GoogleTrends reflects the abundance of the Asian tiger mosquito (*Aedes albopictus*): a call for the web-based surveillance of invasive alien vector species

Jacopo Cerri[1] and Sandro Bertolino[1]

[1]Dipartimento di Scienze della Vita e Biologia dei Sistemi, Università degli Studi di Torino, Via Accademia Albertina 13, 10123, Torino. **email**: jacopocerri@gmail.com

Abstract

The Asian tiger mosquito (*Aedes albopictus*) is becoming widespread in Europe, where it can transmit some major arboviruses, including Chikungunya and Dengue. While surveillance initiatives are being implemented and harmonized between states, the spread of *A.albopictus* is outrunning them and cost-effective surveillance tools are needed.

In this study, we tested whether on-line searches on Google can be adopted to monitor the abundance of *A.albopictus*. By using data from a long-term monitoring program in the Emilia-Romagna region (Italy), we tested whether the monthly level of infestation was causally related to the monthly GoogleTrends index for the species.

The logarithm of the level of infestation from *A.albopictus* had a very strong causal effect over the total volume of Google searches about the species. Our statistical model was highly predictive for the GoogleTrends index, suggesting that this cheap on-line metric can be adopted as a proxy for the real level of infestation from *A.albopictus*.

While GoogleTrends has been adopted to identify and monitor epidemics, including vector-borne diseases, no study tested whether it can reflect the numerical abundance of vector species. To the best of our knowledge, our study, among the few validating GoogleTrends with surveillance data, was the first one opening this possibility. Therefore, we believe that the analysis of on-line search volumes might become an important complement to existing surveillance initiatives for invasive vector species worldwide.

Introduction

Biological invasions are a global driver of change, whose ecological and socio-economic impacts have been increasingly acknowledged over the last few decades [1]. The transmission of pathogens, operated by invasive alien vector species, is among the most worrisome consequences of biological invasions, due to its ecological and socio-economic impacts and its interplay with climate change and urbanization [2][3]. Therefore, due to their role in major epidemics [4] and also their ecological impacts [5], invasive alien mosquitoes have become among the most studied, and managed, invasive alien species.

Over the last 40 years, the Asian tiger mosquito (*Aedes albopictus*) successfully colonized many countries outside of its native range in Sout-East Asia [6][7], due to its capacity to exploit invasion pathways associated with global trade and its ecological flexibility [8]. In Europe, *A. albopictus* successfully established in Mediterranean countries, following multiple accidental introductions between 1970s and early 2000s, and is now steadily expanding its distribution to Central Europe and the United Kingdom, due to climate change and milder winter conditions [9]. Its spread in Europe is expected to continue throughout the next few decades [10][11], raising sanitary concerns. In facts, *A. albopictus* is an important vector for the transmission of Chikungunya, and it could also transmit Dengue virus and dirofilarial worms. Moreover, more than 20 other arboviruses, including Zika, the yellow fever and the West Nile virus were isolated from *A. albopictus* in different parts of the world, and their transmission was proven in laboratory experiments [12][9]. Due to these sanitary impacts, and its capacity to reduce the quality of life of people living in infested areas [13], *A. albopictus* can be considered one of the invasive alien species with the highest social and economic impacts in Europe.

To date, member states of the European Union enforced various monitoring schemes for the species, to monitor its expansion, its seasonal phenology and the circulation of arboviruses [14][9]. However, the harmonization of data collection initiatives is still ongoing, being carried out by different national agencies adopting heterogeneous protocols, with sampling efforts that vary across space and that are sometimes limited to the infestation period of the species. Furthermore, long-term data are missing and, to the best of our knowledge, no open-access dataset from Europe is available on VectorBase (https://vectorbase.org/popbio-map/web/) or can be easily retrieved from the Internet. This gap can limit the capacity to reconstruct invasion dynamics, understand phenological shifts cause by global change and limit the design of effective policies for managing and monitoring the species [15]. There is a need for large-scale, long-term, and open-access data about A. albopictus in Europe.

