A theoretical estimate of the risk of microcephaly during pregnancy with Zika virus infection

Hiroshi Nishiura\textsuperscript{a,b,d,*}, Kenji Mizumoto\textsuperscript{a,c,d}, Kat S. Rock\textsuperscript{a}, Yohei Yasuda\textsuperscript{a,b}, Ryo Kinoshita\textsuperscript{a,b,d}, Yuichiro Miyamatsu\textsuperscript{a,b,d}

\textsuperscript{a} Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 1130033, Japan
\textsuperscript{b} CREST, Japan Science and Technology Agency, Honcho 4-1-8, Kawaguchi, Saitama 332-0012, Japan
\textsuperscript{c} Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 1538902, Japan
\textsuperscript{d} Graduate School of Medicine, Hokkaido University, Kita 15 Jo Nishi 7 Chome, Kita-ku, Sapporo-shi, Hokkaido 060-8638, Japan

\section*{A R T I C L E   I N F O}
Article history:
Received 22 February 2016
Received in revised form 5 March 2016
Accepted 7 March 2016
Available online 18 March 2016

Keywords:
Microcephaly
Zika infection
Gestation
Statistical estimation
Brazil

\section*{A B S T R A C T}
Objectives: There has been a growing concern over Zika virus (ZIKV) infection, particularly since a probable link between ZIKV infection during pregnancy and microcephaly in the baby was identified. The present study aimed to estimate a theoretical risk of microcephaly during pregnancy with ZIKV infection in Northeastern Brazil in 2015.

Methods: Temporal distributions of microcephaly, reported dengue-like illness and dengue seropositivity in Brazil were extracted from secondary data sources. Using an integral equation model and a backcalculation technique, we estimated the risk of microcephaly during pregnancy with Zika virus infection.

Results: If the fraction of Zika virus infections among a total of seronegative dengue-like illness cases is 30%, the risk of microcephaly following infection during the first trimester was estimated at 46.7% (95% CI: 8.1, 84.2), comparable to the risk of congenital rubella syndrome. However, the risk of microcephaly was shown to vary widely from 14.0% to 100%. The mean gestational age at delivery with microcephaly was estimated at 37.5 weeks (95% CI: 36.9, 39.3).

Conclusions: The time interval between peaks of reported dengue-like illness and microcephaly was consistent with cause-outcome relationship. Our modeling framework predicts that the incidence of microcephaly is expected to steadily decline in early 2016, Brazil.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

\section*{1. Introduction}
There has been a growing concern over Zika virus (ZIKV) infection (Musso et al., 2015; Ventura et al., 2016), particularly since a probable link between ZIKV infection during pregnancy and microcephaly was identified (Schuler-Faccini et al., 2016). Clinical symptoms of ZIKV are self-limiting in general and the infection is sometimes even sub-clinical, but severe complications including microcephaly and Guillain–Barré syndrome justify global effort to actively monitor and control the spread of this disease. The transmission potential of Zika virus infection in the South Pacific has been shown to be comparable to dengue and chikungunya viruses (Nishiura et al., 2016a).

Brazil has experienced a large ZIKV epidemic with the notification of a substantial number of microcephaly cases, which have increased almost 20-fold compared to the recent years. In particular, Northeastern Brazil has notified more than 85% of all microcephaly cases that have been reported in Brazil, 2015 as of the end of 2015. Considering that the risk of microcephaly is a substantial public health concern, it is fruitful to quantify the risk of contracting microcephaly given ZIKV infection during early gestational period of pregnancy. Elucidating the epidemiological mechanism of microcephaly could also shed light on future course of the ongoing epidemic.

The present study estimated the risk of microcephaly during pregnancy with ZIKV infection, analyzing epidemiological datasets of reported dengue-like illness and microcephaly in Northeastern Brazil in 2015.

* Corresponding author at: Graduate School of Medicine, Hokkaido University, Kita 15 Jo Nishi 7 Chome, Kita-ku, Sapporo-shi, Hokkaido 060-8638, Japan.
Fax: +81 11 706 7819.
E-mail address: nishuah@gmail.com (H. Nishiura).

http://dx.doi.org/10.1016/j.epidem.2016.03.001
1755-4365/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
2. Materials and methods

