Incidence and prevalence of Toxoplasma gondii infection in women in France, 1980–2020: model-based estimation

F. NOGAREDA 1,2, Y. LE STRAT 1*, I. VILLENA 3, H. DE VALK 1 AND V. GOULET 1

1 Department of Infectious Diseases, French Institute for Public Health Surveillance (InVS), Saint-Maurice, France
2 European Programme for Intervention Epidemiology Training (EPIET), European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden
3 Laboratoire de Parasitologie-Mycologie, EA3800, Centre National de Référence de la Toxoplasmose, CHU Hôpital Maison Blanche, Reims, France

Received 31 July 2013; Final revision 2 October 2013; Accepted 4 October 2013

SUMMARY

Toxoplasmosis is a worldwide zoonosis due to Toxoplasma gondii, a ubiquitous protozoan parasite of warm-blooded animals including humans. In pregnant women, primary infection can cause congenital toxoplasmosis resulting in severe malformations in the newborn. Since 1978, public health authorities in France have implemented a congenital toxoplasmosis prevention programme, including monthly serological screening of all seronegative pregnant women, and treatment in case of seroconversion. However, this programme does not produce systematic surveillance data on incidence and prevalence. Our objective was to estimate the incidence and prevalence of Toxoplasma infection, and the incidence of seroconversion during pregnancy in women in France. We used a catalytic model to estimate incidence and prevalence of Toxoplasma infection between 1980 and 2020 in women of childbearing age. We used age- and time-specific seroprevalence data obtained from the National Perinatal Surveys (NPS) conducted in 1995, 2003 and 2010. We assumed that incidence depends both on age and calendar time, and can be expressed as the product of two unknown functions. We also estimated incidence of seroconversion during pregnancy in 2010 from the NPS and the National Surveillance of Congenital Toxoplasmosis (ToxoSurv). We combined data of 42208 women aged 15–45 years with serology available from the three NPS. For women aged 30 years the modelled incidence decreased from 7·5/1000 susceptible women in 1980 to 3·5/1000 in 2000. In 2010 the incidence was 2·4/1000. The predicted incidence and prevalence for 2020 was 1·6/1000 and 27%, respectively. The incidence of seroconversion during pregnancy in 2010 was estimated at 2·1/1000 susceptible pregnant women (95% CI 1·3–3·1) from the NPS and 1·9 (95% CI 1·8–2·1) from ToxoSurv. Incidence and prevalence of Toxoplasma infection has decreased markedly during the last 30 years. This decrease may be explained by a lower exposure to the parasite by changes in food habits and by improved hygiene practices in meat production. Modelled estimations were consistent with estimates observed in other studies conducted previously in France. The catalytic modelling provides reliable estimates of incidence and prevalence of Toxoplasma infection over time. This approach might be useful for evaluating preventive programme for toxoplasmosis.

Key words: Incidence, modelling, prevalence, toxoplasmosis.

* Author for correspondence: Dr Y. Le Strat, Department of Infectious Diseases, French Institute for Public Health Surveillance, Saint-Maurice, France, 12 rue du Val d’Osne, 94415 Saint-Maurice cedex, France. (Email: y.lestrat@invs.sante.fr)
INTRODUCTION

Toxoplasmosis is a worldwide zoonosis, due to Toxoplasma gondii, a ubiquitous protozoan parasite of warm-blooded animals including humans. The definitive hosts of T. gondii are cats and other felines that acquire infection mainly from eating infected mammals, especially rodents or birds, and rarely from faeces of infected cats [1]. Infection in humans generally occurs by ingesting tissue cysts from undercooked meat or raw meat, mainly pork and lamb, or by consuming food or drink contaminated with sporulated oocysts from the environment [2, 3]. In humans, infection is usually asymptomatic or a mild flu-like illness; however, persons with compromised immune systems may experience severe neurological symptoms that can be life-threatening. Immunoglobulin (Ig) G antibodies to T. gondii are detectable early and reach a peak within 6 months after infection providing lifelong immunity. However, people with immunodeficiencies such as AIDS or malignancies are at risk of developing illness from reactivated infection. T. gondii infection in humans is widespread throughout the world. Prevalence of T. gondii infection (a positive serology for toxoplasmosis, which means the individual was infected at some time in his/her life with T. gondii) varies by geographical region due to differences in environmental conditions, cultural and food habits. Prevalence <30% was observed in USA, Great Britain, Scandinavia and South East Asia [4–7], while in countries in Africa and Latin America the prevalence was >60% [8–13]. In France, Toxoplasma infection prevalence decreased markedly in pregnant women during the past decades, from 83% in 1965 to 54.3% in 1995, 44% in 2003 and 37% in 2010 [14–18].