Over the last few years, the analysis of Internet search volumes became a relatively popular approach to disease surveillance [16][17], including arboviruses [18][19]. As Internet penetration grew steadily at the global level, being now above 80-90% in many developed countries, on-line searching is nowadays a common human behavior, adopted by most people on a daily basis to collect information about the most different topics, especially on Google, which vastly dominates the market of search engines in Western countries.

Considered that Google searches are influenced by the phenology of animal and plant species [20], it is surprising that no study considered the use of on-line search volumes for monitoring the population dynamics of invasive alien vectors, like *A. albopictus*. This approach should work for European countries: *A. albopictus* is an iconic invasive mosquito, which is easily distinguished from native species due to its color and its diurnal habits, and which constitutes a nasty everyday experience for people living in its invaded range, especially in urbanized environments. It is reasonable to assume that people increase their searches on Google about the species, in response to growing levels of infestation or to the arrival of this new mosquito in their area. To date no study considered this idea, and indeed the use of GoogleTrends in conservation biology has never been validated against ecological data from the field, like in epidemiology [19].

In this research, we aim to fill this gap, by testing for the existence of a causal relationship between real levels of infestation and the GoogleTrends index about *A. albopictus*. Notably, we hypothesized that H_1: the level of infestation by A. albopictus affected the value of the GoogleTrends index, at the same time unit.

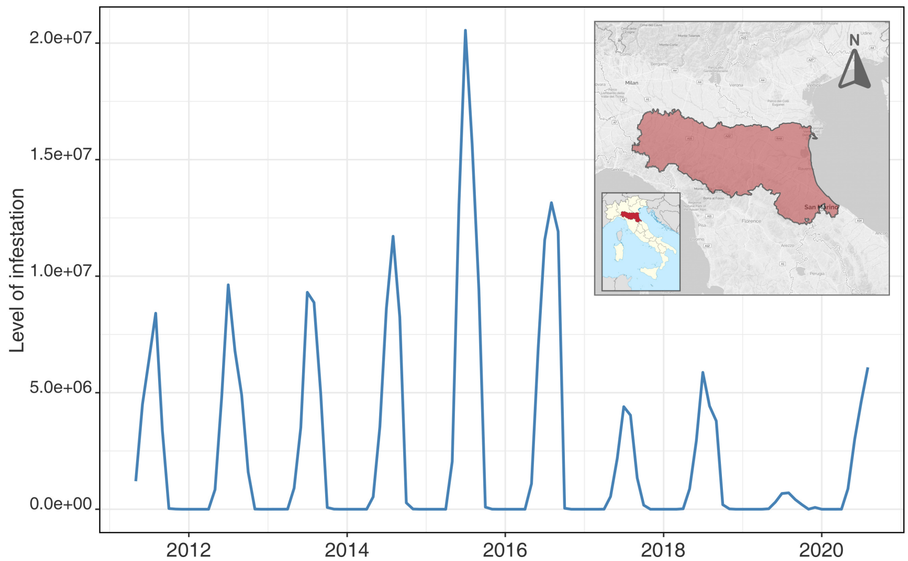

Figure 1 | Temporal evolution of the level of infestation from *A. albopictus* in the Emilia-Romagna region, between 2011 and 2020.

Methods

In this study we tested whether the GoogleTrends metric could be adopted as a valid proxy for the level of infestation by *A. albopictus*. Notably, we tested for the following hypothesis:

- H_1: the level of infestation by *A. albopictus* influenced the value of the GoogleTrends index, at the same time unit.

As a benchmark, we adopted long-term data about *A. albopictus* collected in Emilia-Romagna (Italy) between May 2011 and August 2020. Data collection was carried out by the regional health agency, through ovitraps located in 9 cities, which were checked every 14 days throughout all the year. For each city, the average number of eggs, the percentage of traps with eggs and the overall number of traps were provided [21]. This monitoring scheme was chosen as it was the only one whose temporal and spatial scales could be perfectly matched to those of the GoogleTrends index, for calibration purposes.

