Historical Patterns of Arboviral Seroprevalence across Africa and Asia

George C. Chrisafis
The University of Vermont

Follow this and additional works at: https://scholarworks.uvm.edu/hcoltheses

Recommended Citation
Chrisafis, George C., "Historical Patterns of Arboviral Seroprevalence across Africa and Asia" (2017). UVM Honors College Senior Theses. 139.
https://scholarworks.uvm.edu/hcoltheses/139

This Honors College Thesis is brought to you for free and open access by the Undergraduate Theses at ScholarWorks @ UVM. It has been accepted for inclusion in UVM Honors College Senior Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
Historical Patterns of Arboviral Seroprevalence across Africa and Asia

George C. Chrisafis

April 19, 2017
Abstract

The emergence and resurgence of arboviruses in recent history is challenging our scientific understanding of mosquito-borne diseases and their transmission. To better contextualize recent epidemics and gain insight into historical trends in arbovirus incidence, we conducted a literature review to identify serosurveys from Africa and Asia. We compiled all serosurvey data into a table and tested for variation in disease incidence across countries and between age categories. Our analysis showed that disease incidence was consistently higher in the >15 age category than the <15 age category and revealed significant variation in incidence across countries. In addition, the mean incidence of yellow fever virus was substantially higher than the incidences of the other diseases included in the analysis. Higher incidence in the >15 age category is likely due to the long-term persistence of antibodies in human sera, while a higher incidence of yellow fever can likely be attributed to widespread vaccine use. Characteristics of countries with high disease incidence included a tropical climate, extended rainy season, and flat terrain, and countries with low disease incidence occurred at higher elevations and/or reflected a desert climate. This analysis can hopefully reveal the conditions most important in facilitating an arbovirus outbreak, leading to targeted prevention strategies in high-risk areas. It also highlights the need for continued serosurveys as a method of documenting disease spread.

Introduction

Throughout history, the incidence of diseases and patterns of disease spread have proven to be highly variable and unpredictable (Petersen, 2012; Claflin and Webb, 2015). The complexity of pathogen emergence and persistence is exemplified by arboviruses, which are viruses that rely on arthropods for transmission between hosts (Gubler, 2001). The inter-
dependent relationship between the components required for arbovirus transmission (virus, vector, host)—coupled with external factors that affect the interactions between the components—make it difficult to predict and control the spread of arboviruses (Kuno and Chang, 2005). In recent history, the emergence and resurgence of certain arboviruses has posed significant public health problems in many countries. Examples of resurgent diseases include dengue virus and yellow fever virus, which were historically prominent but thought to be controlled by the mid-20th century. In the case of yellow fever, this included the introduction and widespread use of a licensed vaccine. However, the incidence of these diseases in the past 30 years has increased dramatically, resulting in their spread to previously unaffected regions of the world and afflicting a greater number of people (Gubler, 1998). Emergent diseases such as Zika virus and chikungunya virus reflect a different pattern of incidence, as they were first characterized in the 1950s but remained relatively dormant until the onset of large-scale epidemics in the early-21st century. Like the resurgent diseases, the geographic ranges of Zika and chikungunya viruses have also expanded because of recent epidemics (Powers, 2009).

The recent trend of arbovirus emergence and resurgence is perplexing, because diseases such as Zika virus and chikungunya virus have been observed in humans for decades but haven’t caused widespread epidemics until recently. These epidemics have been noticed by public health authorities, as they have occurred in historically non-endemic areas. Similarly, dengue virus and yellow fever virus were thought to be controlled, yet have also been responsible for recent epidemics. Prior to these epidemics, Zika, chikungunya, dengue, and yellow fever viruses were thought to be of relatively low public health concern. This raises the question of which factors or changes have enabled these diseases to emerge or resurge at the same point in time, and thus warrant further exploration into the potential underlying causes of recent epidemics. Factors
theorized to be responsible for the trend of arbovirus emergence and resurgence include viral recombination and mutation, urbanization, increased commerce, global population growth, and climate change, but no causal relationships have been made between these factors and disease spread (Dash et al., 2013). Understanding the patterns underlying the spread of arboviruses is crucial in identifying which factors may have contributed to the trend of recent arbovirus epidemics.