In pregnant women, primary infection can cause congenital toxoplasmosis resulting potentially in mental retardation and blindness in the infant [19]. In 1999, Dunn et al. estimated that the overall maternal–fetal transmission rate during pregnancy was 29% [20]. In 2012, Wallon et al. estimated the overall transmission rate at 23·9%, from 2·6% at 3 weeks to 69·8% at 39 weeks of pregnancy [21]. In 1978, public health authorities in France implemented a congenital toxoplasmosis prevention programme. Since 1992, this programme has included a monthly serological screening of seronegative pregnant women, from diagnosis of pregnancy until delivery. In addition, health authorities recommend informing all pregnant women without immunity about risk factors and hygienic and dietetic preventive measures [22, 23]. Seroconversions during pregnancy are considered as active infection with an increased risk of toxoplasmosis for the fetus. Antibiotic treatment is prescribed to reduce the risk of mother-to-child transmission and to improve infant outcomes. Then ante-natal diagnosis is carried out by amniocentesis to confirm the transmission and to adapt the treatment [24, 25].

Prevalence and incidence data are necessary for estimating the burden of the disease and for evaluating the impact of the public health prevention programmes. Incidence of Toxoplasma infection is the number of primary infections/1000 susceptible women per year. It can be estimated directly by conducting a follow-up of a cohort of seronegative individuals but this is expensive and difficult to conduct with large populations. The current screening programme in France does not produce systematic surveillance data on incidence and prevalence of Toxoplasma infection. Therefore, incidence estimation from prevalence data obtained from seroprevalence surveys can be useful in this context. In 1986, Papoz et al. developed a simple mathematical model for this purpose [26]. Estimates of incidence modelled by age- and time-specific Toxoplasma infection seroprevalence data using a catalytic model have been described more recently [27]. The aim of this study was to estimate the incidence and prevalence of Toxoplasma infection in women of childbearing age from seroprevalence surveys, and to estimate the incidence of primary infections during pregnancy in France.

METHODS

Data sources

National Perinatal Survey (NPS)

The NPS is a routine cross-sectional survey based on a representative sample of births in France. This survey is coordinated at the national level by the Ministry of Health and the French National Institute of Health and Medical Research (Inserm). The objectives of this survey are to monitor main health indicators for pregnant women, to evaluate medical practices during pregnancy and delivery, and to identify perinatal risk factors [28].

This survey included all pregnant women admitted to public and private obstetrics, gynaecology and surgery departments to end the pregnancy in France, during a given week. Mothers who delivered outside these hospital departments, at home or elsewhere,
and were then transferred to an obstetrics ward were also included.

Data were collected using a standardized questionnaire administered individually to each delivering mother and completed with data from their medical file record. In the event of multiple births, only a single questionnaire was included in the analysis. This questionnaire collected information regarding sociodemographic characteristics, prenatal care, health status, serological screening results for toxoplasmosis, rubella and syphilis, and potential treatments received for these diseases during pregnancy. Information about serological tests used was not collected as part of the survey.

Three NPS have been conducted in France in 1995, 2003 and 2010 including 13459, 15108 and 15432 women, respectively. The overall seroprevalence of Toxoplasma infection in women of childbearing age was 54% in 1995, 44% in 2003 and 37% in 2010 [14, 17, 18].

