AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 123, No. 1
Printed in U.S.A.

A SIMPLE MODEL RELEVANT TO TOXOPLASMOSIS APPLIED TO EPIDEMIOLOGIC RESULTS IN FRANCE

LAURE PAPOZ,[1] FRANÇOIS SIMONDON,[2] WILLIAM SAURIN,[3] AND HALA SARMINI[4]

Papoz L. (INSERM U21, 16 Avenue Paul Vaillant Couturier, F 94807 Villejuif Cédex, France), F. Simondon, W. Saurin, and H. Sarmini. A simple model relevant to toxoplasmosis applied to epidemiologic results in France. *Am J Epidemiol* 1986;123:154–61.

A simple mathematical model was applied to the results of a seroepidemiologic study of toxoplasmosis carried out in France in 1982–1983. An adequate fitting to the prevalence data observed on 7,605 women of childbearing age was obtained. Thus, the data were used to estimate the seroconversion rate, allowing the approach to and discussion of insights gained from the model, such as the risk of *Toxoplasma* infection during pregnancy, the age-related expected risks of maternal seroconversion, as well as an overall prediction of the yearly number of congenitally infected infants in the total population. In addition, the high prevalence of congenital toxoplasmosis in France and the excess risk encountered by young migrant women from lower prevalence areas were confirmed by the model. The model might therefore be useful for public health purposes in other countries.

models, theoretical; toxoplasmosis

The main problem with toxoplasmosis is primary infection (1) during pregnancy. In France, prevention of toxoplasmosis is based on serum antibody determination. In 1978, this test was required for any woman attending the mandatory prenuptial medical examination (Decree no. 78-396, March 17, 1978). Furthermore, obstetricians are strongly advised to perform the serologic surveillance of any nonimmune pregnant woman. Although widely applied, such measures do not allow the assessment of the seroconversion rate on a national level, because data are not centralized. Only crude prevalence rates of seropositive tests are provided by regional laboratory centers. The figures, ranging from 40 per cent to 80 per cent, and even more in the Parisian area, are mentioned as among the highest in the world (1). Some useful information was obtained in studies conducted in maternity hospitals (2, 3). According to the various sources, the proportion of acquired toxoplasmosis among pregnant women would be from 0.5 per cent to 1 per cent. About half of the offsprings are affected in the absence of treatment. In the Desmonts and Couvreur study (3), short-term complications were observed in 14 per cent of the children and death in 6 per cent. In spite of the lack of exhaustive and accurate data, it is clear from these figures that acquired toxoplasmosis is particularly frequent in France, compared with other countries, such as Norway (4) and the United States (5), where only 12 per cent and 30 per cent,

Received for publication July 19, 1984, and in final form March 19, 1985.

[1] Unité de Recherches Statistiques, INSERM U21, 16, Avenue Paul Vaillant Couturier, F 94807 Villejuif Cédex, France. (Reprint requests to Laure Papoz.)

[2] Unité de Recherche Population Santé ORSTOM, Paris, France.

[3] Unité de Programmation Moléculaire et de Toxicologie Génétique, CNRS/INSERM, Institut Pasteur, Paris, France.

[4] Laboratoire CERBA, Cergy-Pontoise, France.

This study was supported by the French Ministry of Health and Social Affairs, Direction of Health.

The authors thank E. Garat-Lesieux and S. Cénée-Prod'homme for their helpful technical assistance.

respectively, of the women are reported "positive" at childbearing age.

Such a high risk of congenital infection represents a significant public health problem; as a consequence, it is necessary to assess the actual impact of the disease, in terms of expected number of primary infections in one year. The use of mathematical models for description and prediction purposes has already been the occasion for a number of publications, especially for rubella, herpes simplex, variola, measles, and some other communicable diseases (6). To our knowledge, very few authors have proposed a model to study the endemicity of toxoplasmosis, still one of the most frequent congenital infections. Some theoretic curves resulting from a simulation were shown in the comprehensive review by Frenkel in 1973 (7).

