A framework to monitor changes in transmission and epidemiology of emerging pathogens: Lessons from Nipah virus

Running title: Monitoring Nipah virus epidemiology

Birgit Nikolay[1], Henrik Salje[1], A.K.M. Dawlat Khan[2], Hossain M.S. Sazzad[2], Syed M Satter[2], Mahmudur Rahman[2], Stephanie Doan[3§], Barbara Knust[3§], Meerjady Sabrina Flora[4], Stephen P Luby[5], Simon Cauchemez*[1], Emily S Gurley*[2,6]

[1]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, UMR2000, CNRS, 75015 Paris, France.

[2]Infectious Diseases Division, icddr,b, Dhaka, Bangladesh

[3]Centers for Disease Control and Prevention, Atlanta, Georgia, USA

[4]Institute of Epidemiology Disease Control and Research, Dhaka, Bangladesh

[5]Infectious Diseases and Geographic Medicine Division, Stanford University, Stanford, California, USA

[6]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

*co-senior authors

[§]The opinions expressed in this article are the authors' own and do not reflect the view of the Centers for Disease Control and Prevention, the Department of Health and Human Services, or the United States government.

All co-authors declare no conflict of interest.

Corresponding author: Henrik Salje (hsalje@pasteur.fr)

Word count: Abstract 144 words; Main text: 2769 words

Abstract

It is of uttermost importance that the global health community develops the surveillance capability to effectively monitor emerging zoonotic pathogens that constitute a major and evolving threat for human health. Here, we propose a comprehensive framework to measure changes in (i) spillover risk, (ii) interhuman transmission, and (iii) morbidity/mortality associated with infections based on six epidemiological key indicators derived from routine surveillance. We demonstrate the indicators' value for the retrospective or real-time assessment of changes in transmission and epidemiological characteristics using data collected through a long-standing, systematic, hospital-based surveillance system for Nipah virus in Bangladesh. We show that while interhuman transmission and morbidity/mortality indicators were stable, the number and geographic extent of spillovers varied significantly over time. Such combination of systematic surveillance and active tracking of transmission and epidemiological indicators should be applied to other high risk emerging pathogens to prevent public health emergencies.

Key words: Nipah virus, Emerging pathogens, Surveillance, Monitoring

Introduction

Emerging zoonotic pathogens represent an important and growing risk to humans. As cross-species interactions increase at the human-animal interface, more opportunities arise for pathogens to spillover. Large-scale epidemics of human immunodeficiency virus (HIV), Ebola, or Middle East respiratory syndrome coronavirus (MERS-CoV) can trace their roots back to spillovers from a zoonotic reservoir [1]. The timely deployment of targeted interventions to prevent public health emergencies would clearly benefit from the real-time evaluation of temporal trends suggestive of altered spillover risk or evolutionary changes linked to increased transmission or disease. Given that many emerging pathogens spillover infrequently, changes in pathogen characteristics may go undetected unless there exist systematic efforts to track these changes. Enhanced monitoring approaches that go beyond tracking case numbers to identify changes in disease risk and underlying mechanisms would therefore help improve global preparedness and response capacities. The International Health Regulations and Global Health Security Agenda both call for stronger surveillance and trend monitoring [2,3]. It has however rarely been possible to develop a comprehensive monitoring framework for emerging zoonotic pathogens due to the lack of stable surveillance collecting detailed case information.

Nipah virus (NiV) is an emerging zoonotic pathogen found in fruit bats throughout South and Southeast Asia. NiV is considered by the World Health Organization as an important health threat to humans due to the severity of disease, the absence of treatments or vaccines, and its ability to be transmitted between people [4,5]. Bangladesh is the only country reporting regular spillovers of NiV, which are nearly exclusively detected through a systematic, hospital-based surveillance system, implemented in 2007 following the ad-hoc identification of ~100 cases during 2001-2006 [6]. The surveillance system is based on three sentinel hospitals where all meningoencephalitis cases are routinely tested for NiV. Identification of a NiV case triggers detailed investigations in affected communities including the identification of transmission networks and the tracing of contacts. The

hospital-based surveillance system leverages clinical services and is therefore less expensive and easier to maintain than other population-based surveillance systems.