When considering the impact of diseases on human populations, it is important to understand their history, basic biology/virology, and public health implications. Zika virus is a Flavivirus closely related to Dengue, West Nile, and Yellow Fever viruses. First characterized in humans in 1954, Zika virus remained of low public health concern for decades (Alera et al., 2015). However, a notable outbreak of Zika virus occurred in Yap Island, Micronesia in 2007, with an estimated infection rate of 73% among the island’s 7391 inhabitants (Weaver et al., 2016). Subsequent epidemics have led to the spread of Zika virus throughout Africa, Asia, Oceania, and the Americas, having been detected in Brazil in 2015 (Kindhauser et al., 2016). Transmitted by Aedes mosquitoes, Zika virus has multiple animal reservoirs, including nonhuman primates and other wild and domestic animals (Plourde and Bloch, 2016). Transmission occurs when a mosquito injects the virus into the skin of a mammalian host. Zika virus is a positive-sense, single-stranded genomic RNA which codes for three structural and seven nonstructural proteins. The interaction of the viral envelope protein with cell surface receptors mediates the initial entry of the virus into multiple types of skin cells. Once infected, the formation of autophagosomes within a cell can enhance viral replication, which occurs in the cellular cytoplasm (Hamel et al., 2015).
Dengue virus is another Flavivirus with a large global burden, as the virus is present in over 100 countries worldwide and causes an approximated 50 million infections per year. (Simmons et al., 2012). It is estimated that two-fifths of the world’s population is at risk for dengue infection, making dengue virus the most widespread vector-borne disease (Murrell et al., 2011). Dengue virus is transmitted by *Aedes* mosquitoes and utilizes nonhuman primates as reservoirs. Like Zika virus, dengue virus consists of a positive-sense, single-stranded RNA genome which encodes three structural and seven nonstructural proteins. The envelope glycoprotein initiates interactions with cell surface receptors, which facilitates the entry of the virus into cells. An incubation period of 4-7 days occurs post-infection, characterized by viral replication in the dendritic cells closest to the bite site. During this period, macrophages and lymphocytes can also be infected with the virus, which ultimately enters the blood stream. There are four serotypes of dengue virus, each of which is antigenically distinct and can result in conditions such as dengue fever, dengue hemorrhagic fever, and dengue shock syndrome. No effective treatment or vaccine has been discovered for the dengue serotypes, contributing to an inability to control the spread of the disease (Murrell et al., 2011).

Yellow fever virus is a Flavivirus endemic to regions of Africa and South America. From the 18th to the 20th century, yellow fever virus posed a significant threat to human health, but the deployment of two licensed, effective vaccines in the 1930s and 1940s caused disease incidence to decline. In recent years, the disease has been making a resurgence in certain regions of the world, often becoming endemic in areas without effective immunization programs (Monath and Vasconcelos, 2015). Currently, yellow fever virus is responsible for over 30,000 deaths annually (Quaresma et al., 2013). The virus contains a positive-sense, single-stranded RNA genome and is transmitted between nonhuman primates, mosquitoes, and humans (Monath and Vasconcelos,
Following infection by a mosquito, the virus initially replicates in local lymph nodes, but spreads to organs such as the liver, heart, and kidneys, where it can cause the formation of lesions. The liver is the primary target of yellow fever virus, and infection can result in the apoptosis of liver cells (Quaresma et al., 2013).

Chikungunya virus is an Alphavirus that was discovered in Tanzania in 1952. The virus has since spread throughout the world, reaching India, Asia, the Pacific, and the Americas and causing numerous epidemics (Sudeep and Parashar, 2008). Between 2004 and 2007, large-scale epidemics of chikungunya virus occurred in Kenya, the Indian Ocean, and the Indian Subcontinent, and over one and a half million cases were thought to have occurred in India alone during this time. On Reunion Island in the southwest Indian Ocean, an estimated 266,00 cases of chikungunya occurred in 2005, representing 34% of the island’s population (Presti et al., 2014). Chikungunya is a mosquito-transmitted virus that utilizes nonhuman primates as intermediates in its transmission cycle. The virus is a single-stranded, positive-sense RNA genome that encodes four nonstructural proteins and three structural proteins. Chikungunya virus enters cells via endocytosis, following the binding of envelope glycoproteins to cell surface receptors (Weaver and Lecuit, 2015). Once in the skin, the virus infects susceptible cells such as macrophages and fibroblasts and then enters the blood stream, where it can spread to the liver, muscle joints, lymphoid tissue, and the brain (Presti et al., 2015).

