

Exploration of variables best able to predict a mother's exposure to Zika Virus in Salvador, Brazil



Introduction

In 2015, Zika Virus (ZIKV) was detected for the first time in Brazil. This marked the start of the largest outbreak of ZIKV ever recorded, with at least 500,000 cases observed across forty countries and territories in one year (Ali *et al.*, 2017). This outbreak is part of a worrying trend towards increasing frequency of arbovirus outbreaks, observed globally over the past four decades (Ali *et al.*, 2017). Though the reasons for this trend are not fully understood, a prevalent view is that it reflects changes in environmental and social factors. With ZIKV being linked to an increased frequency of children born with Microcephaly and severe disability, it is critical to identify the factors which influence the risk of a mother's exposure to the virus, in order to avoid additional cases of Microcephaly. The objective of this study was to determine the variables that best predict a mother's exposure to ZIKV in Salvador, Bahia. The reasoning behind the focus on the mother, was that, if the mother's risk to ZIKV decreases, the risk of Microcephaly in the fetus is also reduced.

Method

Multivariate logistic regression analysis was performed on a single combined dataset from two sources; (1) a nationwide census conducted in 2010, and (2) subject level data on mothers who gave birth in the Hospital Geral Roberto Santos in Salvador between 1st October 2015 to 31st January 2016. To avoid the risk of both spurious associations and over-fitting, a literature evaluation was conducted to condense the number of explanatory variables in this database, down to those where a link to ZIKV is plausible. From the logistic regression analysis, five models were produced; a Saturated model containing all the explanatory variables, a simple model containing only the age of the mother, and then a Socioeconomic, Environmental and finally a Demographic model, which use the variables corresponding to that group classification. Receiver Operating Characteristic (ROC) curve analysis with cross validation using a randomly selected testing and training set, was then performed to compare the predictive ability of the five models against each other. This analysis was repeated 5000 times, each time with a different testing and training set, to reduce the risk of chance predictions. For each iteration, the model with the largest area under the curve, which indicates a better predictive ability, was recorded.



Figure 1: A map of the countries where there is a risk of Zika Virus. Taken from Clifford (2017).

Results

For more than half of the iterations of the ROC curve analysis, the socioeconomic model had the largest area under the curve, which indicates that this is the model most able to correctly predict whether a mother had been exposed to ZIKV. However, as can be seen in figure 2, whilst it is evident that the average ROC curve for the socioeconomic model is superior (shows a larger area under the curve with higher sensitivity for any given specificity), it does not appear to be very different from the four other models, indicating a similar predictive power.

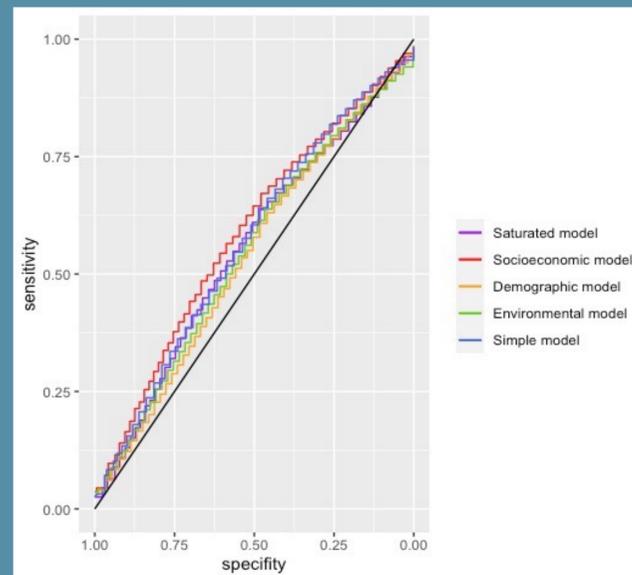


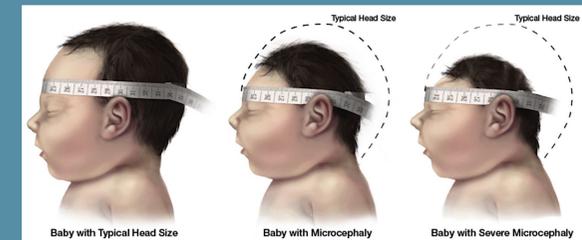
Figure 2: A graph showing the average ROC curves for each of the five models, created by binning by specificity and then averaging over the 5000 iterations

Impact of Research

The objective of this study was to determine which variables are the best predictors of a mother's exposure to ZIKV in Salvador. Results from this study will be used to inform a subsequent study, which will attempt to map the risk of ZIKV exposure to mothers in Salvador. It is the aim that this subsequent study may be used to inform policy makers where preventative policies should be targeted, in order to reduce the risk of Microcephaly.

The ROC curve analysis showed that the socioeconomic logistic model was correctly able to predict whether a mother had been exposed to ZIKV, on more occasions than the other four models. This suggests that mother residing in an area with a low socioeconomic status are at a higher risk of exposure to ZIKV. It is recommended that the subsequent study uses socioeconomic variables as predictor variables of exposure. However, as demonstrated by figure 2, it cannot be concluded that socioeconomic variables are the only predictor variables of exposure, as there does not appear to be large visual differences between the average individual ROC curves. Therefore, further analysis is required to investigate the effect of environmental and demographic variables on a mother's risk of exposure to ZIKV.

Figure 3: A diagram depicting how the appearance of a baby changes with Microcephaly. Photo taken from CDC (2020).



Summary

The objective of this study was to determine the variables that best predict a mother's exposure to Zika Virus (ZIKV) in Salvador, in order to inform a subsequent spatial analysis which will aim to map the risk of Zika Virus exposure to mothers in Salvador. It was concluded that socioeconomic variables should be used by the subsequent study as predictor variables. Further analysis is needed before commenting on the effect of environmental and demographic variables on risk of ZIKV exposure.

Acknowledgments

I would like to express my appreciation to my supervisor Emanuele Giorgi, for his support and valuable expertise during this project. I also want to thank Nivison Ruy Rocha Nery Júnior, Federico Costo, Pablo Aguilar and the rest of the team, for sharing their data with me for use in this project.

References

- Ali, S., Gugliemini, O., Harber, S., Harrison, A., Houle, L., Ivory, J., Kersten, S., Khan, R., Kim, J., Leboa, C., Nez-Whitfield, E., O'Marr, J., Rothenberg, E., Segnitz, R.M., Sila, S., Verwillow, A., Vogt, M., Yang, A. & Mordecai, E.A. (2017) Environmental and Social Change Drive the Explosive Emergence of Zika Virus in the Americas. *PLOS Neglected Tropical Diseases*, 11 (2).
- CDC. (2020). *Facts about Microcephaly*. Available: <https://www.cdc.gov/ncbddd/birthdefects/microcephaly.html>. Last accessed 28.03.2021.
- Clifford, H (2017). *Zika Virus Disease Map*. Available: <https://www.nomadtravel.co.uk/travel-vaccines/attachment/zika-virus-disease-map>. Last accessed 28.03.2021.