2.1. Epidemiological data

We analyzed the epidemiological datasets of microcephaly and dengue-like illness, retrieved from publicly available secondary data sources. Both of them are reported as the cumulative number, and we focused on the dataset from Northeastern Brazil (Ministério da Saúde, 2016a). Microcephaly is defined as a head circumference of more than 2 SDs below the mean for age and gender, and can be caused by a variety of reasons including infection with Zika virus (suspected), syphilis, toxoplasmosis, rubella, cytomegalovirus and other infectious agents. Dengue-like illness is collected as a part of syndromic surveillance (Ministério da Saúde, 2016b), and is clinically defined as a patient who has acute febrile illness accompanied by at least two of the following symptoms: headaches, retro-orbital pain, myalgia, arthralgia, prostration and rash (Ministério da Saúde, 2009). The latest time at which the data were collated for the study was 20th February, 2016. Accessible data of microcephaly have been available from week 47 of 2015 to 6 weeks (Ministério da Saúde, 2016a). Due to increased awareness, the reporting coverage is likely to have been considerably elevated from 2016 and the validity of microcephaly diagnoses has been questioned elsewhere (Victora et al., 2016). For this reason, we focused on microcephaly data within 2015 only.

In addition, the following three pieces of information were retrieved. First, the number of serological samples and the outcome of serological testing for dengue virus over time were also collected, focusing on Northeastern Brazil (Ministério da Saúde, 2016b). Secondly, a point estimate of the proportion of pregnant women among a total population was calculated using Brazilian census (United Nations Population Division, 2016). Third, gestational age distribution of delivery in Brazil was extracted to calculate the variance of the time from pregnancy to delivery (Pereira et al., 2013). Due to scarcity of the dataset, gestational ages of 22–29 weeks and 43 weeks and longer were omitted from the distribution during the implementation of statistical estimation.

2.2. Statistical analysis

Using the following integral equation model, we estimate, $\pi$, the conditional risk of microcephaly given ZIKV infection during the first trimester of pregnancy. Using the weekly incidence of dengue-like illness, $c_{t-s+a}$, which occurred between $a_1$ and $a_2$ weeks of gestation (i.e., between times $t-s+a_1$ and $t-s+a_2$), the expected number of microcephaly, $E(m_t)$, at time $t$, is written as

$$E(m_t) = \frac{\pi z b}{q} \sum_{s=1}^{a_2} \sum_{a=a_1}^{a_2} (1-p_{t-s+a})^c_c t_{t-s+a} \delta_2(\mu, \sigma^2) ,$$

where $p_t$ is the time-dependent proportion of seropositive for dengue virus infection among all serological samples, and $q$ is the proportion of ZIKV cases who sought medical treatment. $z$ represents the fraction of ZIKV infections among seronegative dengue-like illness cases. $b$ is a point estimate of the proportion of pregnant women among a total population. $r$ is the proportion of actual microcephaly cases among all notified cases of microcephaly. $g_w$ is the frequency of successful infection in fetus given an exposure at gestational age $a$, and $f_i$ is the probability density function of the time from pregnancy to delivery. $g_w$ was assumed be uniformly distributed during the first trimester of pregnancy, because this is the period when there appears to be a high risk of microcephaly (de Paula Freitas et al., 2016). Different sources define different lengths of the time for the first trimester ranging from 12 to 16 weeks, however the default parameterization was taken to be $a_1 = 1$ and $a_2 = 12$. In a sensitivity analysis, it was assumed that $a_2$ ranges from 12 to 16. $f_i$ was assumed to follow a gamma distribution as the visual assessment of fit to term delivery data was satisfactory. Since the variance of gestational age at delivery appeared not to be deviated from term delivery in Brazil (De Araujo et al., 2016), and thus, the variance of $f_i$ was fixed at 3.78 weeks$^2$ (Pereira et al., 2013). However, considering that microcephaly could lead to delivery at an early gestational age, the mean ($\mu$) was jointly estimated with other parameters.

Tables 1 and 2 list estimated and known parameters and variables, respectively. This formulation is close to the idea used for congenital rubella syndrome (Gao et al., 2013; Cutts and Vynnycky, 1999), calculating the risk of fetus infection using epidemiological data and risk parameters governing infection in pregnant women.

A maximum likelihood method was employed to estimate unknown parameters, $\pi, p_t, q, r$ and $\mu$ (among which $p_t$ was dealt with as varying with time). Assuming that the weekly count of microcephaly cases follows a Poisson distribution, the corresponding likelihood function to estimate parameters based on the dataset from week 47 to 52 is

$$L_1(\pi, p_t, q, r, \mu; c_t, x_t) = \prod_{t=47}^{52} E(m_t; c_t)^{x_t} \exp\left(-E(m_t; c_t)\right),$$

where $x_t$ is the observed incidence of microcephaly in week $t$.