To reinforce the preventive action initiated in 1978, in 1982–1983 the French Ministry of Health decided to carry out a national seroepidemiologic survey among women of childbearing age. The importance of the prevalence data provided by age offered the opportunity to test the adequacy of a probability model, as described below.

Methods

The model

The model used is based on the following single hypothesis: Each individual in a given population is submitted to a risk (r) which is the probability of contact with *Toxoplasma gondii* during one year. The risk is only related to environment and therefore is the same for all subjects, and identical for each year of life, from birth to death. Seroconversion occurs at the first contact. Thus, r corresponds to the seroconversion rate or incidence rate.

Let P_x be the probability of being seropositive at age x (x in years) and $Q_x = 1 - P_x$ the probability of being seronegative at age x in the same population. Thus,

$$P_x = \sum_{i=1}^{x} r(1 - r)^{i-1} \quad (1)$$

and

$$Q_x = (1 - r)^x. \quad (2)$$

The basic equations and the connected hypotheses can be found in Lilienfeld's handbook of epidemiology (8). It appears that r can be theoretically estimated if one knows the prevalence of serum *Toxoplasma* antibodies in a group of subjects born the same year and living in the same country since birth.

It follows from equation 2 that the risk of primary infection during pregnancy between age x and x' will be

$$R_{x,x'} = 0.75\ (1 - r)^x \sum_{i=x+1}^{x'} r(1 - r)^{i-x-1}\, g_i, \quad (3)$$

where g_i is the probability of a woman becoming pregnant at age i among the female population of childbearing age, and 0.75 is the duration of gestation in one year.

If a cohort of migrant people move at a given age a ($a < x$ and migration is assumed to occur on an individual's birthday) from a country 1 characterized by a risk r_1 to another country 2 characterized by a risk r_2, equations 1 and 2 must be replaced, respectively, by

$$P'_x = \sum_{i=1}^{a} r_1(1 - r_1)^{i-1} + \sum_{i=a+1}^{x} r_2(1 - r_1)^a\ (1 - r_2)^{i-a-1} \quad (4)$$

and

$$Q'_x = (1 - r_1)^a\ (1 - r_2)^{x-a}. \quad (5)$$

Population

Women were recruited who came for biologic blood sample tests in 69 medical analyses laboratories in 14 administrative regions in France. Selection referred only to age, which was required to be at least 15 years and below 45 years. Serum samples were systematically sent to one laboratory where *Toxoplasma* antibodies were tested by the immunofluorescence method (9), under the responsibility of one of the authors (H. S.). Evidence of previous infection was assessed by a result at least equal to 10 IU.

In the course of the study, between No-

vember 1982 and April 1983, results were obtained from 7,605 women of French extraction and 680 migrant women.

The reason for examination was a systematic medical visit in 72.5 per cent of the cases: prenuptial 2.5 per cent, antenatal 31.3 per cent, routine check-up 38.7 per cent. The remaining 27.5 per cent were related to a health problem.

Statistical methods

Estimations of r were first calculated directly, and the corresponding variances were estimated by the δ method (10). An overall estimation was then obtained by the maximum likelihood method. The search for the zero of the likelihood derivative was achieved by a dichotomous method (11). (Computer routines are available upon request to the authors.) The confidence interval for the maximum likelihood estimate was obtained by the asymptomatic normality and variance for maximum likelihood estimates. The adequacy of the model was tested by comparison of observed and expected frequencies by the χ^2 test. Other results were obtained by computer simulation procedures.

RESULTS

Testing the model

This section applies to the data observed on women of French extraction.

Point estimations. According to equation 2, a series of 30 independent estimates of r were obtained from the percentage, q_x, of seronegative women observed in the 30 age groups, x, by

$$\hat{r}_x = 1 - \exp(\log q_x/x), \qquad (6)$$

and the corresponding approximate variance was obtained by the δ method.

$$\hat{s}^2_{r_x} = \frac{p_x(\hat{r}_x - 1)^2}{n_x x^2 q_x}, \qquad (7)$$

where p_x is the proportion of seropositive women and n_x the size of the age group x.