The recent trend of large-scale arbovirus epidemics due to the emergence and resurgence of diseases poses a significant public health challenge in many countries. To better understand this trend, one approach is to analyze data describing the historical incidence of diseases. Historical data pertaining to disease incidence is useful because it can reveal trends in disease spread. We therefore attempt to better understand and contextualize recent arbovirus outbreaks.
by looking at historical patterns of disease incidence in endemic countries. Two key questions raised by this analysis are whether the incidence of diseases varies by age and geographic location. To test for the presence of this variation, we will perform three different analyses. First, we will test for variation in incidence across countries, for each disease individually. Second, we will again test for variation in incidence within countries, but for each disease and age category. Third, we will test for variation in incidence by age category, for each disease and across all countries. We anticipate variability in disease incidence across countries due to climatic and geographic differences, but don’t expect major differences in disease incidence across age groups, because the incidence data comes from historically endemic countries.

Methods

To compile a dataset of historical arbovirus incidence, arbovirus serosurveys were first identified through a literature review. Using Google Scholar, searches were conducted with the following word combinations: “arbovirus, sero, human,” “dengue, sero, human,” “chikungunya, sero, human,” “Zika, sero, human,” “arbovirus, antibody, human,” “dengue, antibody, human,” “chikungunya, antibody, human,” “Zika, antibody, human,” “arbovirus, molecular, human,” “dengue, molecular, human,” “chikungunya, molecular, human,” and “Zika, molecular, human.” Each search excluded “West Nile Virus” and was designated to include “Africa” or “Asia.” The general search criteria were thus serosurveys conducted in Africa and Asia within the last 75 years, and empirical incidence data as opposed to systematic reviews. A maximum of 1,000 results were returned for each search, ordered by citation count and search relevance. All titles/abstracts were scanned to determine if a serosurvey had been conducted, and a PDF was located and saved for each article containing primary incidence data. For the purpose of this study, serosurveys conducted among febrile or clinically-suspected patients were excluded, to
prevent the inclusion of biased incidence data. The data from all serosurveys identified by the literature review was manually compiled into a data table for further analysis. The table contained the following column headings: “Country,” “Sub-region,” “Disease,” “Incidence,” “Sample Size,” “Age Group,” “Age Category (<15 or >15),” “Test Type,” and “Citation.” Sample size refers to the total number of individuals tested for the presence of a particular disease, while incidence refers to the number of individuals who tested positive for the disease. The results of some serosurveys were not reported by age group, so age data was recorded where appropriate.

All statistical analyses were conducted in RStudio (1.0.44). Disease, age category, and country were utilized as covariates. First, to observe trends in disease incidence across countries, the observed incidence of a single disease was computed for all countries individually. To accomplish this, all row entries for a disease (i.e. Zika) were selected for, and a new data frame was created containing solely these entries. The total number of positive samples was then summed for each country and divided by the total sample size. For countries that lacked incidence data pertaining to the selected disease, all NA’s were replaced with zeros. A matrix was then created with a row number equal to the total number of countries and the following column headings: “Country,” “Positive,” “Total,” “Mean,” “Upper,” “Lower.” Country names were added to the matrix (“Country”), as were the total number of positive samples for a country (“Positive”) and the total number of samples tested (“Total”). Using the “binom” package (1.1-1; Dorai-Raj, 2014), a 95% credible interval of the mean incidence was calculated for each country, using Bayesian inference. The software infers a posterior distribution of incidences and we used the default parameters, which are a beta prior distribution and a binomial likelihood (see Gelman et al., 1997 for more information). The upper and lower limits of each credible interval were
added to the matrix under their corresponding columns, as well as the mean incidence. This analysis was repeated for each of the four diseases.

To observe trends in disease incidence between countries and age categories, the incidence rates of a single disease were computed for both age categories individually (<15 and >15) and across all countries. All row entries for a disease were first selected for, and the resulting data frame was modified to select all entries pertaining to one of the age categories. A 95% credible interval of the mean incidence was then computed for each country, reflecting the incidence of a single disease in one of the age categories. For example, a 95% credible interval was computed for Zika virus in all countries individually, for the <15 age category. There were two age categories for four total diseases.

To observe trends in disease incidence between age categories, the incidence of each disease was computed for each age category. Instead of initially selecting for a disease, an age category was selected for. The number of positive samples and total sample size were then summed for each disease within the age category. The matrix containing the credible intervals and associated data was modified to include a column titled “Disease,” replacing the column titled “Country.” A total of two tables were created, and both tables contained the incidence of all four diseases within an age category. To produce figures, the R packages “maps” (3.1.1; Brownrigg, 2016) and “ggplot2” (Wickham, 2009) were used.