The conditional risk of medical attendance given ZIKV infection, $q$ was assumed to follow a binomial distribution. Supposing that $l$ cases attended clinics among a total of $k$ estimated number of infections, the likelihood function to estimate $q$ is

$$L_2(q; k, l) = \binom{k}{l} q^l (1-q)^{k-l},$$

Since there was no dataset that permitted us to explicitly estimate $q$ in Brazil, we used Yap Island seroepidemiological data in 2007 and compensated the estimate from Eq. (3) into the information as a part of the right hand side of Eq. (1).

In addition to the syndromic data, serological testing has been performed for a small fraction of dengue-like illness cases over time. We assume that empirically examined seroepidemiological samples are representative of all reported dengue-like illness cases. The probability of seropositive becoming a confirmed dengue case given dengue-like illness status (i.e., given clinical/syndromic diagnosis) was also similarly calculated from binomial distribution. Assuming that there were $i_t$ positive samples among a total of $h_t$ serological samples, we have

$$L_3(p_t, h_t, i_t) = \prod_{t} \binom{h_t}{i_t} p_t^{i_t} (1-p_t)^{h_t-i_t},$$

Similarly, referring to an epidemiological study that assessed diagnostic accuracy of reported microcephaly (and emphasized that more than half of reported microcephaly cases did not satisfy the clinical criteria of microcephaly) (Victora et al., 2016), a binomial sampling process of actual microcephaly cases $b$ given a total of $c$ reported microcephaly cases was modeled as

$$L_4(\pi ; b, c) = \binom{c}{b} p^b (1-p)^{c-b},$$

Thus, the total likelihood is calculated as $L = L_1 L_2 L_3 L_4$. The maximum likelihood estimates of $\pi, p_t, q, r$ and $\mu$ were identified by minimizing the negative log-likelihood. The 95% confidence interval (CI) was computed by using the profile likelihood. The parameter $z$ was not estimated from empirical data due to its correlation with $\pi$. Instead, we varied $z$ from 0 to 1.0, examining the sensitivity of $\pi$ to the value of $z$. 

H. Nishiura et al. / Epidemics 15 (2016) 66–70

67
Table 1
Estimated parameter values of Zika virus infection model.

<table>
<thead>
<tr>
<th>Parameter's interpretation</th>
<th>Notation</th>
<th>Mean (%)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower bound of the conditional risk of microcephaly given ZIKV infection during first trimester of pregnancy (when z = 1)</td>
<td>π</td>
<td>14.0</td>
<td>2.7, 25.3</td>
</tr>
<tr>
<td>Proportion of seropositive for dengue virus infection among all serological samples given dengue-like illness</td>
<td>π</td>
<td>See below</td>
<td></td>
</tr>
<tr>
<td>Proportion of ZIKV cases with medical attendance among the total of infected individuals</td>
<td>p</td>
<td>3.4</td>
<td>3.1, 3.6</td>
</tr>
<tr>
<td>Proportion of actual microcephaly cases among a total of reported microcephaly cases</td>
<td>μ</td>
<td>37.5</td>
<td>36.9, 39.3</td>
</tr>
<tr>
<td>Mean gestational age from pregnancy to delivery (weeks)</td>
<td>h</td>
<td>36.6</td>
<td>35.4, 37.9</td>
</tr>
</tbody>
</table>

* CI, confidence interval (derived from the profile likelihood).

Table 2
Variables and parameters of Zika virus infection model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Interpretation</th>
<th>Source/assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>c_t</td>
<td>Weekly incidence of dengue-like illness in week t</td>
<td>Derived from data (Ministério da Saúde, 2016b)</td>
</tr>
<tr>
<td>m_t</td>
<td>Weekly incidence of microcephaly in week t</td>
<td>Derived from data (Ministério da Saúde, 2016a)</td>
</tr>
<tr>
<td>h_t</td>
<td>Serological sample size in week t</td>
<td></td>
</tr>
<tr>
<td>p_t</td>
<td>Number of dengue positive tests in week t</td>
<td></td>
</tr>
<tr>
<td>m_p</td>
<td>Model prediction, microcephaly in week t</td>
<td></td>
</tr>
<tr>
<td>g_a</td>
<td>Frequency of successful infection in fetus given an exposure at gestational age a</td>
<td>Assumed to be uniformly distributed U(0, 12)</td>
</tr>
<tr>
<td>f_b</td>
<td>Probability density function of the time from pregnancy to delivery</td>
<td>Assumed to be gamma distributed (μ = estimated σ = 1.94)</td>
</tr>
<tr>
<td>b</td>
<td>Point estimate of the proportion of pregnant women among a total population in Brazil</td>
<td>0.36% calculated (United Nations Population Division, 2016)</td>
</tr>
</tbody>
</table>

3. Results

Northeastern Brazil was most severely affected by ZIKV infection (Table 3), as it is observed from microcephaly notifications, and in this Federal Unit, the fraction of dengue seropositive has been as low as 17.1% (Secretaria de Vigilância em Saúde, 2016), implying that reported dengue-like illness cases were contaminated by substantial number of ZIKV infections (Nishiura et al., 2016b).