We extracted the monthly GoogleTrends index for the term “tiger mosquito” (“zanzara tigre”), for the same timespan and for the whole Emilia-Romagna region. The GoogleTrends index is obtained by: (*i*) dividing the monthly number of searches for a certain keyword for the total volume of Google searches in the same timespan and area, then by (*ii*) dividing again this value for the maximum value of the time series and (*iii*) multiplying by 100. In our case, we rescaled our index between 0 and 1, for modeling purposes. Then, to match the two time series, we calculated the monthly expected number of eggs, and we averaged the value across the 9 cities, to obtain a value for the whole region. The expected number of eggs was calculated by multiplying the average number of eggs, the percentage of traps with eggs and the overall number of traps in each city. Hereinafter we will refer to the average number of eggs as the "level of infestation".

We adopted a Bayesian generalized additive model, based on a Binomial distribution of the error and adopting first-order lagged residuals. To test for H_1, in our model we predicted the monthly value or the GoogleTrends index in function of the logarithm of the level of infestation. The level of infestation was converted on a logarithmic scale to reduce outliers, which could have been caused by media coverage of *A. albopictus* [22], rather than by seasonal changes.

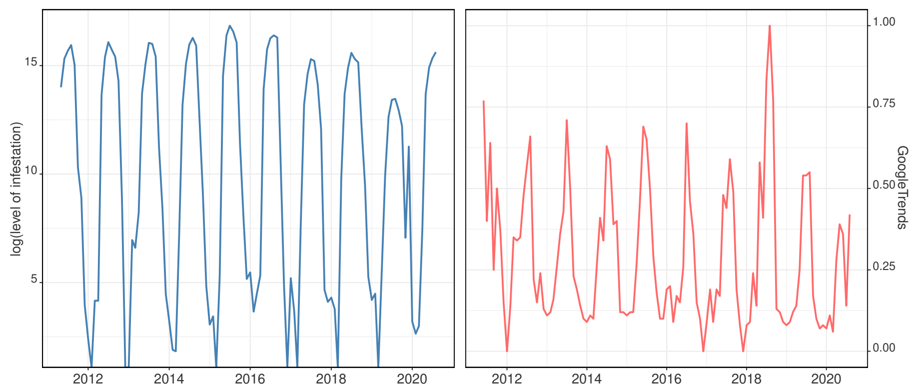

Figure 2 | Temporal evolution of the level of infestation (log-converted, left) and the GoogleTrends index (right) for *A. albopictus* in the Emilia-Romagna region, between 2011 and 2020.

Results

The level of infestation from *A. albopictus* showed marked seasonal fluctuations, in line with the phenology of the species. Moreover, we observed a peak in 2015, in line with findings from the monitoring scheme (https://www.arpae.it/dettaglio_notizia.asp?id=7037&idlivello=1504), which confirmed the goodness of our aggregation procedure for capture data (Fig. 1). On the other hand, the GoogleTrends index of *A. albopictus* remained relatively stable through time, with seasonal fluctuations. The GoogleTrends index showed only a single outlier in July 2018, probably related to a regional communication campaign about vector-borne diseases (https://salute.regione.emilia-romagna.it/campagne/zanzare/zanzare-tigre-zanzare-comuni-e-pappataci-la-campagna-informativa-per-contrastare-la-diffusione-degli-insetti-vettori).

Our model, which included a first-order temporal autocorrelation, indicated that there was a remarkable association between the level of infestation and the GoogleTrends index for *A. albopictus*, in the Emilia-Romagna region between 2011 and 2020. The model explained a considerable proportion of variability in the data ($R^2 = 0.96$) and had a good fit to observed values of the GoogleTrends index (Fig. 3). Therefore, we found support for H_1.