**National surveillance of congenital toxoplasmosis (ToxoSurv)**

In 2007, a surveillance system was set up in France to collect information on cases of congenital toxoplasmosis diagnosed during pregnancy by amniocentesis, or diagnosed in newborns and infants aged <1 year whose mother had a primary infection with *T. gondii* during pregnancy. This surveillance system is based on a network of 35 specialized laboratories that are certified in prenatal and neonatal diagnosis of toxoplasmosis and 16 additional medical biology laboratories that occasionally perform this diagnosis. The National Reference Centre for Toxoplasmosis collects data on all cases of congenital toxoplasmosis reported in France. These cases are registered in a web-based system, i.e. ToxoSurv [29, 30].

**Case definitions**

We defined a seropositive woman as one whose last serological test indicated the presence of *Toxoplasma*-specific antibodies (IgG and/or IgM). We considered a woman seronegative if no anti-*Toxoplasma* antibodies were indicated in her medical records at the last available test during her pregnancy. Seroprevalence was defined as the proportion of seropositive women in all women included in the survey with serological information available. Incidence is defined as the number of seroconversions/1000 susceptible women per year. We defined a seroconversion during pregnancy as the presence of IgG- and/or IgM-specific antibodies for toxoplasmosis after negative tests during pregnancy with dates available for the last negative test and the first positive test.

**Study**

We estimated both incidence and prevalence of *Toxoplasma* infection in women of childbearing age by age and over time with a model-based approach using the dataset obtained by merging the data of the three NPS conducted in 1995, 2003 and 2010.

To estimate the incidence of seroconversion during pregnancy in 2010 we also used data on seroconversion obtained from the NPS conducted in 2010 and from ToxoSurv. Additionally, we estimated the number of seroconversions by applying the model estimates to the number of births registered in France in 2010.

**Estimation of incidence in women of childbearing age: model-based estimation**

We estimated the incidence of primary *Toxoplasma* infection using age- and time-specific seroprevalence data obtained from the three NPS conducted in 1995, 2003 and 2010, following a catalytic epidemic parametric model proposed by Ades & Nokes [27].

The relationship between age- and time-specific seroprevalence data $p(a, t)$ and the incidence $I(a, t)$ for a woman of age $a$ at time $t$ is

$$p(a, t) = 1 - \exp\left[-\int_0^a I(u, t-a+u)\,du\right]. \quad (1)$$

We assumed that the incidence $I(a, t)$ depends both on age $a$ and calendar time $t$. It can be expressed as the product of two unknown functions: $I_A(a)$ depending only on age and $I_T(t)$ depending only on time, $I(a, t) = I_A(a) \cdot I_T(t)$. We used five modelling steps.

**Step 1: Building age- and time-specific groups**

We built $K$ age- and time-specific groups by combining all women aged 15–45 years from the three NPS conducted in 1995, 2003 and 2010 with age and serology available. We grouped data into $K=18$ age- and time-specific groups: six age groups ($<20, 20–24, 25–29, 30–34, 35–39, \geq 40$ years) for the three years of the surveys (Table 1). In each group we noted $t_k$ the calendar time of the survey, $N_k$ the number of sampled women, $N_k^+$ the number of seropositive women and $a_k$ the calculated mean age.
As incidence is expressed as the product of two independent functions, from equation (1) we derived the estimated prevalence, noted as \( \hat{p}_k \), in group \( k \).

\[
\hat{p}_k = \hat{p}_k(a_k, t_k) = 1 - \exp \left[ - \int_0^{a_k} I_A(u) \cdot I_T(t_k - a_k + u) \, du \right].
\]  

\( (2) \)

**Step 2: Expressing the incidence**

We chose several parametric forms to the two functions \( I_A(a) \) and \( I_T(t) \) to express the incidence. All of these functions have an exponential polynomial form to ensure positive values and flexible shapes.

**Step 3: Writing a deviance function**

We assumed that the probability of being infected with \( T. gondii \) for a woman is independent from the probability of being infected for another woman, and we considered that the number of seropositive women follows a binomial distribution. Hence, the binomial \( \chi^2 \) deviance function is:

\[
\text{Dev}(\theta) = 2 \sum_{k=1}^{K} \left\{ N_k^+ \ln \left( \frac{N_k^+}{N_k \hat{p}_k} \right) + (N_k - N_k^+) \ln \left( \frac{N_k - N_k^+}{N_k (1 - \hat{p}_k)} \right) \right\}.
\]  

\( (3) \)

where \( \theta \) is the vector of parameters involved in the incidence function.