The results, $\hat{r}_x$, range from 0.027 to 0.051 (table 1). There is no remaining linear effect of age (correlation coefficient = 0.00008). After weighting according to the inverse of the variances $\hat{s}^2_{r_x}$, a weighted average of 0.0369 is found.

Maximum likelihood. A better estimation can be expected using the maximum likelihood method, in which each woman is defined by her serologic status (0 = negative, 1 = positive). At age $x = x_i$, let n_{i0} be the number of seronegatives and n_{i1} the number of seropositives. Hence, for $i = 1, \ldots k$, the logarithmic likelihood is

$$L = \log(1 - r) \sum_{i=1}^{k} n_{i0}\, x_i + \sum_{i=1}^{k} n_{i1} \log\{1 - (1 - r)^{x_i}\}. \qquad (8)$$

It can be shown that the derivative $\frac{\partial L}{\partial r}$ is close to zero for $\hat{r} = 0.0375$ with a 95 per cent confidence interval equal to 0.0364–0.0386. When the variability of the preceding 30 estimates is studied by reference to the maximum likelihood estimate, the distribution of $(\hat{r}_x - 0.0375)/\hat{s}_{rx}$ appears to be approximately normal with mean = 0.12, and only three values out of 30 are found out of 2 standard deviations from the mean.

Accordingly, the comparison of the expected and observed frequencies within the 30 age groups provides a χ^2 value equal to 33.9, corresponding to a p value of about 22 per cent (table 1). Thus, it can be considered that the model is acceptable. The following results were established using the number provided by the maximum likelihood method.

Using the model

Prediction of congenital infections. Once an estimation of r and the age distribution of the pregnant women in the general population are available, it is possible to approach the annual number of primary infections by *Toxoplasma* during pregnancy. The proportions, g_x, of pregnancies among women of age x between ages 15 and 44 years were replaced in equation 3 according to the French national statistics (12). The

TABLE 1

Estimations of the seroconversion rate (r) in 30 age groups and comparison between observed and expected frequencies, according to the maximum likelihood estimate

Age (x) (years)	No. of women ($n_{x0} + n_{x1}$)	% of negative women (q_x)	Estimate of r ($\hat{r}_x$)	$s^2_{\hat{r}_x}$ ($\times 10^{-6}$)	Observed no. of negative (n_{x0})	Expected no. of negative
15	68	45.6	0.0510	70.2	31	38.3
16	131	48.9	0.0438	28.5	64	71.0
17	199	48.7	0.0414	16.8	97	103.9
18	246	46.7	0.0414	13.2	115	123.6
19	324	46.9	0.0390	8.9	152	156.7
20	378	42.6	0.0418	8.2	161	176.0
21	434	49.5	0.0329	5.0	215	194.5
22	442	41.6	0.0390	6.1	184	190.6
23	450	39.6	0.0395	5.9	178	186.8
24	416	42.3	0.0352	5.3	176	166.2
25	417	41.0	0.0350	5.1	171	160.4
26	441	37.6	0.0369	5.2	166	163.2
27	391	47.8	0.0270	3.6	172	139.3
28	360	35.8	0.0360	5.9	129	123.4
29	308	37.0	0.0337	6.1	114	101.6
30	313	31.9	0.0373	7.0	100	99.4
31	299	32.8	0.0353	6.6	98	91.4
32	291	23.0	0.0449	10.2	67	85.6
33	235	25.1	0.0410	10.7	59	66.6
34	227	28.2	0.0366	9.0	64	61.9
35	203	29.6	0.0342	8.9	60	53.3
36	205	24.4	0.0384	10.8	50	51.8
37	139	17.3	0.0464	22.8	24	33.8
38	115	21.7	0.0394	20.0	25	26.8
39	115	21.7	0.0384	19.1	25	25.9
40	117	14.5	0.0471	28.6	17	25.4
41	91	12.1	0.0502	42.5	11	19.0
42	72	16.7	0.0418	36.1	12	14.5
43	94	19.1	0.0377	22.6	18	18.2
44	84	21.4	0.0344	21.1	18	15.6
Total	7,605	36.5	0.0369*		2,773	2,784.7