**Results**

**Disease by Country**

Zika virus ranged from a low incidence of 0.0083 (9.18x10^-6 to 0.0215, 95% CrI) in Egypt to a high incidence of 0.8651 (0.7802 to 0.9432, 95% CrI) in Gambia. The mean incidence of Zika virus across all countries was 0.2401. Burkina Faso and Gambia reflected an incidence
greater than 0.5, while in China, Egypt, Ethiopia, Mozambique, Pakistan, Uganda, and Zambia, the incidence of Zika virus was less than 0.1. Gambia was the only country whose credible interval of the mean incidence did not overlap with the credible interval of another country. The interval of Gambia was notably higher than the intervals of other countries, as the upper limit of the next-highest credible interval was 0.5551, observed in Burkina Faso (0.5102 to 0.5551, 95% CrI). Ethiopia displayed a particularly low credible interval, as the upper limit was only 0.0185 (0.0005 to 0.0185, 95% CrI; see Fig. 1; Fig. 2).

Dengue virus ranged from a low incidence of 0.0015 (1.56x10^-6 to 0.0038, 95% CrI) in Saudi Arabia to a high incidence of 0.6351 (0.6061 to 0.6640, 95% CrI) in Sri Lanka. The mean incidence of dengue virus across all countries was 0.2378. Philippines, Sri Lanka, and Thailand reflected an incidence greater than 0.5, while in Egypt, Japan, Madagascar, Niger, Saudi Arabia, Senegal, South Africa, and Sudan the incidence of dengue virus was less than 0.1. Sri Lanka was the only country whose entire credible interval was above 0.6, although the interval still overlapped with that of Thailand. Unlike Zika virus, overlap occurred between all credible intervals, and no interval was entirely distinct. The intervals of South Africa and Saudi Arabia were notably lower than the rest, as the upper limits were 0.0036 (0.0003 to 0.0037, 95% CrI) and 0.0038, respectively (see Fig. 3; Fig. 4).

Chikungunya virus ranged from a low incidence of 0.0048 (0 to 0.0183, 95% CrI) in Palau to a high incidence of 0.7351 (0.7152 to 0.7549, 95% CrI) in Burkina Faso. The mean incidence of chikungunya virus across all countries was 0.2438. Burkina Faso, Guinea, Tanzania, and Vietnam reflected an incidence greater than 0.5, while in China, Ethiopia, India, Namibia, Niger, Pakistan, Palau, Solomon Islands, South Africa, and Vanuatu the incidence of chikungunya virus was less than 0.1. The entire credible interval of Burkina Faso was greater
than 0.7, but it still overlapped with the interval of Tanzania, which had the highest upper limit of any interval at 0.8231 (0.6294 to 0.8231, 95% CrI). Overlap was observed between all intervals. Countries in Africa showed a sharp contrast in incidence; Burkina Faso and Tanzania reflected incidence values greater than 0.7, while Ethiopia and Namibia had incidence values less than 0.02 (0.0116 and 0.0079, respectively). Palau, Solomon Islands, and Vanuatu, all island nations in the Pacific Ocean, reflected particularly low incidence values of chikungunya virus, as the highest incidence among the three was 0.0104 (0 to 0.0398, 95% CrI) in Vanuatu (see Fig. 5; Fig. 6).

Yellow fever virus ranged from a low incidence of 0.0035 (0 to 0.0136, 95% CrI) in Egypt to a high incidence of 0.7520 (0.7325 to 0.7713, 95% CrI) in Burkina Faso. The mean incidence of yellow fever virus across all countries was 0.3093. Burkina Faso, Central African Republic, Cote D’Ivoire, Lebanon, Sudan, and Thailand reflected an incidence greater than 0.5, while in Egypt, Niger, Somalia, Tanzania, and Uganda the incidence of yellow fever virus was less than 0.1. Burkina Faso was the only country whose credible interval of the mean incidence did not overlap with the credible interval of another country. The credible interval of Lebanon exceeded 0.7, with an upper limit of 0.7037 (0.5279 to 0.7037, 95% CrI). Egypt and Somalia displayed notably low credible intervals, as the upper limits were 0.0136 and 0.0215 (0 to 0.0215, 95% CrI), respectively (see Fig. 7; Fig. 8).