**Step 4: Obtaining the parameters**

We estimated the parameters of the incidence functions by minimizing the deviance function [equation (3)]. The associated number of degrees of freedom (D.F.) was equal to 18 (\( K \) age- and time-specific groups built in step 1 minus the number of parameters involved in the incidence function).

**Step 5: Fitting the catalytic model**

We selected the incidence function fitting significantly better than the other functions. We preferred an incidence function \( \mathcal{A} \) (\( p \) parameters, deviance \( D_A \)) over an incidence function \( \mathcal{B} \) (\( q \) parameters, deviance \( D_B \)) if the \( P \) value of a \( \chi^2 \) with \( n_B - n_A \) D.F. was <0.05. The 95% confidence interval (CI) associated with the incidence was derived from the estimated incidence under a binomial distribution assumption.

**Estimation of prevalence in women of childbearing age: model-based estimation**

We estimated the prevalence of \( T. gondii \) infection by age (20–40 years) from 1980 to 2020 from equation (1) using the parameters estimated in step 4. We compared prevalence obtained through the model with prevalence data observed in the NPS and other seroprevalence studies conducted in France.

**Estimation of incidence of seroconversion in pregnant women in 2010**

We used three different approaches to estimate the number of seroconversions in pregnant women in 2010. First, we obtained the number of seroconversions reported in the NPS conducted in 2010. Second, to estimate the number of seroconversions during pregnancy, we divided the number of cases of congenital toxoplasmosis reported through ToxoSurv...
in 2010 by the maternal–fetal transmission rate during pregnancy [21]. We used the number of births from seronegative women registered in 2010 in France to calculate the incidence of seroconversion [31]. Third, we multiplied the estimated incidence (obtained from the catalytic model) by the number of susceptible pregnant women (estimated from the prevalence observed in the NPS multiplied by the number of births).

We used R 2.13.0 [32] to model age- and time-specific incidence and prevalence in women of childbearing age and estimated the incidence of seroconversion in pregnant women using Microsoft Excel.

RESULTS

Estimation of incidence in women of childbearing age: model-based estimation

We combined data of 42,208 women aged 15–45 years with age and serology available from the NPS conducted in 1995, 2003 and 2010 (Table 1). We estimated parameters that minimized the binomial log-likelihood \( \chi^2 \) deviance function using different polynomial exponential functions (Table 2).

We selected incidence function 5, which fitted significantly better than others. Thus the estimated incidence for a seronegative woman of age \( a \) at time \( t \) was expressed by:

\[
I(a, t) = \exp(72.327 - 0.0381t - 0.0621a).
\]

The estimated incidence of primary Toxoplasma infection in women aged 30 years decreased from 7.5/1000 susceptible individuals in 1980 to 3.5/1000 in 2000 (Figs 1 and 2). In 2010, the incidence was estimated at 2.4/1000 susceptible women. If the current trend continues, it is predicted that in 2020 the incidence will decline to 1.6 cases/1000 in women aged 30 years. The incidence of primary Toxoplasma infection in women aged 20 years was 3.4 times higher than in those aged 40 years (Table 3).

Prevalence estimation in women of childbearing age: model-based estimation

The estimated prevalence obtained from modelling indicated that 76.3% of women aged 30 years were seropositive in 1980, 48.9% in 2000, and is predicted to be 26.9% in 2020 (Table 3). The estimated prevalence increased with age and decreased by time. This is precisely the pattern observed in the data obtained from the three NPS (Fig. 3).

Estimation of the incidence of seroconversion in pregnant women from observed data

Seroconversions reported in the NPS

We identified 20 women with a seroconversion during their pregnancy in the NPS conducted in 2010. Considering a stable number of seroconversions over the 52 weeks of the year, we estimated the incidence of seroconversion at 2.1 (95% CI 1.3–3.1)/1000 pregnancies at risk, and the number of women with a primary infection during their pregnancy in 2010 at 1040 (95% CI 643–1535).