Here we propose six epidemiological indicators of emerging zoonotic pathogens and their transmission that can be derived from such routine surveillance data to measure changes in (i) spillover risk, (ii) interhuman transmission, and (iii) morbidity/mortality associated with infections (Figure 1A). Although many of these measures are commonly used in epidemiological studies, their role in the routine monitoring of emerging infectious diseases is not yet defined. The NiV surveillance system in Bangladesh provides a unique opportunity to collect comparable data on cases and their characteristics over time. The objective of this study was to demonstrate how these indicators can be obtained from routine surveillance data and used to monitor changes in disease transmission and epidemiology to inform future public health policy and practice.

Methods

Establishing baseline measures and assessing historical changes

We used data from the systematic NiV surveillance system collected during 2007-2018 to establish baseline measures of the six indicators and assess historical changes over time. Details of the surveillance methods have been previously published [7,8], but in short, at three tertiary care hospitals, all patients admitted with signs and symptoms consistent with febrile neurological illness during December – March had clinical and epidemiological information and serum collected for testing for IgM antibodies against NiV. We further assessed all six indicators using data collected during 2001-2006, a time period before routine hospital-based surveillance was implemented and cases were not systematically detected (see Supplementary Material). During this time period, cases were detected through community investigations following the notification of case clusters and NiV cases not associated with case clusters were likely to be missed.

1. Spillover risk

In Bangladesh, humans usually acquire NiV through drinking date palm sap contaminated with the virus by fruit bats [4]. It is important to monitor spillover frequency since any increase may result in higher disease burden, or may indicate emergence of new transmission pathways. In addition, each spillover event represents an opportunity for more transmissible or virulent strains to emerge. A spillover event is defined as a single case infected from the reservoir or a cluster of cases that can be traced back to a single spillover source (i.e., one or several individuals infected through a contaminated palm sap pot plus subsequent cases infected by interhuman transmission) and the annual spillover rate as the number of spillover events in a given transmission season (December of the previous year until May of the given year). We assumed that spillover events follow a Poisson distribution and estimated the spillover rate observed in a given year with Poisson exact 95% confidence intervals (95%CI) and the average annual spillover rate during a given time period (i.e. for the duration of stable surveillance or time periods between change points) with approximate Poisson 95%CIs. We tested the statistical significance of changes in spillover frequency pre and post a specific time point using exact Poisson test, where we moved the time point by one-year increments over the duration of stable surveillance. If several time points resulted in significant differences, we selected the time point resulting in the lowest p-value. We further assessed linear changes in annual spillover rates using Poisson regression; we evaluated statistical significance of linear changes using a likelihood ratio test and compare this model to step-change models using the Akaike Information Criterion (AIC).

2. Geographic extent

The geographic extent of case occurrence informs the size and location of the population at risk. Moreover, pathogen introductions into new ecological areas or populations (e.g. higher population density, mobility) can affect transmission dynamics and potentially boost the spread and impact of the pathogen [9]. Using surveillance data on geographic location of cases we can monitor the geographic extent of NiV risk. The geographic extent of case reports is defined as the number of

districts from which cases are reported, we quantified the geographic extent using 3-year sliding windows over the duration of the surveillance dataset and assessed spatial patterns based on administrative boundaries obtained from the Database of Global Administrative Areas (www.gadm.org).

3. Cluster size

The NiV surveillance data can also be used to monitor indicators for changes in spillover mechanisms. The number of individuals who are infected from the reservoir during each spillover event, may vary depending on the spillover source. For example, clusters associated with drinking palm sap may differ in size from clusters associated with pig exposure, which was the main spillover route in the Malaysian NiV outbreak [10]. Cluster sizes are also affected by interhuman transmission through the number of secondary cases associated with a spillover event (see section below). We defined a cluster of cases as individuals who acquired infection through a single spillover event (either through one or more bat-to-human transmission events in one time and place or through subsequent interhuman transmission). We estimated the median and interquartile range (IQR) of cluster sizes for a given year, or a given time period. We used Wilcoxon rank sum test to assess changes in outbreak sizes by comparing cluster size distributions pre and post a specific time point that we moved along the time of stable surveillance.