In fact, a strong positive correlation can be identified between microcephaly incidence and estimated dengue seronegative cases, which were originally notified as dengue-like illness (Fig. 1A; Pearson’s r = 0.788, p = 0.012). Assuming that these dengue seronegative cases represented ZIKV infections and using a family of backcalculation techniques, we estimated the risk of microcephaly given infection during the first trimester of pregnancy. Since the fraction of infections leading to medical attendance is not very well known in Brazil, the risk of medical attendance given infection derived from Yap island epidemic data in 2007 was jointly estimated (Duffy et al., 2009).

Fig. 1B compares the observed and estimated weekly incidence of microcephaly, visually capturing the pattern of microcephaly notifications which were considered to have taken place in week 46, a few weeks before first recognition of clinical link between microcephaly and ZIKV infection. The proportion of medical attendance among a total of ZIKV infections was estimated at 3.4% (95% confidence intervals [CI]: 3.1, 3.6). Assuming that z = 1, the lowest bound of the risk of microcephaly following infection during the first trimester was estimated at 14.0% (95% CI: 2.7, 25.3). The mean gestational age at delivery with microcephaly was estimated at 37.5 weeks (95% CI: 36.9, 39.3).

Fig. 2 shows the time-dependent variations in the proportion of seropositive (true positive) dengue-like illness cases in Northeastern Brazil. Overlaying it with the estimated Zika virus infection curve in Bahia (one of Federal State of Northeastern region) elsewhere (Rodrigues Faria et al., 2016), two interesting features were identified: (I) the valley of the estimated proportion seropositive coincides with the timing of peak of estimated Zika virus infection in Rodrigues Faria et al. (2016) ensuring that low seropositive fraction reflects high contamination of Zika virus infections, while (ii) the estimated proportion in earlier part of 2015 was also low, implying that the estimated Zika virus cases in the corresponding period was considerably underestimated (perhaps due to under-ascertainment) in Rodrigues Faria et al. (2016). The maximum bound of the fraction of ZIKV infections among reported dengue-like illness was 96.1% (95% CI: 94.6, 97.6) from week 21 to 30; the estimate is regarded as maximum, because seronegative cases include not only ZIKV infections but also many other causes including (undetected) dengue and chikungunya virus infections.

In a sensitivity analysis with an increased “at risk” period during pregnancy, it was found that r decreased from 14.0% to 10.5% by varying the last week of first trimester from week 12 to 16.

Fig. 3 shows the estimated risk of microcephaly as a function of r, the proportion of ZIKV infections among a total of seronegative dengue infections.
dengue-like illness cases. If $r$ is 0.5, namely, half of seronegative dengue-like illness cases are caused by ZIKV, the risk of microcephaly during pregnancy would be 28.0%. If ZIKV infections account for only 30% of the seronegative dengue-like illness cases, the risk of microcephaly is 46.7% (95% CI: 9.1, 84.2), comparable to empirical estimate of congenital rubella syndrome (Cutts and Vynnycky, 1999).

4. Discussion

The conditional risk of microcephaly given ZIKV infection during the first trimester of pregnancy was estimated, and it was shown that the estimate is variable with the fraction of ZIKV infections that account for seronegative dengue-like illness cases (Fig. 3). The integral equation model was shown to capture the temporal pattern of the microcephaly. Not only did this fit well to the observed data, but also the model predicted the incidence of microcephaly before intensified surveillance for microcephaly began (week 45 and earlier) and offered a short-term prediction of future incidence (week 5, 2016 and later). The model reasonably yielded predictions based on exposures that had occurred more than 8 months ago. The prediction has indicated that the epidemic of microcephaly in Brazil, 2016 is likely to steadily decline soon in the future course. If the fraction ZIKV infections among seronegative dengue-like illness cases is as small as 30%, the risk of microcephaly during pregnancy with ZIKV infection would be 46.7%, comparable to the risk of congenital rubella syndrome (Cutts and Vynnycky, 1999).

The estimate is likely useful for a variety of epidemiological purposes. First, ZIKV infection involves substantial number of asymptomatic and mild cases, and the estimated risk enables us to calculate the total number of infections based on microcephaly data. Second, using the backcalculation technique, a short-term prediction of microcephaly occurrence can be achieved.