Seroconversions estimated from the cases of congenital toxoplasmosis reported to ToxoSurv

We estimated the number of seroconversions during pregnancy from the number of cases of congenital toxoplasmosis reported to ToxoSurv divided by the maternal–fetal transmission rate 23.9% (95% CI 20.5–27.4).

Table 2. Parametric incidence functions with their corresponding estimated parameters, deviance and degrees of freedom

<table>
<thead>
<tr>
<th>Incidence function</th>
<th>Parameters</th>
<th>Deviance</th>
<th>D.F.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( I(a,t) = \exp(\theta_1) )</td>
<td>(-3.898)</td>
<td>1064.8</td>
<td>17</td>
</tr>
<tr>
<td>2. ( I(a,t) = \exp(\theta_1 + \theta_2a) )</td>
<td>(-4.113)</td>
<td>1039.7</td>
<td>16</td>
</tr>
<tr>
<td>3. ( I(a,t) = \exp(\theta_1 + \theta_2a + \theta_3a^2) )</td>
<td>(-4.091)</td>
<td>1039.7</td>
<td>15</td>
</tr>
<tr>
<td>4. ( I(a,t) = \exp(\theta_1 + \theta_2t) )</td>
<td>(63.538)</td>
<td>132.3</td>
<td>16</td>
</tr>
<tr>
<td>5. ( I(a,t) = \exp(\theta_1 + \theta_2t + \theta_3a) )</td>
<td>(72.327)</td>
<td>73.8</td>
<td>15</td>
</tr>
</tbody>
</table>

* D.F., degrees of freedom = 18 specific groups – number of parameters.
In 2010, 244 cases of congenital toxoplasmosis were reported through ToxoSurv [33]. We calculated 1021 (95% CI 938–1119) seroconversions in pregnant women in 2010. The incidence rate was 1·9 (95% CI 1·8–2·1)/1000 pregnant women per year.

Fig. 1. Estimated *Toxoplasma* infection incidence by year for women aged 20, 25, 30, 35, and 40 years, France, 1980–2020.

Fig. 2. Estimated *Toxoplasma* infection incidence by age and year, France, 1980–2020. Solid, dashed and long dashed lines represent the estimated incidences obtained from the model in 1995, 2003 and 2010.
Estimation of the incidence of seroconversion in pregnant women from incidence modelling

Besides these two approaches, we performed an additional estimation of seroconversion during pregnancy using the incidence estimates obtained from the model and applied to the number of births registered in France in 2010.

For each age from 15 to 45 years, we calculated the number of susceptible pregnant women by multiplying the prevalence in the 2010 NPS with the number of births \( (n = 841563) \) registered in France in 2010. We applied the incidence estimates obtained from the model to the susceptible women in 2010 to calculate the predicted number of seroconversions during pregnancy by age. We estimated 1319 (95% CI 971–1668) seroconversions in pregnant women in 2010, corresponding to an incidence of 2·5 (95% CI 1·9–3·2)/1000 susceptible pregnant women.

DISCUSSION

We estimated incidence of *Toxoplasma* infection in women by modelling seroprevalence data from repeated NPS. Using this model we also estimated the predicted prevalence by age from 1980 to 2020. These surveys give valid estimates of prevalence and produced useful surveillance data for tracking *T. gondii* prevalence over time in women of childbearing age in France. In 15 years, *Toxoplasma* infection prevalence in women of childbearing age decreased markedly from 54% in 1995 to 37% in 2010 [17, 18].

The catalytic epidemic model developed in this study suggested that incidence of primary *Toxoplasma* infection in women aged 15–45 years has decreased by 70% over the last 30 years. Based on this model, and considering the same trend and conditions, we estimated that the incidence may continue to decline over the following years. This reduction may be explained by a lower exposure to the parasite by changes in food habits and improved hygiene practices in meat production. Ovine meat consumption is an important cause of *Toxoplasma* infection in France [17, 18, 34, 35]. The 30% decline in ovine meat consumption between 2000 and 2010 reported by the French Ministry of Agriculture and Food may have contributed to this reduction [36].