4. Reproduction number

The reproduction number R (i.e., the average number of individuals infected by a case) is the standard measure of the interhuman transmission potential of a pathogen [11]. While sustained transmission of a pathogen in a population can occur only if $R \geq 1$, even increases in R to values <1 can lead to larger cluster sizes with transmission ceasing only after a substantial number of transmission generations. Tracking this parameter can therefore provide an early warning sign of a change in transmission [12–14]. For NiV in Bangladesh, this essential parameter can be directly estimated from the transmission trees that are reconstructed during outbreak investigations. We estimated R and

95%CIs based on the observed number of secondary cases caused by a case, assuming that the number of secondary cases follows a negative binomial distribution with mean R and a dispersion parameter k [14]. We used a likelihood ratio test to assess changes in R pre and post a specific time point that we moved along the duration of stable surveillance. We further assessed linear changes in R by year using negative binomial regression.

5. Proportion of cases who transmit infection

Interhuman transmission of emerging zoonotic pathogens is often highly heterogeneous, meaning that a large part of secondary cases originates from only a few superspreading events as previously noted for NiV [8,15]. The mechanisms leading to superspreading events can be of social (e.g., a higher number of contacts resulting in more transmission opportunities) or biological nature (e.g., higher levels of virus shedding or stronger symptoms facilitating transmission) [14]. The extent of transmission heterogeneity has important implications for the optimisation of control strategies [14]. Where detailed information on transmission trees is available, we can quantify transmission heterogeneity as the percentage of cases who transmit the pathogen [14]. For a given reproduction number, a smaller percentage of cases who transmit will indicate a stronger transmission heterogeneity [14]. We estimated the proportion of spreaders among NiV cases and exact binomial 95%CIs for a given year, or a given time period. We used an exact binomial test to assess changes in the proportion of spreaders pre and post a specific time point that we moved along the time of stable surveillance. We further assessed linear changes in proportion spreaders by year using logistic regression. We also assessed the probability of observing multiple superspreading events (i.e. a NiV case infecting ≥5 individuals) during outbreaks of various sizes (Supplementary material).

6. Morbidity and mortality

Changes in the case fatality ratio (CFR) and the prevalence of specific symptoms such as difficulty breathing (a symptom previously found associated with interhuman transmission[6]), can indicate an adaptation of a pathogen to the human host. We estimated the case fatality ratio (i.e. the proportion

of cases who died among cases) and exact binomial 95%CIs for a given year, or a given time period, separately for cases infected by the reservoir (primary cases) and cases infected through interhuman transmission (secondary cases). We used an exact binomial test to assess changes in the case fatality rate pre and post a specific time point that we moved along the time of stable surveillance. We further assessed linear changes in the case fatality rate by year using logistic regression. We applied the same methods to assess changes in the proportion of cases with difficulty breathing.

Real-time monitoring of indicators

Changes in the transmission and epidemiology of NiV can also be assessed in real-time using the proposed framework by plotting new observations against baseline distributions of indicators established above. To facilitate identification of epidemiologic changes prospectively within this system, we identified thresholds for "highly unlikely" events by quantifying the 2.5th and 97.5th percentiles of these baseline distributions.

Results

Monitoring spillover risk

During 2007-2018, 76 NiV spillover events were detected in Bangladesh resulting in 166 NiV cases (Figure 1B). We observed that the average annual number of spillovers detected in Bangladesh varied significantly over time; it increased from 2.2 (95%CI 1.3-3.7) in 2007-2009 to 10.5 (95%CI 8.2-13.4) in 2010-2015, and returned to the lower level in 2016-2018 (Figure 2A).

Coinciding with the temporary increase in spillover frequency, we found that the number of districts reporting cases was more than two times greater in 2011-2015 compared to the other time windows (Figure 2B). This increase however does not seem to reflect a gradual spread of NiV into previously unaffected areas; instead there were more regular spillovers in geographic regions where NiV had previously been observed (Figure S1). Two NiV cases who acquired infections in rural areas were reported in Dhaka, the capital city of Bangladesh.

We observed a constant median cluster size of 1 case (Interquartile range [IQR] 1-2) over time (Figure 2C). The median cluster size was higher in 2008 (with two clusters of 6 and 4 spillover cases, respectively) than in other years.