Discussion and conclusion

To the best of our knowledge, this study constitutes a first validation of web-based monitoring of *A. albopictus*, based on the GoogleTrends index. While Proulx et al. [23] already suggested that GoogleTrends could be adopted to investigate the seasonality of mosquito outbreaks, we deem our research to significantly advance their findings, by comparing their claims to long-term, large-scale, field data about A. albopictus. Overall, our candidate model performed extremely well, showing that the real level of infestation influenced the volume of Google searches about *A. albopictus*. Therefore, we concluded that the amount of mosquitoes experienced by people led them to document about them on the Internet. While this fact is not surprising per-se, at a time where Internet searches are becoming pervasive, we were astonished by how well the GoogleTrends index was predicted by mosquito abundance. This finding indicates that GoogleTrends might be a valuable proxy for the level of infestation, at least in areas characterized by a good access to the Internet, a prolonged period of infestation and relatively large urbanized areas, the main habitat for *A. albopictus* in its native and invaded range [24][25]. All these conditions were met in the study area, the Emilia-Romagna region, but of course might not apply to other contexts: Europe is also characterized by regions with little urbanization, as well as by shorter infestation periods, and lower densities of *A. albopictus*. Future studies should confirm whether our results can be replicated at these areas, as this could finally pave the way for a pan-European monitoring of *A. albopictus*, which would integrate existing surveillance through ovitraps. However, considered that the species is particularly problematic in urbanized areas, the possibility to monitor its abundances in regions with large cities in Central Europe and the UK, where the species is rapidly expanding its distribution due to climate change, might be particularly important.

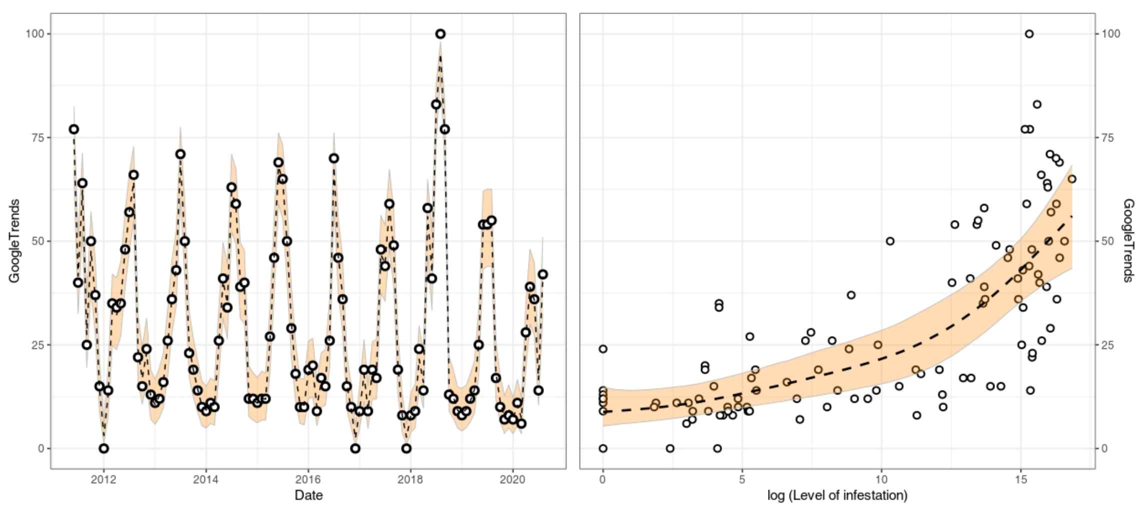

Figure 3 | Fitted versus observed values for the GoogleTrends index, from our predictive model (left), and relationship between the logarithm of the level of infestation and the GoogleTrends index for *A. albopictus*.

