

EPIDEMIOLOGY

Making high-res Zika maps

Accurate estimates of disease burden are possible by building high-resolution geographical models. However, novel pathogens such as Zika virus pose substantial challenges, requiring both new analytical techniques and, where possible, serological surveys.

Steven Riley

As the Zika virus epidemic continues to unfold across Latin America and the Caribbean, much uncertainty remains about how many people have been infected and how many associated congenital abnormalities have occurred. In this issue of *Nature Microbiology*, Perkins and colleagues¹ present high-resolution geographical maps of expected numbers of infections in the first wave of the Zika epidemic. Aggregating over their maps, they estimate that between 82 and 117 million people will eventually be infected, including between 1.5 and 2.1 million childbearing women. Differences in the numbers predicted for one location over another are driven by known features including human population density, fertility rates, socioeconomic status, and estimated vector abundance. As local, national and global health policy makers weigh the costs and benefits of potential intervention strategies, detailed estimates such as these provide crucial initial reference points.

This work is important because it merges mechanistic models of infectious disease transmission with high-resolution disease risk maps. The relationship between known features of each 5 km by 5 km geographical square and the infection attack rate is fundamental to the study. For other disease systems that have been mapped in this way^{2,3}, high-resolution data were available for both the drivers of transmission and direct measures of infection, leading to accurate infection estimates and an increasing emphasis on the use of spatially driven burden models for public health planning. However, the availability of reliable data from only a single complete Zika virus epidemic (not in the same geographical region)⁴, makes extrapolation far more challenging and necessitates the extension to mechanistic models.

Perkins and colleagues first estimate transmissibility for each small population, before adapting standard theory to define a novel relationship between transmissibility and the expected cumulative attack rate. Transmissibility is measured by the basic

reproductive number R_0 , the average number of new infections caused by one initially infectious individual in an otherwise susceptible population. A classic result for simple epidemic models describes the expected attack rate as a rapidly and monotonically increasing function of R_0 (ref. 5). For example, it predicts attack rates of 94% for $R_0 = 3$ and 98% for $R_0 = 4$. By transforming estimated R_0 values onto lower values with a scaling parameter, Perkins and colleagues imply a more gradual relationship between transmissibility and attack rate in this key range of R_0 for Zika virus.

The concept of a complete first wave of infection is central to the estimates presented. However, a complete first wave doesn't necessarily correspond to a single epidemic. Seasonal variation in transmissibility and chance effects due to the connectedness of distant populations ensure that first waves of infection for any single population might be made up of two or more separate epidemics^{6,7} (Fig. 1). Vector population density is driven by temperature and rainfall, both of which vary substantially throughout the year in Latin America and the Caribbean. Therefore, if

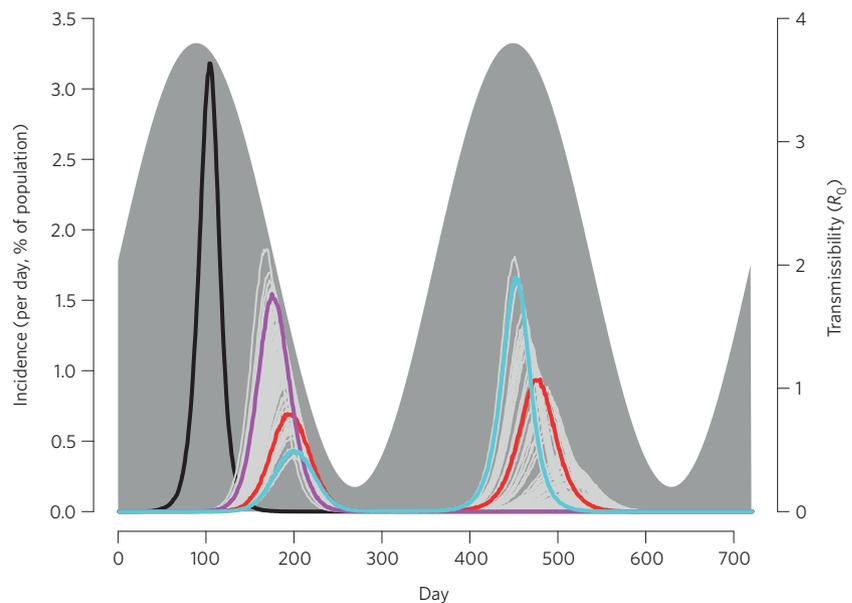


Figure 1 | Strong seasonality may split first waves of Zika virus infection into multiple epidemics. The black line shows a single Zika-like scenario (SEIR model¹² with generation time of 16 days) in a population of one million people with ten initially infectious individuals. Transmissibility shown by the dark shaded areas, peaking at $R_0 = 3.8$ at day 90. The light grey lines show 100 realizations of the same scenario except that the virus only entered the population at an average rate of three infectious individuals per week starting at day 70. I assumed that seasonality was sinusoidal and strong for both scenarios (dark grey shaded area, right-hand axis), dropping to $R_0 = 0.2$ in the low season, which is less severe than recent preliminary estimates of the seasonality of flavivirus transmission in this region¹³. For the second population, with later invasion and low levels of importation of cases, it was possible to have a large epidemic in the first year (magenta) or the second year (cyan) or to have infection split almost evenly between years (red). If reporting rates vary substantially between populations, seasonality and epidemic dynamics could drive observed case numbers in the first year to be similar for populations with very different levels of susceptibility at the start of the second year.