Monitoring interhuman transmission

Transmission potential of NiV did not vary significantly in Bangladesh during 2007-2018, with an average reproduction number of 0.20 (95%CI 0.10-0.40) (Figure 2D). No interhuman transmission events were observed in 2008-2009 and 2015-2018, which is consistent with the estimated average reproduction number and may be explained by a smaller number of NiV cases during these years (Figure S2).

Eight percent (95%CI 5-14) of cases transmitted NiV to another person, which was stable over time (Figure 2E). During 2007-2018, two outbreaks were driven by a superspreading event. The occurrence of more than one superspreading event in a single outbreak has never been observed and is unlikely given the current reproduction number (Figure S3).

Monitoring morbidity/mortality associated with infections

We observed a stable CFR over time among cases infected by the reservoir (86%, 95%CI 79-92) (Figure 2F) and those infected through interhuman transmission (46%, 95%CI 29-63) (Figure S4A). The percentage of NiV cases who developed breathing difficulties also remained constant over time with an average of 57% (95%CI 49-64) (Figure S4B). No significant difference in the proportion with breathing difficulties were detected between cases infected by the reservoir and those infected through interhuman transmission (chi-squared p-value= 0.37). The CFR was higher among cases with breathing difficulty (90%) than those without (61%, chi-squared p-value <0.001).

Based on data collected during 2001-2006, a period before routine hospital-based surveillance was implemented and cases were not systematically detected, we observed a much higher variability in

some of these indicators (in particular for spillover frequency, outbreak size, and reproduction number) (Figure S5).

Real-time monitoring of indicators

The baseline measures for these six indicators that we established here can be used by public health officials to detect future changes in NiV characteristics (Figure S6). Such changes can be evaluated in real-time by comparing expected values (e.g., percentiles of estimated distributions) to new observations, either for a single outbreak or for a transmission season. For example, based on current estimates of spillover frequency, observing more than 17 spillovers in a season is very unlikely (based on the 97.5th percentile of the high spillover frequency period), while half of the time, more than 10 spillovers are expected to occur (Figure S6A).

Discussion

Systematic NiV surveillance in Bangladesh has been an important step forward in the response to the emerging threat of NiV. These routinely collected data enable the establishment of baseline measures that comprehensively describe current NiV spillover risk, interhuman transmission and morbidity/mortality, which can now be used by public health authorities to guide rapid and reliable decisions to respond to NiV outbreaks. Using these baseline measures in real-time assessments allows the detection of deviations from these at early stages of a transmission season or an ongoing outbreak.

While interhuman transmission and morbidity/mortality of NiV has been stable since the start of systematic surveillance, frequency and geographic extent of spillovers temporarily increased during that time period before returning to initial levels. The temporary increase may reflect NiV transmission and shedding dynamics in bat populations [16], changes in bat behaviour (e.g., moving to new areas to look for food), or changes in human behaviour (e.g., sap drinking, access to hospitals). A study of weather associations with spillovers between 2007 – 2013 identified a significant correlation between increasing spillovers and colder winter temperatures, suggesting that

climatic factors could be driving some of these patterns over longer time scales [17]. The monitoring framework has thereby provided insights that can form the basis of studies to test different hypotheses of what drives these patterns, and measure changes in patterns that could result from primary or secondary prevention programs. Continuing surveillance efforts to monitor time trends or cyclical dynamics in this disease system are critical and should be prioritized for funding.

Observed cluster sizes were stable over time, suggesting that no changes in spillover mechanisms had occurred during the time period of surveillance. This finding is consistent with other epidemiological evidence that identified palm sap as main spillover route over time, including during spillovers in 2008 that were larger in size [18,19]. No increase in the interhuman transmission potential has been suggested by any of the indicators and the reproduction number currently is lower than what is required for a large outbreak. Based on the established baseline measures, the occurrence of multiple superspreading events during a single outbreak is highly unlikely and may therefore represent an early warning sign for the emergence of a more transmissible strain. For monitoring situations where details on transmission networks are unavailable, R could also be derived from the distribution of cluster sizes or the proportion of cases infected by the reservoir [13,20]. The detection of two NiV cases in Dhaka, the capital of Bangladesh, demonstrates the risk of virus introductions into densely populated areas where large-scale outbreaks may be more likely to occur. The observed heterogeneity in some indicators during 2001-2006, a time period before hospital-based surveillance started in Bangladesh, highlights that systematic surveillance for NiV, as introduced in Bangladesh in 2007, is key for establishing reliable baseline estimates of transmission characteristics and will remain crucial for the detection of departures from these trends in the future. Based on established estimates of R it would however be highly unlikely to observe superspreading events as reported before 2007 (with 11 and 22 secondary cases generated by a NiV case) (Figure S6C).