Moreover, GoogleTrends worked well for *A. albopictus*, an invasive alien pest which is common in urbanized areas, clearly distinguished from native mosquitoes, due to its morphology and habits, and nasty to most people on a daily basis. As other mosquitoes from the genus *Aedes* are invading Europe (*A. koreicus* [26]; *A. japonicus* [27]), we do not know how Google searches could be influenced by the misclassification of these species from Internet users, in case these become abundant in the near future.

Also, it is important to say that we warn against a näive use of GoogleTrends, especially for mediatic species. Whenever invasive alien pests are covered by media, often through sensational news, curiosity towards them peaks, as it does the amount of searches about them on the Internet [22]. This can blur the association between GoogleTrends and real population dynamics, and like in the case of Zika in the US [28]. While no epidemics occurred in the study area in our timespan, thereby not affecting our case study, the Emilia-Romagna region for example faced the first Chikungunya epidemics in Italy in 2007, and other epidemics occurred in Italy and Europe since then [29][30]. Researchers must be aware of their occurrence, before inferring population changes from online search volumes, and should discard areas and timespan affected by epidemics, or adopt ad-hoc approaches to data analysis. Longitudinal quantile regression [31], for example, could be adopted to model the separate effect of the numerical abundance of a species and news volumes, extracted from GDELT [28], on different volumes of searches on Google. By using this approach, researchers could observe a differential effect of these two components for different values of the GoogleTrends index, with media exposure being more predictive for periods with high search volumes.

We deem our approach to be adaptable to other common and iconic invasive alien pests, which are familiar to laypeople. For example, the marmorated stinkbug (*Halyomorpha halys*), a major invader which could seriously harm fruit orchards, was already mapped through citizen science [32]. As the species is easily observed when it enter the houses in the fall, or in gardens throughout the breeding season, GoogleTrends might be useful to predict its population dynamics. Similarly, we believe that GoogleTrends can be adopted for monitoring population dynamics of other vector species, such as ticks. Tick-borne diseases can have serious implications for human and animal health [33], and their frequency and diffusion in Europe is increasing, due to climate change and environmental modifications [34]. Tick removal is a common behavior, among outdoor recreationists or pet owners, and it is likely that people regularly search for ticks on Google, to remove them in a safe and effective way. The use of GoogleTrends as an indicator variable for ticks can potentially be even more important than for mosquitoes, because tick-borne diseases are characterized by an even more demanding surveillance than arboviruses [35].

Overall, we would like to remark that GoogleTrends seems to be a valuable source of information to map population dynamics of invasive alien pests and vector species. This information could be adopted for surveillance, to detect areas of geographical expansion and also phenological changes related to climate change. At a time where invasive alien pests are introduced at an unforeseen rate [36], we believe on-line search volumes to be a fundamental complement to field-based sampling, whose implementation will hardly keep the same pace of biolgical invasions.

References

1. Simberloff, D. (2013). *Invasive species: what everyone needs to know*. Oxford University Press. https://global.oup.com/academic/product/invasive-species-9780199922031?cc=it&lang=en