---

Fig. 1. Reported dengue-like illness and microcephaly in Northeastern Brazil, 2015. (A) Scatter plot of microcephaly against estimated seronegative cases of reported dengue-like illness at nine different states in Northeastern Brazil. Box-Cox transformation was employed to ensure the normality of both datasets. A positive correlation (Pearson’s $r = 0.788$, $p = 0.012$) was identified. Seronegative cases of reported dengue-like illness are considered to mirror ZIKV infections. (B) Temporal distributions of reported dengue-like illness and microcephaly in Northeastern Brazil in 2015. The peak of microcephaly is considered to be 30 weeks apart from the peak of reported dengue-like illness. Predicted microcephaly was derived from the solution of backcalculation method, which was applied to the observed microcephaly data by the end of 2015. Filled circles represent the observed microcephaly data within 2015, while unfilled circles are the reported data in 2016. Due to probable increased awareness, we have used the dataset in 2015 alone to estimate parameters, but the prediction is shown to be aligned with the observed data in 2016.

Fig. 2. Time dependent proportion seropositive for dengue among dengue-like illness cases, Northeastern Brazil, 2015. Estimated proportions seropositive for dengue among all reported dengue-like illness cases are evaluated for every 10 weeks, 2015. Dashed line represents the lower and upper 95% confidence intervals. Filled circles show the estimated confirmed number of Zika virus cases in Bahia, one Federal State that belongs to Northeastern Brazil, as estimated elsewhere (Rodrigues Faria et al., 2016). The epidemic peak is coincided with the minimum estimate of dengue seroprevalence.

Fig. 3. Sensitivity of the risk of microcephaly during pregnancy with Zika virus infection to the proportion of Zika virus infections among seronegative dengue-like illness cases. Both vertical and horizontal axes show the estimated percentages. Solid line represents maximum likelihood estimate, while lower and upper dashed lines represent 95% confidence intervals derived from profile likelihood.
For instance, the prediction can show that the increased notification of microcephaly from December 2015 to January 2016 is certainly expected to steadily decline in the near future. Third, once an effective vaccine is developed, the expected impact of introducing ZIKV vaccine into a population could be calculated, explicitly accounting for the disease burden associated with the infection during pregnancy.

Our modeling exercise is not free from limitations, and especially, it must be remembered that the estimation involved a key assumption that a certain fraction of seronegative dengue-like illness cases had been ZIKV infections (Rodrigues Faria et al., 2016). Due to uncertainty associated with this assumption, we have examined the sensitivity of the estimated risk of microcephaly during pregnancy to different values of the fraction ZIKV infections (Fig. 3). Despite this strong assumption, it must be noted that the time interval between peaks of reported dengue-like illness and microcephaly (Fig. 1B) was consistent with cause–outcome relationship, and the estimated mean duration of pregnancy with microcephaly (37.5 weeks) was only slightly shorter than our knowledge of gestational age at term delivery (Pereira et al., 2013). Moreover, we successfully varied $p_d$, the proportion seropositive among all dengue-like illness cases and the estimated timing of minimum seroprevalence coincided with the timing of Zika virus epidemic peak as reported elsewhere (Rodrigues Faria et al., 2016). Another limitation is the reporting coverage of microcephaly might have been erroneously elevated (Ministério da Saúde, 2009), especially after the declaration of Public Health Emergency of International Concerns on 1 February, 2016. We have fitted our model only by using the dataset in 2015 before considerable increase in awareness. Better estimates for Brazil and other countries at risk from ZIKV infection could be derived from improved surveillance systems with laboratory testing capacity. Moreover, we had to compromise the data gap. The incidence of dengue-like illness was complemented by fitting a spline curve to an observed cumulative number, because the direct calculation of weekly incidence sometimes yielded a negative value.

Despite the existence of reporting bias and a data gap, our modeling study certainly provides predictions which are in agreement with the reported trend of microcephaly in Brazil, and the current epidemic of microcephaly in Brazil is expected to steadily decline in the early part of 2016. Considering the severe disease burden of ZIKV infection, it is highly recommended that intensified surveillance is introduced and that ZIKV infection using serological samples from dengue-like illness is examined for early and accurate detection of possible future ZIKV outbreaks.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgments

HN received funding support from the JSPS KAKENHI Grant Numbers 26670308 and 26700028, Japan Agency for Medical Research and Development, the Japan Science and Technology Agency (JST) CREST program and RISTEX program for Science of Science, Technology and Innovation Policy. KSR was supported by the JSPS Postdoctoral Fellowship for Foreign Researchers. KM received funding support from the Japanese Society for the Promotion of Science (JSPS) KAKENHI Grant Number 15K20936.

References