the virus arrives in a given small population early in the season, it may peak because it has depleted the pool of susceptible individuals prior to the period of highest transmissibility. This would constitute a complete first wave because we would expect the highest possible attack rate and highest associated herd immunity. However, if the epidemic arrives late in the season, it may peak and decline before it has been able to deplete the susceptible population, because the transmission season is drawing to a close⁸. This would generate only partial herd immunity, and ensure subsequent epidemics are required to complete the first wave.

In light of these potentially complex invasion dynamics, the continuing refinement of infection-burden maps using incidence data should be a research priority. However, the use of surveillance data alone will pose considerable challenges. As with other emergent pathogens such as pandemic influenza, MERS and Ebola, human behaviour will change so as to affect both the reporting process and underlying infection process, making the interpretation of surveillance data difficult. Wider awareness of the virus and greater testing capacity will increase the probability that clinical cases are tested and reported, while interventions such as personal reduction in vector exposure and active vector control may reduce transmissibility. Similarly, although patterns of microcephaly could be leveraged to make epidemiological comparisons of different populations, rates of reporting microcephaly will increase as

case definitions are standardized⁹, while the underlying rate of microcephaly per infection will probably decrease as some women choose to delay pregnancy or abort pregnancies after suspected infection.

Given the need to improve the accuracy of high-resolution incidence maps, and in anticipation of these changes in surveillance data, geographically representative serological samples should be obtained, despite current difficulties in the interpretation of serological assays for flaviviruses⁶. Such samples could come from existing serum archives, discarded clinical samples, or newly initiated prospective studies. The accuracy with which we will be able to infer recent Zika infection from serum will probably improve rapidly in the near future because of refinements in the assay itself and also because of an increased ability to use novel statistical models to disentangle cross-reaction from infection¹⁰. Even if it is not possible to be certain of the infection status of some individuals, it should be possible to estimate overall attack rates with good accuracy for distinct populations and also to assess the degree of herd immunity. However, if samples are not collected and curated in the very near future — in parallel with the development of the assay technology — the public health advantages of the improved serological tests will be greatly diminished. Essentially, there currently exists the opportunity to validate and refine the mechanistic approach presented by Perkins and colleagues, but that opportunity will be rapidly lost.

Refined Zika incidence maps, which take into account different types of up-to-date incidence data, are important to enable optimal design for vaccine trials. As with recent Ebola vaccines in West Africa¹¹, average levels of Zika infection in South and Latin America will likely fall to far lower levels⁷ at exactly the time that candidate vaccines are ready for late-stage trials. Accurate identification of those communities most at risk will ensure that trials have the highest chance of success and also that their participants derive the greatest possible health benefit. □

Steven Riley is at the MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, W2 1PG, UK.
e-mail: s.riley@imperial.ac.uk

References

- Perkins, T. A., Siraj, A. S., Ruktanonchai, C. W., Kraemer, M. U. G. & Tatem, A. J. *Nature Microbiol.* **1**, 16126 (2016).
- Hay, S. I. *et al. PLoS Med.* **6**, e1000048 (2009).
- Bhatt, S. *et al. Nature* **496**, 504–507 (2013).
- Kucharski, A. J. *et al. PLoS Negl. Trop. Dis.* **10**, e0004726 (2016).
- Kermack, W. O. & McKendrick, A. G. *Proc. R. Soc. Lond. A* **115**, 700–721 (1927).
- Lessler, J. *et al. Science* <http://dx.doi.org/10.1126/science.aaf8160> (2016).
- Ferguson, N. M. *et al. Science* <http://dx.doi.org/10.1126/science.aag0219> (2016).
- Johansson, M. A., Dominici, F. & Glass, G. E. *PLoS Negl. Trop. Dis.* **3**, e382 (2009).
- França, G. V. A. *et al. Lancet* [http://dx.doi.org/10.1016/S0140-6736\(16\)30902-3](http://dx.doi.org/10.1016/S0140-6736(16)30902-3) (2016).
- Kucharski, A. J. *et al. PLoS Biol.* **13**, e1002082 (2015).
- Henaou-Restrepo, A. M. *et al. Lancet* **386**, 857–866 (2015).
- Hethcote, H. W. *SIAM Rev.* **42**, 599–653 (2000).
- Mordecai, E. *et al. Preprint at* <http://dx.doi.org/10.1101/063735> (2016).