The analysis of the collected epidemiological data also provided other insights that can be investigated in future studies, such as the difference in the CFR between cases infected through the reservoir and those infected through interhuman transmission. This difference may be due to a higher dose of virus received through date palm sap than through contact with a patient.Developing capacities to efficiently detect and respond to unusual public health events is key for improved global epidemic preparedness [2,3]. The World Health Organization in 2015 therefore advocated for prioritizing a number of emerging zoonotic diseases for urgent research and development that would allow for improved disease control. In addition to NiV, these diseases included Crimean Congo haemorrhagic fever, Ebola, Marburg, Lassa fever, Middle East respiratory syndrome and other coronaviruses, and Rift Valley fever [5]. We believe that systematic surveillance for these diseases, as exemplified by NiV surveillance in Bangladesh, should be implemented to enable the collection of comparable and reliable data over time. The framework proposed here would also allow tracking outbreaks of these diseases and targeting disease control measures. The global community would be much better prepared for these threats through investments in systematic surveillance coupled with active tracking and reporting of transmission and epidemiological indicators.

Acknowledgements

icddr,b acknowledges with gratitude the commitment of the National Science Foundation, the National Institutes of Health (NIH) (NIH grant for Ecology and Evolution of Infectious Diseases under grant no. 2R01-TW005869), and the US Centers for Disease Control and Prevention (US CDC) to its research efforts. icddr,b is thankful to the Governments of Bangladesh, Canada, Sweden, and the UK for providing core/unrestricted support. B. N., H. S., and S. C. acknowledge the support of the Laboratory of Excellence Integrative Biology of Emerging Infectious Diseases (Grant ANR-10-LABX-62-IBEID), the National Institute of General Medical Sciences Models of Infectious Disease Agent Study Initiative, the AXA Research Fund and the INCEPTION project (PIA/ANR-16-CONV-0005). The donors had no role in preparation or review of this manuscript.

References

1. Wolfe ND, Dunavan CP, Diamond J. Origins of major human infectious diseases. Nature. **2007**; 447(7142):279–283.

2. WHO | International Health Regulations (IHR) [Internet]. WHO. [cited 2019 Feb 18]. Available from: http://www.who.int/topics/international_health_regulations/en/

3. About | Global Health Security Agenda [Internet]. [cited 2019 Feb 18]. Available from: https://www.ghsagenda.org/about

4. Luby SP, Gurley ES, Hossain MJ. Transmission of human infection with Nipah virus. Clin Infect Dis. **2009**; 49(11):1743–1748.

5. WHO | List of Blueprint priority diseases [Internet]. WHO. [cited 2018 Dec 28]. Available from: http://www.who.int/blueprint/priority-diseases/en/

6. Luby SP, Hossain MJ, Gurley ES, et al. Recurrent zoonotic transmission of Nipah virus into humans, Bangladesh, 2001-2007. Emerging Infect Dis. **2009**; 15(8):1229–1235.

7. Naser AM, Hossain MJ, Sazzad HMS, et al. Integrated cluster- and case-based surveillance for detecting stage III zoonotic pathogens: an example of Nipah virus surveillance in Bangladesh. Epidemiol Infect. **2015**; 143(9):1922–1930.

8. Nikolay B, Salje H, Hossain MJ, et al. Transmission of Nipah Virus - 14 Years of Investigations in Bangladesh. N Engl J Med. **2019**; 380(19):1804–1814.

9. Alexander KA, Sanderson CE, Marathe M, et al. What factors might have led to the emergence of Ebola in West Africa? PLoS Negl Trop Dis. **2015**; 9(6):e0003652.