Comparing the incidence of multiple diseases within a country also yielded interesting results. Burkina Faso reflected a concurrently high incidence of multiple diseases, as Zika virus had an incidence of 0.5327 (0.5102 to 0.5551, 95% CrI), chikungunya virus an incidence of 0.7351 (0.7152 to 0.7549, 95% CrI), and yellow fever virus an incidence of 0.7520 (0.7325 to 0.7713, 95% CrI). Thailand followed a similar pattern, as dengue virus and yellow fever virus
had incidence values of 0.6151 (0.5997 to 0.6305, 95% CrI) and 0.5891 (0.4934 to 0.6839, 95% CrI). In contrast, Egypt reflected a concurrently low incidence of multiple diseases, as the incidences of Zika, dengue, and yellow fever viruses were 0.0083 (9.18x10^-6 to 0.0215, 95% CrI), 0.0521 (0.0209 to 0.0867, 95% CrI), and 0.0035 (0 to 0.0136, 95% CrI), respectively. Ethiopia and Niger also demonstrated this pattern of multiple low incidences. Other countries reflected a mixed pattern of incidences, such as in Tanzania where the incidence of chikungunya virus was relatively high at 0.7278 (0.6294 to 0.8231, 95% CrI), but Zika and yellow fever viruses had much lower incidence values of 0.1757 (0.0635 to 0.2980, 95% CrI) and 0.0739 (0.0300 to 0.1222, 95% CrI), respectively. Philippines also demonstrated this pattern, with a relatively high incidence of dengue virus in comparison to relatively low incidences of Zika virus and chikungunya virus.

**Disease by Country and Age Category**

In the <15 age category, Zika virus ranged from a low incidence of 0.0062 (0 to 0.0236, 95% CrI) in Egypt to a high incidence of 0.4449 (0.4116 to 0.4783, 95% CrI) in Senegal. The mean incidence of Zika virus in the <15 age category was 0.1511. In the >15 age category, Zika virus ranged from a low incidence of 0.0149 (1.7x10^-4 to 0.0384, 95% CrI) in Egypt to a high incidence of 0.9082 (0.8274 to 0.9785, 95% CrI) in Gambia. The mean incidence of Zika virus in the >15 age category was 0.3171. The largest difference in incidence between age categories was 0.4915 in Gambia, and the smallest difference in incidence was 0.0051 in India. In seven of thirteen countries containing incidence data for both age categories, no overlap was observed in the 95% credible intervals of Zika virus incidence between age categories. In all seven cases, the credible interval of the >15 age category was higher than that of the <15 age
category. A difference in incidence of greater than 0.2 was observed between age categories in Ivory Coast, Nigeria, and Senegal.

In the <15 age category, dengue virus ranged from a low incidence of 0.0155 (0.0055 to 0.0266, 95% CrI) in Niger to a high incidence of 0.6291 (0.5997 to 0.6583, 95% CrI) in Sri Lanka. The mean incidence of dengue virus in the <15 age category was 0.1948. In the >15 age category, dengue virus ranged from a low incidence of 0.0015 (1.56x10^-6 to 0.0038, 95% CrI) in Saudi Arabia to a high incidence of 0.9356 (0.9252 to 0.9459, 95% CrI) in Thailand. The mean incidence of dengue virus in the >15 age category was 0.3609. The largest difference in incidence between age categories was 0.7613 in Thailand, and the smallest difference in incidence was 0.0497 in Kuwait. In thirteen of fifteen countries containing incidence data for both age categories, no overlap was observed in the 95% credible intervals of dengue virus incidence between age categories. In these cases, the credible interval of the >15 age category was higher than that of the <15 age category in all but two cases. For Indonesia and Laos, the credible interval of incidence in the >15 age category was lower than that of the <15 age category. A difference in incidence of greater than 0.2 was observed between age categories in Cameroon, Indonesia, Singapore, Taiwan, and Thailand.

In the <15 age category, chikungunya virus ranged from a low incidence of 0.0113 (0.0031 to 0.0209, 95% CrI) in Niger to a high incidence of 0.6786 (0.4428 to 0.9006, 95% CrI) in Tanzania. The mean incidence of chikungunya virus in the <15 age category was 0.2370. In the >15 age category, chikungunya virus ranged from a low incidence of 0.0764 (0.0219 to 0.1380, 95% CrI) in China to a high incidence of 0.9323 (0.8816 to 0.9775, 95% CrI) in Senegal. The mean incidence of chikungunya virus in the >15 age category was 0.3695. The largest difference in incidence between age categories was 0.4576 in Nigeria, and the smallest
A difference in incidence was 0.0386 in India. In thirteen of eighteen countries containing incidence data for both age categories, no overlap was observed in the 95% credible intervals of chikungunya virus incidence between age categories. In all thirteen cases, the credible interval of the >15 age category was higher than that of the <15 age category. A difference in incidence of greater than 0.2 was observed between age categories in Ivory Coast, Kenya, Mozambique, Niger, Nigeria, Senegal, and Taiwan.