2. Caminade, C., McIntyre, K. M., & Jones, A. E. (2019). Impact of recent and future climate change on vector-borne diseases. *Annals of the New York Academy of Sciences*, 1436(1), 157. https://doi.org/10.1111/nyas.13950
3. Wilke, A. B., Beier, J. C., & Benelli, G. (2019). Complexity of the relationship between global warming and urbanization–an obscure future for predicting increases in vector-borne infectious diseases. *Current opinion in insect science*, 35, 1-9. https://doi.org/10.1016/j.cois.2019.06.002
4. Wilder-Smith, A., Gubler, D. J., Weaver, S. C., Monath, T. P., Heymann, D. L., & Scott, T. W. (2017). Epidemic arboviral diseases: priorities for research and public health. *The Lancet infectious diseases*, 17(3), e101-e106. https://doi.org/10.1016/S1473-3099(16)30518-7
5. LaPointe, D. A., Atkinson, C. T., & Samuel, M. D. (2012). Ecology and conservation biology of avian malaria. *Annals of the New York Academy of Sciences*, 1249(1), 211-226. https://doi.org/10.1111/j.1749-6632.2011.06431.x
6. Kraemer, M. U., *et al.* (2019). Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature microbiology*, 4(5), 854-863. https://doi.org/10.1038/s41564-019-0376-y
7. Kraemer, M. U.,*et al.* (2015). The global compendium of *Aedes aegypti* and *Ae. albopictus* occurrence. *Scientific data*, 2(1), 1-8.https://doi.org/10.1038/sdata.2015.35
8. Reiter, P. (1998). *Aedes albopictus* and the world trade in used tires, 1988-1995: the shape of things to come?. *Journal of the American Mosquito Control Association*, 14(1), 83-94. https://www.biodiversitylibrary.org/content/part/JAMCA/JAMCA_V14_N1_P083-094.pdf
9. European Centre for Disease Prevention and Control (ECDC) (2020). Aedes albopictus, factsheet for experts. https://www.ecdc.europa.eu/en/disease-vectors/facts/mosquito-factsheets/aedes-albopictus
10. Caminade, C., *et al.* (2012). Suitability of European climate for the Asian tiger mosquito *Aedes albopictus*: recent trends and future scenarios. *Journal of the Royal Society Interface*, 9(75), 2708-2717. https://doi.org/10.1098/rsif.2012.0138
11. Cunze, S., Kochmann, J., Koch, L. K., & Klimpel, S. (2016). *Aedes albopictus* and its environmental limits in Europe. *PLoS One*, 11(9), e0162116. https://doi.org/10.1371/journal.pone.0162116
12. Bonizzoni, M., Gasperi, G., Chen, X., & James, A. A. (2013). The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives. *Trends in parasitology*, 29(9), 460-468. https://doi.org/10.1016/j.pt.2013.07.003
13. Halasa, Y. A., *et al.* (2014). Quantifying the impact of mosquitoes on quality of life and enjoyment of yard and porch activities in New Jersey. *PloS one*, 9(3), e89221. https://doi.org/10.1371/journal.pone.0089221
14. Bellini, R., et al. (2020). Practical management plan for invasive mosquito species in Europe: I. Asian tiger mosquito (*Aedes albopictus*). *Travel Medicine and Infectious Disease*, 101691. https://doi.org/10.1016/j.tmaid.2020.101691
15. Groom, Q. J., et al. (2017). Seven recommendations to make your invasive alien species data more useful. *Frontiers in Applied Mathematics and Statistics*, 3, 13.https://doi.org/10.3389/fams.2017.00013
16. Carneiro, H. A., & Mylonakis, E. (2009). Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clinical infectious diseases*, 49(10), 1557-1564. https://doi.org/10.1086/630200
17. Nuti, S. V., *et al.* (2014). The use of google trends in health care research: a systematic review. *PloS one*, 9(10), e109583. https://doi.org/10.1371/journal.pone.0109583
18. Strauss, R., *et al.* (2020). Investigating the utility of Google trends for Zika and Chikungunya surveillance in Venezuela. *BMC public health*, 20(1), 1-6. https://doi.org/10.1186/s12889-020-09059-9
19. Watad, A., *et al.* (2019). Forecasting the West Nile virus in the United States: an extensive novel data streams–based time series analysis and structural equation modeling of related digital searching behavior. *JMIR public health and surveillance*, 5(1), e9176. https://doi.org/10.2196/publichealth.9176
20. Mittermeier, J. C., Roll, U., Matthews, T. J., & Grenyer, R. (2019). A season for all things: Phenological imprints in Wikipedia usage and their relevance to conservation. *PLoS biology*, 17(3), e3000146. https://doi.org/10.1371/journal.pbio.3000146

