



HAL
open science

Epidemiological data accessibility in Brazil: Current challenges for an adequate response to emergencies

Flávio Codeço Coelho, Cláudia Torres Codeço, Oswaldo Gonçalves Cruz, Sabrina Camargo, Pierre-Alexandre Bliman

► To cite this version:

Flávio Codeço Coelho, Cláudia Torres Codeço, Oswaldo Gonçalves Cruz, Sabrina Camargo, Pierre-Alexandre Bliman. Epidemiological data accessibility in Brazil: Current challenges for an adequate response to emergencies. *The Lancet Infectious Diseases*, 2016, 16, pp.524 - 525. 10.1016/S1473-3099(16)30007-X . hal-01395754

HAL Id: hal-01395754

<https://inria.hal.science/hal-01395754v1>

Submitted on 14 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Epidemiological data accessibility in Brazil: Current challenges for an adequate response to emergencies

Flávio Codeço Coelho^a, Cláudia Torres Codeço^b, Oswaldo Gonçalves Cruz^b,
Sabrina Camargo^a, Pierre-Alexandre Bliman^{a,c}

^a*Fundação Getulio Vargas, Praia de Botafogo, 190, Rio de Janeiro, Brazil*

^b*Fundação Oswaldo Cruz, Av. Brasil 4365, Rio de Janeiro, Brazil*

^c*Sorbonne Universités, Inria, UPMC Univ Paris 06, Lab. J.-L. Lions UMR CNRS 7598, Paris, France*

Concerns about data sharing and transparency during epidemiological emergencies are not new.^{1,2,3,4} Recently, Dye et al.⁵ announced in an editorial of the Bulletin of the World Health Organization, an initiative, called Zika Open, through which the manuscripts and respective data submitted to the Bulletin would be open from the date of submission onwards, under a Creative Commons License (CC BY IGO 3.0). This is an important initiative. In this comment we would like to bring to the international academic community other challenges faced, particularly in Brazil, in the timely, lawful access to governmentally collected disease-notification data, that are essential in understanding not only the current Zika epidemic, but also any future public health emergency.

Brazil is commonly praised for its country-wide disease notification system, SINAN. It is indeed a visionary system that allows for continuous assessment of epidemiological dynamics throughout the country, that would otherwise be impossible. Created in the 1990s, it continually records cases of many diseases defined as of compulsory notification by health care facilities all over the country. Case reports come mostly from public healthcare facilities, as the private healthcare sector does not seem to find sufficient incentives to comply with the required notifications. SINAN is centrally managed and releases annual reports with aggregated data. SINAN also has a web interface where one can tabulate notified cases of diseases per municipality, month and covariates. This open notification system has fed many hundreds

epidemiological studies and has contributed to the development of a strong Brazilian epidemiological research.

However, there are major drawbacks in SINAN,⁶ mainly its slow update cycle and restrictive data-access policies and tools. In the advent of public health emergencies - such as the current Zika epidemic - this is not satisfactory. First, there is a delay of months to years for data to be available to researchers. This is due to the logistics, on one hand, but also the policy of cleaning the data before publishing. However, to be used for real time epidemiological analysis, data are needed as they appear, even if they are to be corrected and updated later. Secondly, SINAN only provides access to monthly data, which is too coarse a time-scale for disease modeling. In general, many of the challenges regarding the usability of SINAN are related to the use of legacy database management systems and of data entry systems which provide insufficient tools to prevent data entry errors.

SINAN and other public health information systems also suffer from accessibility issues. They were designed for manual query by a human. Modern mathematical and statistical epidemiology methods can consume large volumes of data and provide unique insight into the early dynamics of transmission,^{7,8} but demand machine readable data, i.e., data endpoints which can be automatically queried by analytical software. This is an opportunity for SINAN and other national notification systems to rethink their structure and data access policies. If this is not done, SINAN is risking complete obsolescence.

A good example of application which benefits from better access to readily accessible data are early warning and nowcasting systems for transmissible diseases, such as the InfoDengue (Info.dengue.mat.br), currently running in Rio de Janeiro. By properly integrating data from different sources, such as climate, social media activity and disease notification, each with its own level of sensitivity and specificity, Infodengue treats uncertainties present in each of the individual sources and provide alerts and warnings to guide public health action. However, restrictive access policies make it very difficult to run nowcasting systems. The need of specific permissions and human interference to access data imposes limitations to such initiatives. Unfounded concerns with confidentiality and potential ill-effects of data transparency are among the main barriers to data-sharing, but this could be properly managed by well established techniques, such as some level of aggregation. Both the Brazilian and the international academic communities would greatly benefit from a clear stance on the part of the government, favoring open data access,

along with the necessary investments required to make it a reality. On our part, InfoDengue is prepared to make our enriched dataset available to all interested under a Creative Commons License, as suggested by Dye et al.⁵ However, the regulatory foundations for such a service are still not clear.

We hope that the current zika epidemic can serve as wake-up call to change these outdated policies. Once we start to see the positive feedback of better data accessibility on the response to epidemiological emergencies, perhaps the private health sector will begin to see the importance of better compliance with notification requirements.

References

- [1] K. Modjarrad, V. S. Moorthy, P. Millett, P.-S. Gsell, C. Roth, M.-P. Kieny, Developing Global Norms for Sharing Data and Results during Public Health Emergencies, *PLoS Med* 13 (2016) e1001935.
- [2] C. J. M. Whitty, T. Mundel, J. Farrar, D. L. Heymann, S. C. Davies, M. J. Walport, Providing incentives to share data early in health emergencies: the role of journal editors, *The Lancet* 386 (2015) 1797–1798.
- [3] Benefits of sharing, *Nature* 530 (2016) 129–129.
- [4] E. Callaway, Zika-microcephaly paper sparks data-sharing confusion, *Nature* (2016).
- [5] C. Dye, K. Bartolomeos, V. Moorthy, M. P. Kieny, Data sharing in public health emergencies: a call to researchers, *Bulletin of the World Health Organization* (2016).
- [6] P. R. S. Galvao, A. T. Ferreira, M. Maciel, R. D. Almeida, D. Hinders, P. A. Schreuder, L. R. S. Kerr-Pontes, An evaluation of the Sinan health information system as used by the Hansen’s disease control programme, Pernambuco State, Brazil., *Leprosy review* 79 (2008) 171–182.
- [7] F. C. Coelho, L. M. d. Carvalho, Estimating the Attack Ratio of Dengue Epidemics under Time-varying Force of Infection using Aggregated Notification Data, *Scientific Reports* 5 (2015) 18455.
- [8] M. F. Gomes, A. P. Piontti, L. Rossi, D. Chao, I. Longini, M. E. Halloran, A. Vespignani, Assessing the international spreading risk associated with

the 2014 West African Ebola outbreak, PLOS Currents Outbreaks 1 (2014).