In the <15 age category, yellow fever virus ranged from a low incidence of 0.0078 (0 to 0.0299, 95% CrI) in Egypt to a high incidence of 0.5629 (0.5322 to 0.5936, 95% CrI) in Senegal. The mean incidence of yellow fever virus in the <15 age category was 0.1793. In the >15 age category, yellow fever virus ranged from a low incidence of 0.0064 (0 to 0.0246, 95% CrI) in Egypt to a high incidence of 0.8788 (0.8447 to 0.9118, 95% CrI) in Senegal. The mean incidence of yellow fever virus in the >15 age category was 0.3922. The largest difference in incidence between age categories was 0.5571 in Niger, and the smallest difference in incidence was 0.0014 in Egypt. In nine of thirteen countries containing incidence data for both age categories, no overlap was observed in the 95% credible intervals of yellow fever virus between age categories. In all nine cases, the credible interval of the >15 age category was higher than that of the <15 age category. A difference in incidence of greater than 0.2 was observed between age categories in Ivory Coast, Niger, and Senegal.

**Age Category by Disease**

Overall, the mean incidence of each disease was higher in the >15 age category than in the <15 age category (see Fig. 9). The largest difference in the incidence of a disease between age categories was reflected by yellow fever virus, which had an incidence of 0.3619 (0.3530 to 0.3708, 95% CrI) in the >15 category and an incidence of 0.1928 (0.1842 to 0.2014, 95% CrI) in
the <15 category. The smallest difference occurred for dengue virus, which had incidences of 0.3699 (0.3621 to 0.3777, 95% CrI) and 0.3174 (0.3067 to 0.3281, 95% CrI) in the >15 and <15 age categories, respectively. In addition, none of the credible intervals for the mean incidence of a disease overlapped between the age categories. For example, the credible interval of the mean incidence of Zika virus in the <15 age category was 0.1172 to 0.1328, compared to a higher interval of 0.2390 to 0.2573 in the >15 age category. This was true for all four diseases.

Within each category, the order of disease incidences varied slightly. For the <15 age category, dengue virus displayed an incidence of 0.3174 (0.3067 to 0.3281, 95% CrI), which was the highest incidence out of the four diseases. Chikungunya virus had the second highest incidence of 0.2571 (0.2483 to 0.2659, 95% CrI), followed by yellow fever virus at 0.1928 (0.1842 to 0.2014, 95% CrI) and Zika virus at 0.1250 (0.1172 to 0.1328, 95% CrI). There was no overlap in credible intervals within the <15 age category. For the >15 age category, dengue virus also had the highest incidence at 0.3699 (0.3621 to 0.3777, 95% CrI), but was followed by yellow fever virus at 0.3619 (0.3530 to 0.3708, 95% CrI), chikungunya virus at 0.3188 (0.3124 to 0.3253, 95% CrI) and Zika virus at 0.2481 (0.2390 to 0.2573, 95% CrI), which had the lowest incidence in both age categories. Within the >15 age category, the credible intervals for dengue virus and yellow fever virus displayed significant overlap.

Discussion

In this study, we investigated the historical incidence of arboviruses in endemic countries throughout Africa and Asia. We focused the analysis on four arboviruses that have caused some of the most recent global epidemics: Zika virus, dengue virus, chikungunya virus, and yellow fever virus. After compiling a dataset of disease incidence from serosurveys, we analyzed the dataset in three ways: disease by country, disease by age category and country, and age category
by disease. The first section of the analysis revealed significant variation in disease incidence between countries. In addition, the mean incidence of yellow fever virus across all countries was substantially higher than the incidences of the other three diseases, which displayed relatively similar values. In the second section of the analysis, the incidence of a disease within a country was always lower for the <15 age category than the >15 age category, except for the incidence of dengue virus in Indonesia and Laos, which reflected a higher incidence of dengue virus in the <15 age category than the >15 age category. The third section of the analysis showed that the incidences of all four diseases were higher in the >15 age category than the <15 age category. This result is not surprising, and tells us that the sample sizes of Laos and Indonesia were much smaller than the sample size of all the countries. Our findings partially supported our initial hypothesis, where we anticipated variability in disease incidence across countries due to climatic and geographic differences, but didn’t expect major differences in disease incidence across age groups because the incidence data comes from historically endemic countries. While disease incidence varied greatly across countries, it also varied between age groups, suggesting that age may be an important determinant of infection.