* Weighted average according to the inverse of the variances of the $\hat{r}_x$'s.

estimated numbers of maternal seroconversions by age are shown in table 2; the absolute and relative distributions of the risk are represented in figure 1. Globally, a figure as high as nearly 8,600 (1.06 per cent) acquired *Toxoplasma* infections may be expected each year. In total absence of intervention, it might result in more than 5,000 cases (6.4/1,000) of congenital toxoplasmosis among the offsprings, on the basis of an incidence rate equal to 60 per cent, as reported by Desmonts and Couvreur (3).

Interpopulation differences. It would be of great interest to predict the rate of maternal infection in different populations according to the corresponding seroconversion rates. Equation 3 clearly indicates that $R_{x,x'}$ is curvilinear and passes through a maximum for a given value of r. Unfortunately, it was impossible to assess the interpopulation relationship mathematically, since the maternal infection rate depends on the age distribution among pregnant women in each country. Thus, we used the French data g_x to obtain by a simulation procedure the values of $R_{x,x'}$ for $x = 15$ years and $x' = 44$ years according to values of r ranging from 0 per cent to 20 per cent. It is observed that the percentage of maternal infections increases up to 1.07 per cent for $r = 0.038$ and decreases thereafter, so that the previous estimate of r in the French population is close to the highest possible risk for pregnant women (figure 2). It is important to note that two values of r as different as 0.01 and 0.098 lead to the same risk (0.6 per cent).

Migrant women. Young women moving

TABLE 2

Expected risk of acquired toxoplasmosis during pregnancy, according to age

Age (x) (years)	Proportion of pregnant women*	% of acquired toxoplasmosis† during pregnancy	No. of seroconverters
15	0.3	1.65	5
16	1.2	1.58	22
17	3.8	1.53	66
18	8.6	1.47	143
19	16.8	1.41	271
20	26.5	1.36	411
21	36.3	1.31	541
22	45.1	1.26	648
23	51.9	1.21	718
24	57.0	1.17	758
25	58.7	1.12	751
26	58.2	1.08	717
27	55.0	1.04	652
28	49.7	1.00	567
29	45.1	0.96	496
30	39.5	0.93	417
31	34.7	0.89	353
32	28.6	0.86	280
33	23.8	0.83	224
34	19.1	0.80	174
35	14.5	0.77	127
36	9.1	0.73	76
37	7.2	0.72	59
38	5.3	0.68	41
39	3.6	0.65	27
40	2.5	0.64	18
41	2.0	0.61	14
42	1.2	0.58	8
43	0.8	0.57	5
44	0.4	0.54	3
Total	706.5	1.06	8,592

* Number of pregnant women aged x years per 10,000 out of the total female population aged 15–44 years, according to French national statistics.

† Expected percentage of acquired toxoplasmosis (calculated according to the proposed model) among the total number of pregnancies observed in a year.

from a country of low prevalence to a country of high prevalence are exposed to an increased risk of primary infection during pregnancy. For instance, figure 3 shows the cumulative percentage of seropositive women among a cohort having moved at age 15 years from a country where $r = 0.01$ to a country where $r = 0.037$. The increment of prevalence between ages 15 and 44 years is then more than 57 per cent instead of 36 per cent. In the screened sample, the prevalence rates were found markedly lower among migrant women (globally, 58 per cent) compared with the French women (globally, 67 per cent).