21. Carrieri, M., *et al.* (2017). Quality control and data validation procedure in large-scale quantitative monitoring of mosquito density: the case of *Aedes albopictus* in Emilia-Romagna region, Italy. *Pathogens and global health*, 111(2), 83-90. https://doi.org/10.1080/20477724.2017.1292992

22. Correia, R. A., *et al.* (2019). Inferring public interest from search engine data requires caution. *Frontiers in Ecology and the Environment*, 17(5). https://doi.org/10.1002/fee.2048

23. Proulx, R., Massicotte, P., & Pépino, M. (2014). Googling trends in conservation biology. *Conservation Biology*, 28(1), 44-51. https://doi.org/10.1111/cobi.12131

24. Li, Y., *et al.* (2014). Urbanization increases *Aedes albopictus* larval habitats and accelerates mosquito development and survivorship. *PLoS Negl Trop Dis*, 8(11), e3301. https://doi.org/10.1371/journal.pntd.0003301

25. Sherpa, S., *et al.* (2020). Landscape does matter: disentangling founder effects from natural and human-aided post-introduction dispersal during an ongoing biological invasion. *Journal of Animal Ecology*. https://doi.org/10.1111/1365-2656.13284

26. Montarsi, F., *et al.* (2013). Distribution and habitat characterization of the recently introduced invasive mosquito *Aedes koreicus* [*Hulecoeteomyia koreica*], a new potential vector and pest in north-eastern Italy. *Parasites & vectors*, 6(1), 292. https://doi.org/10.1186/1756-3305-6-292

27. Cunze, S., Koch, L. K., Kochmann, J., & Klimpel, S. (2016). *Aedes albopictus* and *Aedes japonicus*-two invasive mosquito species with different temperature niches in Europe. *Parasites & vectors*, 9(1), 1-12. https://doi.org/10.1186/s13071-016-1853-2

28. Tizzoni, M., Panisson, A., Paolotti, D., & Cattuto, C. (2020). The impact of news exposure on collective attention in the United States during the 2016 Zika epidemic. *PLoS computational biology*, 16(3), e1007633. https://doi.org/10.1371/journal.pcbi.1007633

29. Barzon, L. (2018). Ongoing and emerging arbovirus threats in Europe. *Journal of Clinical Virology*, 107, 38-47. https://doi.org/10.1016/j.jcv.2018.08.007

30. Rezza, G. (2018). Chikungunya is back in Italy: 2007–2017. *Journal of travel medicine*, 25(1), tay004. https://doi.org/10.1093/jtm/tay004

31. Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1), 74-89. https://doi.org/10.1016/j.jmva.2004.05.006

32. Maistrello, L., Dioli, P., Bariselli, M., Mazzoli, G. L., & Giacalone-Forini, I. (2016). Citizen science and early detection of invasive species: phenology of first occurrences of *Halyomorpha halys* in Southern Europe. *Biological invasions*, 18(11), 3109-3116. https://doi.org/10.1007/s10530-016-1217-z

33. Boulanger, N., Boyer, P., Talagrand-Reboul, E., Hansmann, Y. (2019). Ticks and tick-borne diseases. *Medecine et maladies infectieuses*, 49(2), 87-97. https://doi.org/10.1016/j.medmal.2019.01.007

34. Süss, J., Klaus, C., Gerstengarbe, F. W., & Werner, P. C. (2008). What makes ticks tick? Climate change, ticks, and tick-borne diseases. *Journal of travel medicine*, 15(1), 39-45.https://doi.org/10.1111/j.1708-8305.2007.00176.x

35. Van den Wijngaard, C. C., *et al.* (2017). Surveillance perspective on Lyme borreliosis across the European Union and European economic area. *Eurosurveillance*, 22(27), 30569.https://doi.org/10.2807/1560-7917.ES.2017.22.27.30569

36. Seebens, H., et al. (2017). No saturation in the accumulation of alien species worldwide. *Nature communications*, 8(1), 1-9. https://doi.org/10.1038/ncomms14435