médicale 77 (1969) 793

[6] Papiernik, E.: Coefficient of risk of accidental prematurity. 6th World Congress of Gynecology and Obstetrics, New York 1970

[7] Papiernik, E., M. Kaminski: Multifactorial study of the risk of prematurity at 32 weeks of gestation. I. A study of the frequency of 30 predictive characteristics. J. Perinat. Med. 2 (1974)

[8] Rantakallio, P.: Groups at risk in low birth weight infants and perinatal mortality. Acta Ped. Scand. suppl. 193 (1969)

[9] Rouquette, C., J. Goujard: Prématurité et insuffisance pondérale, étude de quelques facteurs étiologiques. Bulletin de l'INSERM 22 (1967) 733

[10] Saling, E.: Prämaturitäts- und Dysmaturitäts-Präventions Programm (P. D. P.-Programm). Z. Geburth. u. Perinatol. 176 (1972) 70—81

M. Kaminski
Unité de Recherches Statistiques de l'Institut National de la Santé et de la Recherche Médicale
16 bis, avenue Paul Vaillant Couturier
F-94800 Villejuif/France