DISCUSSION

In the first part of the results, it is established that the model adequately describes epidemiologic data. This is consistent with a body of evidence from numerous studies (7, 13) showing that within a given population, age is not only the main variate related to *Toxoplasma* infection but accounts for nearly all the differences between groups. This means that the duration of exposure is sufficient to describe the phenomenon in a given population, in spite of the possible variations over space and time of the complex underlying factors leading to seroconversion. The points to discuss are the degree of representativeness of the studied sample and the reliability of the serologic test used. As a matter of fact, the subjects were included consecutively during a given period and not randomly selected among the whole population. Regardless of the fact that a random selection method, although highly desirable, would raise serious difficulties, it may be underlined that the size of the age groups was taken into account in the calculations. In addition, the prevalence rates compared according to socioeconomic status or to the reason for a biologic test prescription did not show any marked difference.

On the other hand, it cannot be rejected that a serologic test using the immunofluorescence method could be less sensitive than the dye test and therefore may produce less positive results. In this study, we obtained lower percentages of seropositive women than did Desmonts et al. (14) 20 years ago in a large prospective survey. The origin of the population under screening, however, was restricted to the Parisian area, where the risk of contamination is particularly high (because undercooked meat is more often consumed (15)), whereas the present study is a national one. Moreover, it is likely that the spread of toxoplasmosis has changed over time with the improvement of hygiene standards, the decreasing number of stray cats, the use of

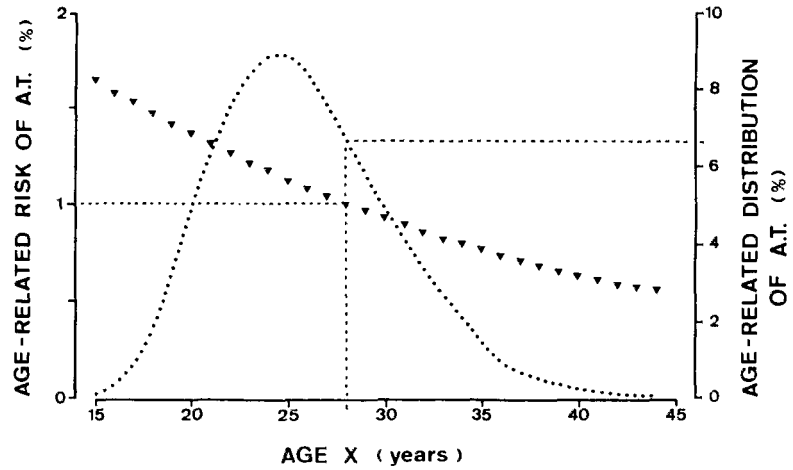

FIGURE 1. Expected risk of acquired toxoplasmosis (A.T.) during pregnancy from ages 15 to 44 years. The decreasing curve corresponds to the absolute risk and the dotted line to the calculated age-related distribution of acquired toxoplasmosis. For instance, for any 28-year-old woman becoming pregnant, the risk is equal to 1 per cent, while 6.6 per cent of the total cases of acquired toxoplasmosis occur in this age group.