Data obtained from other studies conducted previously in France were consistent with the prevalence data estimated by the model. In 1982, a study based on sera collected from a sample of pregnant women in France found a prevalence of 57% in women aged 20 years, 68% in women aged 30 years and 86% in those aged 40 years [26]. The modelled prevalence for 1982 was 56%, 74%, and 87% for the respective ages. Long-term predictions suggest that prevalence will continue to decline to 18%, 27% and 38%, respectively, for these ages in 2020.

Estimations of the incidence of seroconversion in pregnant women derived from the NPS and the surveillance system for congenital toxoplasmosis lead to similar rates of seroconversion during pregnancy in 2010 (2·1/1000 and 1·9/1000 susceptible pregnant women, respectively). The estimated number of seroconversions that used incidence estimations in women of childbearing age obtained from the model was higher (2·5/1000 susceptible women). This suggests that women of childbearing age have a higher risk of being infected by *T. gondii* than pregnant women. This may be explained by the preventive effect of the hygiene messages given by the health authorities to susceptible women during pregnancy.

The number of primary infections with *T. gondii* during pregnancy reported in the NPS decreased by 41% from 5·4/1000 susceptible women in 1995 to 2·1/1000 in 2010. This decrease may reflect the overall reduction in incidence due to less frequent exposure to the parasite. The number of congenital toxoplasmosis

<table>
<thead>
<tr>
<th>Age (yr)</th>
<th>Incidence (per 1000)</th>
<th>Seroprevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>14</td>
<td>9·5</td>
</tr>
<tr>
<td>25</td>
<td>10·2</td>
<td>7</td>
</tr>
<tr>
<td>30</td>
<td>7·5</td>
<td>5·1</td>
</tr>
<tr>
<td>35</td>
<td>5·5</td>
<td>3·8</td>
</tr>
<tr>
<td>40</td>
<td>4</td>
<td>2·8</td>
</tr>
</tbody>
</table>

Table 3. Estimated incidence and prevalence of *Toxoplasma gondii* infection in women by age and year, France, 1980–2020
cases reported to ToxoSurv has also fallen since 2007 [33].

Our catalytic epidemic model considered that the incidence of primary *Toxoplasma* infection depended both on age and calendar time. Models previously developed considered incidence as a function that depended only on age [26]. The catalytic epidemic approach has also been applied for incidence estimation for several other diseases including hepatitis A, hepatitis B, measles, mumps and rubella, congenital rubella syndrome, and cytomegalovirus infection [37–42]. To our knowledge, no studies on incidence estimation of *Toxoplasma* infection have been conducted recently in Europe. We considered our model valid for a number of reasons. First, our modelled prevalence fitted to the prevalence observed in the three NPS conducted in 1995, 2003 and 2010, especially for women aged between 25 and 35 years for which the estimates are more precise due to the larger number of women. In addition, prevalence obtained from other studies in France also approximates to the prevalence estimated by our model [15, 16, 26]. Second, the estimated number of seroconversions during pregnancy, obtained by applying the estimated incidence in women of childbearing age modelled by age and time to the number of births in France, led to incidence rates slightly higher than those observed in the NPS and those estimated from the ToxoSurv surveillance system, which was to be expected due to the positive impact of the preventive messages to susceptible pregnant women. However, our model could be improved to reduce the residual deviance by including an interaction term between age and time. A potential overdispersion due to the binomial distribution hypothesis could also be tested.

In conclusion, catalytic epidemic modelling developed from age- and time-specific prevalence data can be used to obtain reliable estimates of incidence and prevalence of *Toxoplasma* infection in women over time. This information is essential for epidemiologists and health economists who estimate the impact and the cost-effectiveness of preventive programmes for toxoplasmosis.

**ACKNOWLEDGEMENTS**

The authors acknowledge B. Blondel and M. Kermarrec of the French National Institute of Health and Medical Research (Inserm), U.953, for providing data from the National Perinatal Surveys. In addition, the authors acknowledge M. Wallon and C. Binquet for providing the mother–child transmission rate, and the Toxosurv network members for
registering the cases of congenital toxoplasmosis in the surveillance system.

No funding was received for the work presented.

DECLARATION OF INTEREST

None.

REFERENCES