306 10. Parashar UD, Sunn LM, Ong F, et al. Case-control study of risk factors for human infection with
307 a new zoonotic paramyxovirus, Nipah virus, during a 1998-1999 outbreak of severe encephalitis
308 in Malaysia. J Infect Dis. **2000**; 181(5):1755–1759.

309 11. Anderson RM, May RM. Infectious Diseases of Humans: Dynamics and Control. Oxford
310 University Press; 1991.

311 12. Ferguson NM, Fraser C, Donnelly CA, Ghani AC, Anderson RM. Public health. Public health risk
312 from the avian H5N1 influenza epidemic. Science. **2004**; 304(5673):968–969.

313 13. Blumberg S, Lloyd-Smith JO. Inference of R(0) and transmission heterogeneity from the size
314 distribution of stuttering chains. PLoS Comput Biol. **2013**; 9(5):e1002993.

315 14. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual
316 variation on disease emergence. Nature. **2005**; 438(7066):355–359.

317 15. Arunkumar G, Chandni R, Mourya DT, et al. Outbreak investigation of Nipah Virus Disease in
318 Kerala, India, 2018. J Infect Dis [Internet]. [cited 2018 Nov 5]; . Available from:
319 https://academic.oup.com/jid/advance-article/doi/10.1093/infdis/jiy612/5144922

320 16. Plowright RK, Parrish CR, McCallum H, et al. Pathways to zoonotic spillover. Nat Rev Microbiol.
321 **2017**; 15(8):502–510.

322 17. Cortes MC, Cauchemez S, Lefrancq N, et al. Characterization of the Spatial and Temporal
323 Distribution of Nipah Virus Spillover Events in Bangladesh, 2007-2013. J Infect Dis. **2018**;
324 217(9):1390–1394.

325 18. Hegde ST, Sazzad HMS, Hossain MJ, et al. Investigating Rare Risk Factors for Nipah Virus in
326 Bangladesh: 2001–2012. Ecohealth. **2016**; 13(4):720–728.

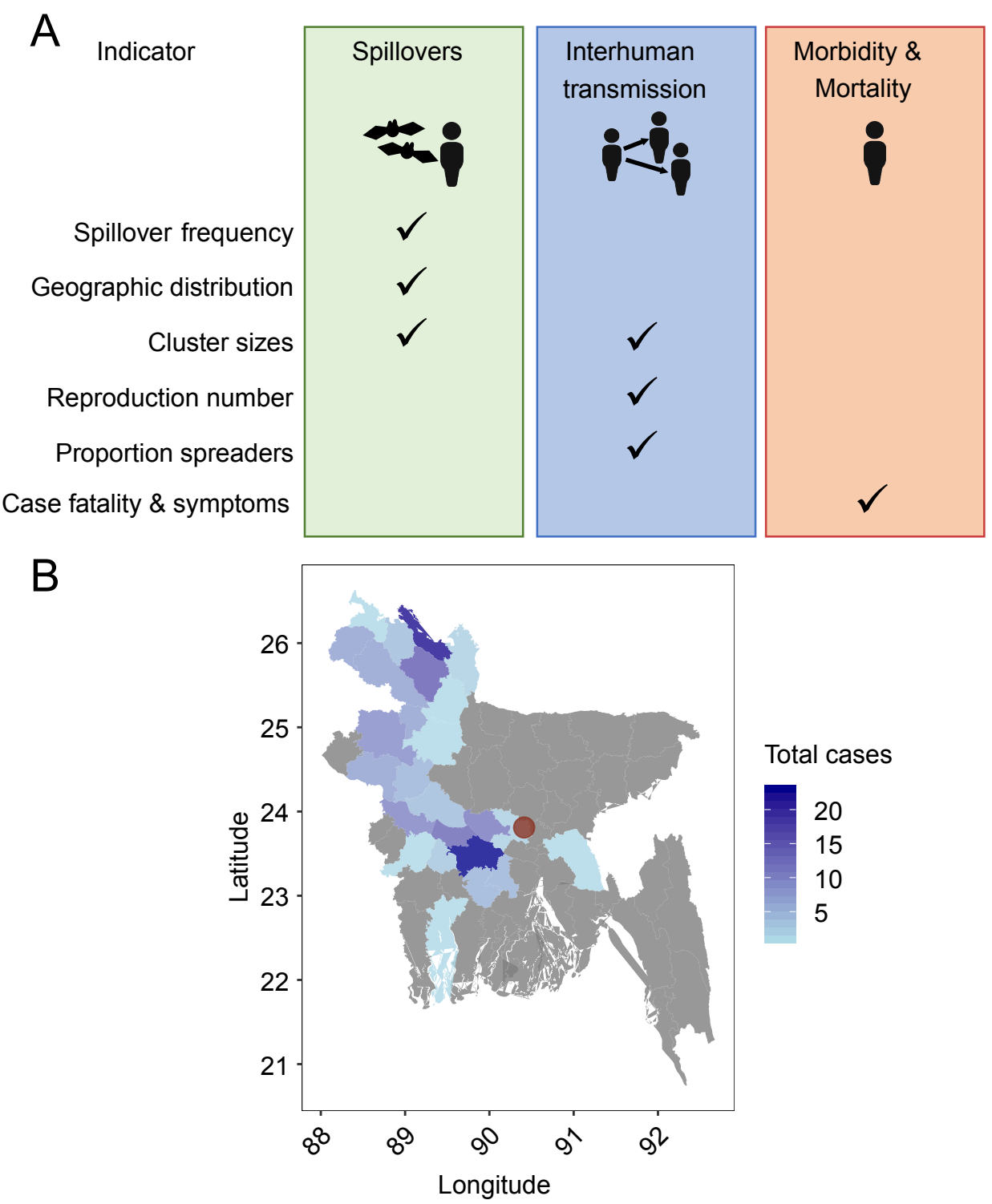
19. Rahman MA, Hossain MJ, Sultana S, et al. Date palm sap linked to Nipah virus outbreak in Bangladesh, 2008. Vector Borne Zoonotic Dis. **2012**; 12(1):65–72.