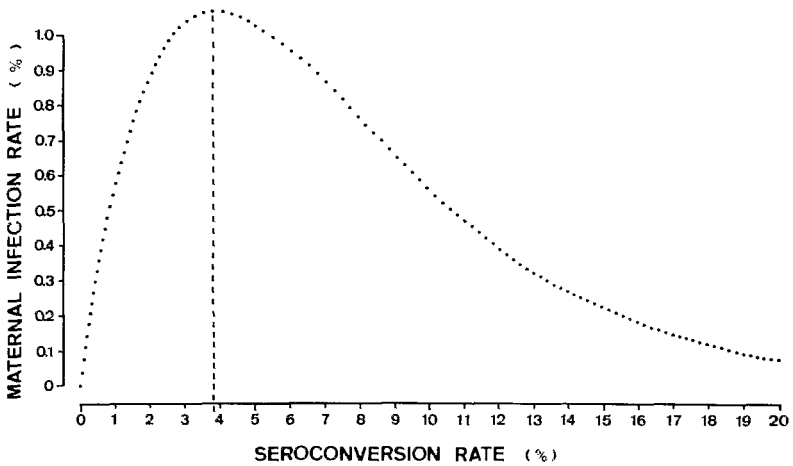

FIGURE 2. Expected maternal infection rate $R_{xx'}$ (per cent) for $x = 15$ years and $x' = 44$ years according to the seroconversion rate r (per cent) in the French population aged 15–44 years. A maximum value of 1.07 per cent is obtained for $r = 3.8$ per cent.

canned diets for domestic animals, and of frozen food for humans.

In the second part of the results, the model was used to predict the number of *Toxoplasma* infections during pregnancy. We found a rate of about 10 seroconversions in 1,000 pregnancies, a figure which is finally consistent with the data reported in previous studies: According to the number of seroconverters observed in different samples by Desmonts et al. (14) and Desmonts and Couvreur (16), the rate lies between four and 12 maternal infections per 1,000 pregnancies. More recently, Lapierre et al. (17) followed 15,132 pregnant women from 1977 to 1982 in another Parisian center and found a rate of 5.35 maternal infections per 1,000 pregnancies. It must be kept in mind that these figures were calculated on crude data obtained under routine conditions, without taking age into account or evaluating the proportion of drop-out subjects. Thus, the theoretic estimation method used here is probably relevant to

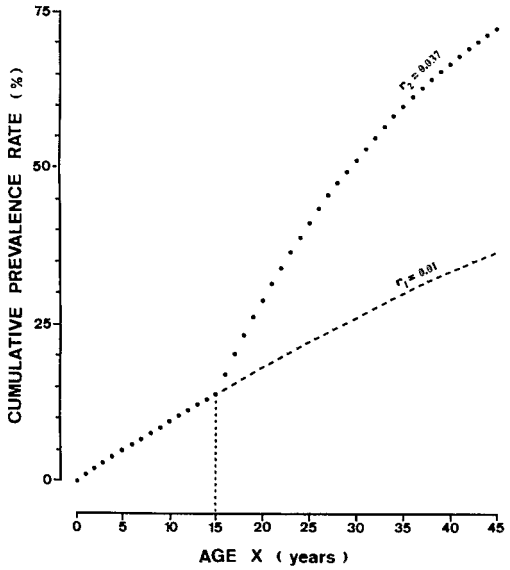

FIGURE 3. Calculated cumulative prevalence rate P'_x (per cent) according to age x up to 45 years in a cohort of migrant women having moved at age 15 from a country where the seroconversion rate (r) is equal to 1 per cent to a country where $r = 3.7$ per cent. The lower curve represents the cumulative prevalence rate P, corresponding to $r = 1$ per cent in the absence of migration, according to equation 1 (see text). The upper curve represents the cumulative prevalence rate P' after migration, according to equation 4 (see text).

the actual impact of toxoplasmosis on public health and can be applied in any population. Reliable comparisons could then be undertaken, in order to understand, for instance, the mechanisms leading to wide differences in the seroconversion rate in similar countries.

The equations derived from the model provided a mathematical description for a number of things that were previously observed in different studies. Equation 3 showed that exposure to *Toxoplasma* infection for pregnant women is maximum for an intermediate value of the seroconversion rate, r. According to the estimate of r, we found that the exposure was close to the maximum for pregnant women in the French population. Simultaneously, this explains why the risk may be identical in countries of low and high prevalence, as already mentioned (4, 7). When equation 4 is used, it is possible to assess quantitatively the increased risk of seroconversion for people coming from a country of low prevalence to a country of high prevalence, a phenomenon which deserves the consideration of the public health services.