The mean incidence of yellow fever virus across all countries was higher than the mean incidences of the other three diseases. Because vaccinated individuals can show up as seropositive for yellow fever (Poland et al., 1981), one possible explanation for this observation is the availability and use of a vaccine against the disease. Between 1933 and 1961, mass yellow fever vaccination campaigns were carried out in African countries such as Burkina Faso, Cameroon, and Ivory Coast (WHO, 2017). After vaccination with the live, attenuated virus, an immune response is triggered which results in the production of antibodies against the virus. Yellow fever antibodies have been shown to persist in human sera for 30-35 years post-
vaccination, and may even confer lifelong immunity (Poland et al., 1981). The serum from a person who received the yellow fever vaccine would therefore test positive for yellow fever virus antibodies by standard diagnostic tests (such as antibody neutralization), even if they were never infected with the virus by a mosquito. These mass vaccination campaigns, which continue to distribute about 20-60 million vaccines annually, may explain why yellow fever virus appears to have a higher incidence than the other diseases (Monath and Vasconcelos, 2015). However, the interruption of mass vaccination campaigns in Africa between the 1960s and 1990s may have also assisted in the evolution of current epidemics and the resurgence of the disease (WHO, 2016).

The incidences of all four diseases were observed to be higher in the >15 age category than in the <15 age category, with two exceptions. This trend may be explained by the persistence of antibodies to a disease in human sera for many years after initial infection. For example, antibodies to dengue virus have been detected in the sera of infected individuals more than 60 years after initial infection with the disease (Imrie et al., 2007). Similarly, antibodies to chikungunya virus have been detected in human sera 20 years post-infection, including consistently high titers of antibodies in a large percentage of sera (Nitatpattana, 2014). The long-term persistence of antibodies can explain why disease incidence appears higher in the >15 age category, as the probability of obtaining a positive diagnostic test increases with age. While this trend was reflected by a vast majority of the dataset, there were two exceptions: the incidence of dengue virus in Indonesia and Laos was higher in the <15 age category than in the >15 age category. In the case of Indonesia, a 1995 serosurvey indicated a high incidence of dengue antibodies in children under 15 years of age (Graham et al., 1999). This was likely due to the occurrence of a recent outbreak, and a study published in 1987 reported that dengue had recently
become endemic to many large cities and small towns throughout Indonesia. The incidence of dengue virus had increased dramatically since the initial recognition of the disease in Indonesia in 1968, and had spread to 26 of 27 provinces by 1985 (Sumarmo, 1987). The recent spread of dengue virus in relation to the serosurvey analysis could explain why the disease displayed a higher incidence among the <15 age category. A similar situation was observed in the case of Laos, as a 2015 serosurvey also reflected a high incidence of dengue in the <15 age category (Conlan et al., 2015). In 2010 and 2013, Laos experienced widespread epidemics of dengue virus, reflecting a trend of increasing emergence. During these epidemics, the largest number of cases occurred among 10- to 20-year-olds, and children accounted for the majority of deaths (Khampapongpane et al., 2014). This provides a plausible explanation for the elevated incidence of dengue observed in the <15 age category.

Multiple diseases were observed to have high incidences in countries such as Burkina Faso and Thailand and low incidences in countries such as Egypt, Ethiopia, and Niger. Many factors and conditions may contribute to these patterns of high or low incidences, including climate and geography. Burkina Faso has a tropical climate characterized by warm, dry winters and hot, wet summers. Most of the country is a relatively flat, landlocked savanna that occurs at low elevation (CIA, 2017). Thailand also has a tropical climate, with both a rainy, warm monsoon season and a dry, cool monsoon season. While Thailand contains both plateaus and mountains, most of the country is a flat, central plain located in a river valley (CIA, 2017). Similarities between Burkina Faso and Thailand, which have high disease incidences, include a tropical climate containing a rainy season, and flat terrain (i.e. savannas or plains) which occurs at low elevations. In contrast, countries with low disease incidences display notable differences in both climate and geography when compared to countries like Burkina Faso and Thailand.
Egypt is a vast desert plateau with an arid desert climate characterized by hot, dry summers and moderate winters (CIA, 2017). Niger is composed of predominantly desert plains and sand dunes, with a desert climate that is hot, dry, and dusty (CIA, 2017). Ethiopia has a tropical monsoon climate, but the terrain consists of a high plateau with a central mountain range (World Factbook, 2017). Shared characteristics of countries with low disease incidences include a desert climate and the presence of plateaus which occur at higher elevations. In the case of Ethiopia, the high elevation together with plateaus may even mitigate the effects of a tropical climate on disease incidence, due to a lack of vector survival. *Aedes* mosquitoes cannot establish and proliferate at higher elevations, thus inhibiting the spread of vector-borne diseases such as arboviruses and leading to lower incidence rates (Eisen and Moore, 2013). As evidenced by this analysis, climatic and geographic conditions can vary greatly within regions of the world, and this may affect the incidence of diseases. For example, Niger and Burkina Faso are neighboring countries, yet display striking differences in climate and geography which may provide an explanation for the sharp contrast in disease incidences between the countries. Within countries, multiple factors can affect the spread and prevalence of diseases, and determining precisely what these conditions are will be essential in predicting and preventing the next major epidemic.