In any kind of population, similar information may be derived from serologic data by age groups. Unfortunately, the data available from other countries that we used were not sufficiently documented regarding age to allow their analysis by our method. It is essential that the age groups should be of one year and preferably of equal size. If this is so, the adequacy of the model can be first verified using equation 2 in a few well chosen age groups, for example 15, 25, and 35 years. The subsequent calculations are easy to achieve with the help of a microcomputer. Then, the prevalence rates at any age or the increment of prevalence over an age interval can be estimated with a minimal data set collection. This is the main interest of the model, facing the lack of hard data concerning toxoplasmosis in almost all countries, while the use of relevant information is the necessary first step for any efficient policy of prevention (18).

REFERENCES

1. Remington JS, Desmonts G. Toxoplasmosis. In: Remington JS, Klein JO, eds. Infectious diseases of the fetus and newborn infant. Philadelphia: WB Saunders, 1983:143–263.
2. Roux C, Desmonts G, Mulliez N, et al. Bilan de deux ans de prophylaxie de la toxoplasmose congénitale à la Maternité de l'Hôpital Saint-Antoine. J Gynecol Obstet Biol Reprod (Paris) 1976;5:249–64.
3. Desmonts G, Couvreur J. A prospective study of the offsprings of 542 women who acquired toxoplasmosis during pregnancy: pathophysiology of congenital disease. In: Thalhammer O, Baumgartner K, Pollak A, eds. Perinatal medicine, Sixth European Congress, Vienna 1978. Stuttgart: George Thieme, 1979:51–60.
4. Stray-Pedersen B. A prospective study of acquired toxoplasmosis among 8,043 pregnant women in the Oslo area. Am J Obstet Gynecol 1980;136:399–406.
5. Wilson CB, Remington JS. What can be done to prevent congenital toxoplasmosis? Am J Obstet Gynecol 1980;138:357–63.
6. Becker N. The uses of epidemic models. Biometrics 1979;35:295–305.
7. Frenkel JK. Toxoplasma in and around us. Bioscience 1973;23:343–52.
8. Lilienfeld AM, Lilienfeld DE. Foundations of ep-

idemiology. 2nd ed. New York: Oxford University Press, 1980:358–9.
9. Kelen AE, Ayllon-Leindl L, Labzoffsky NA. Indirect fluorescent antibody method in serodiagnosis of toxoplasmosis. Can J Microbiol 1962;8:545–54.
10. Rao CR. Linear statistical inference and its applications. 2nd ed. New York: John Wiley & Sons, 1973:388–9.
11. Mivoux M. Programmation mathématique (théorie et algorithmes). Book 1. Paris: Dunod, 1983:67–93.
12. INSEE. La situation démographique en 1981. Les collections de l'INSEE (D). 1982;94:41–6.
13. Stray-Pedersen B. Toxoplasmosis in Norway. In: Thalhammer O, Baumgartner K, Pollak A, eds. Perinatal medicine, Sixth European Congress, Vienna 1978. Stuttgart: George Thieme, 1979:89–91.
14. Desmonts G, Couvreur J, Ben Rachid MS. Le toxoplasme, la mère et l'enfant. Arch Fr Pediatr 1965;22:1183–1200.
15. Desmonts G, Couvreur J, Alison F, et al. Etude épidémiologique sur la toxoplasmose: de l'influence de la cuisson des viandes de boucherie sur la fréquence de l'infection humaine. Rev Franc Etudes Clin Biol 1965;10:952–8.
16. Desmonts G, Couvreur J. Toxoplasmosis in pregnancy and its transmission to the fetus. Bull NY Acad Med 1974;50:146–59.
17. Lapierre J, Tourte-Schaefer C, Heyer F, et al. Toxoplasmose congénitale: réflexions à propos de la surveillance sérologique de quinze mille femmes enceintes. Sem Hop Paris 1983;40:2741–5.
18. Henderson JB, Beattie CP, Hale EJ, et al. The evaluation of new services: possibilities for preventing congenital toxoplasmosis. Int J Epidemiol 1984;13:66–72.