One major limitation of this study involves the cross-reactivity of human antibodies to multiple diseases. Our observation that the incidences of dengue, chikungunya, and Zika viruses fall within a narrow range could be explained by similarities in vector, but could also be accounted for by known issues of antibody cross-reactivity in serosurveys. For example, antibodies to dengue virus have been shown to consistently bind and neutralize Zika virus, and vice versa. This is due to the high degree of similarity between Flaviviruses such as dengue virus and Zika virus, in terms of sequence and structural homology (Priyamvada et al., 2016). A high
rate of cross-reactivity has also been detected between yellow fever virus and dengue virus, as pre-existing yellow fever virus antibodies have been shown to neutralize dengue virus (Allwinn et al., 2002). While cross-reactivity between chikungunya virus (an Alphavirus) and Flaviviruses is low, it can still occur (Cho et al., 2008). Disease cross-reactivity can potentially explain the observed similarity between disease incidence values, as a person infected with dengue virus may also test positive for Zika virus and yellow fever virus, despite not being infected with these diseases. In addition to cross-reactivity, serosurvey data does not indicate when infection with a disease occurred. Limitations in the methodology of this study also had to be taken into consideration in the analysis. Due to a finite number of published serosurveys, small sample sizes were obtained for some countries, and samples were only taken from certain regions within countries. The search terms used in the literature review were also unable to identify every published serosurvey. While these obstacles were acknowledged, they were unavoidable and did not affect the outcomes of the study.

This analysis will hopefully add to the understanding of emerging and resurging arboviral diseases and highlight the need for continued serosurveys. Serosurveys are an important tool for documenting the spread of disease and lend insight into historical patterns of disease incidence. Expansion of this dataset will be necessary to broaden the reach of this study, in terms of increasing sample sizes and including more countries in the analysis. Inclusion of South American countries would be another interesting option to consider, as arbovirus epidemics in South America are a more recent phenomenon. Dengue virus first emerged in South America in the 1990s, while chikungunya virus emerged in 2013 and Zika virus in 2015 (Fauci and Morens, 2016; Zanluca et al., 2015). Given the recent nature of these epidemics, we would expect to see trends in serosurvey data like those of dengue virus in Indonesia and Laos, with higher disease
incidence in the <15 age category than the >15 age category. Most importantly, discovering trends in disease incidence and determinants of disease spread may enable increased preparation for epidemics. This analysis can potentially reveal what conditions are most important in facilitating arbovirus epidemics, and thus serve as a predictor for where the next major outbreak may occur. Investing in increased prevention and preparation is an important priority, but targeted strategies must be developed to minimize associated costs. Using this analysis as a predictor can help determine the high-risk areas where these strategies should be employed, with an ultimate goal of reducing the public health impact of arboviruses and saving human lives.

Citations


Figure 1: Barplot of mean seropositivity of Zika virus in 51 countries; the absence of a bar for a country indicates a lack of data.
Figure 2: Heat map of Zika virus seropositivity by country, darkness of colors indicates seropositivity ranging from 0 to 100%.
Figure 3: Barplot of mean seropositivity of dengue virus in 51 countries; the absence of a bar for a country indicates a lack of data.
Figure 4: Heat map of dengue virus seropositivity by country, darkness of colors indicates seropositivity ranging from 0 to 100%.
Figure 5: Barplot of mean seropositivity of chikungunya virus in 51 countries; the absence of a bar for a country indicates a lack of data.
Figure 6: Heat map of chikungunya virus seropositivity by country, darkness of colors indicates seropositivity ranging from 0 to 100%.
Figure 7: Barplot of mean seropositivity of yellow fever virus in 51 countries; the absence of a bar for a country indicates a lack of data.
Figure 8: Heat map of yellow fever virus seropositivity by country, darkness of colors indicates seropositivity ranging from 0 to 100%.
Figure 9: Barplot of mean seropositivity of four diseases by age category; red bars correspond to the >15 age category and gray bars correspond to the <15 age category.