20. Cauchemez S, Epperson S, Biggerstaff M, Swerdlow D, Finelli L, Ferguson NM. Using routine surveillance data to estimate the epidemic potential of emerging zoonoses: application to the emergence of US swine origin influenza A H3N2v virus. PLoS Med. **2013**; 10(3):e1001399.

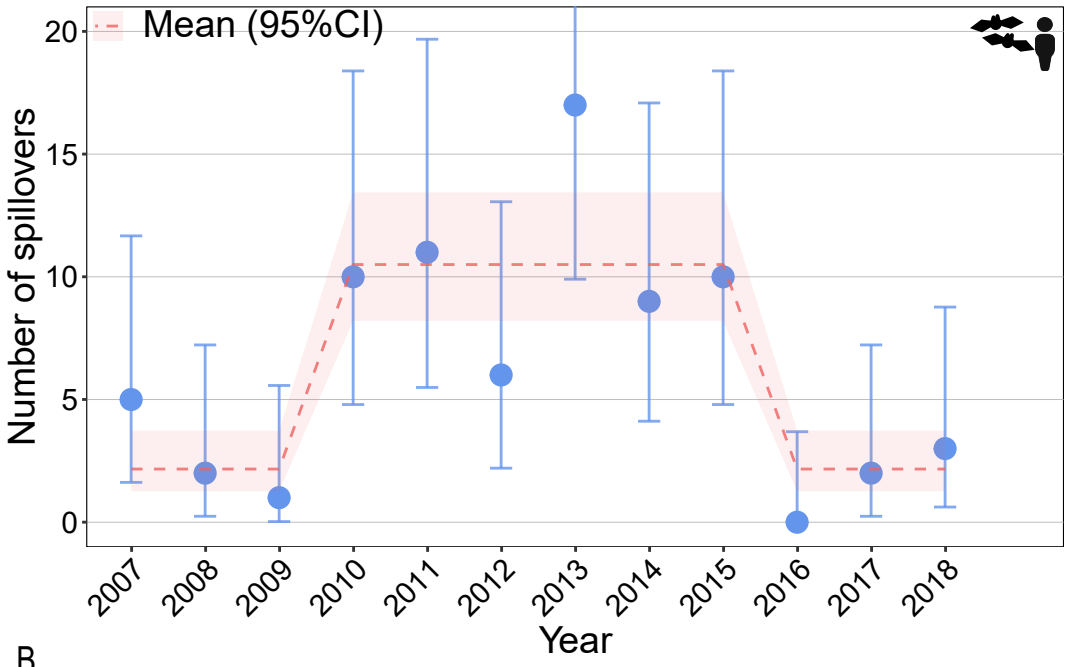
Figure legends

Figure 1. (A) Indicators for changes in spillover risk, interhuman transmission and morbidity/mortality of Nipah virus and other zoonotic pathogens. (B) Total number of Nipah cases ever reported by districts in Bangladesh (2007 to 2018).

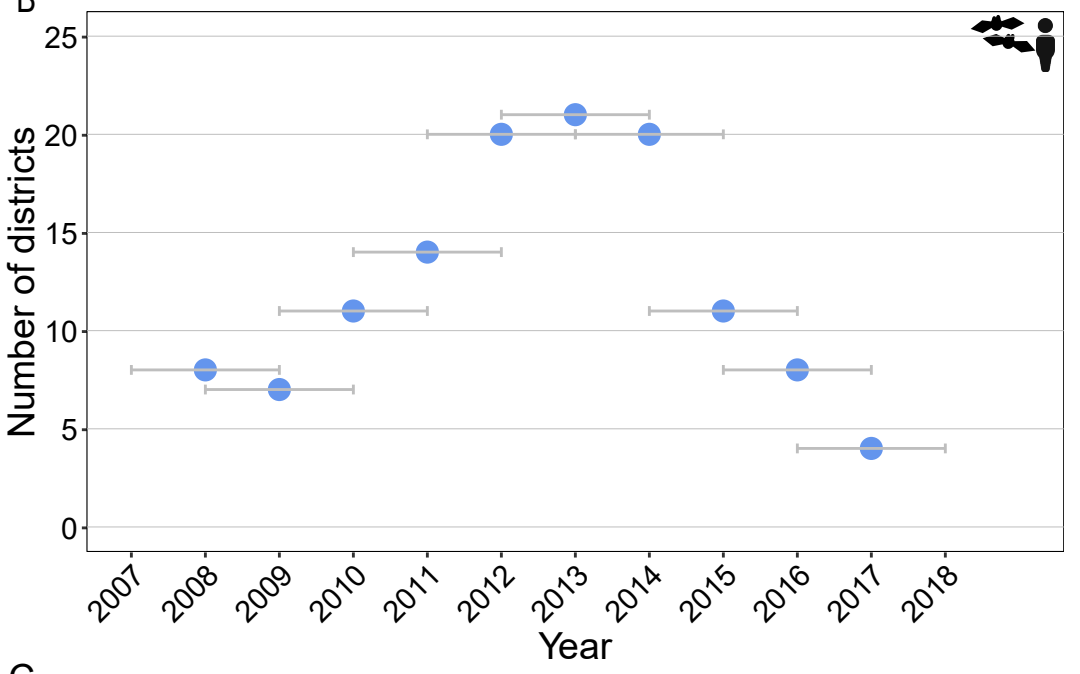
Figure 2. Baseline estimates and historical changes of key indicators to monitor Nipah virus epidemiology and transmission. Average number of spillovers (Poisson exact 95%CI) (A), number of districts reporting NiV cases within 3 year window (time windows are indicated as grey bars) (B), median cluster size (interquartile range) (C), reproduction number (negative binomial 95%CI) (D), proportion spreaders (binomial exact 95%CI) (E), CFR among primary cases, i.e. cases infected by the reservoir (binomial exact 95%CI) (F). The year represents a transmission season, which lasts from December of the previous year until May of the indicated year.


A
Indicator
Spillovers
Interhuman transmission
Morbidity & Mortality
Spillover frequency
Geographic distribution
Cluster sizes
Reproduction number
Proportion spreaders
Case fatality & symptoms
B
Latitude
26
25
24
23
22
21
88
89
90
91
92
Longitude
Total cases
20
15
10
5

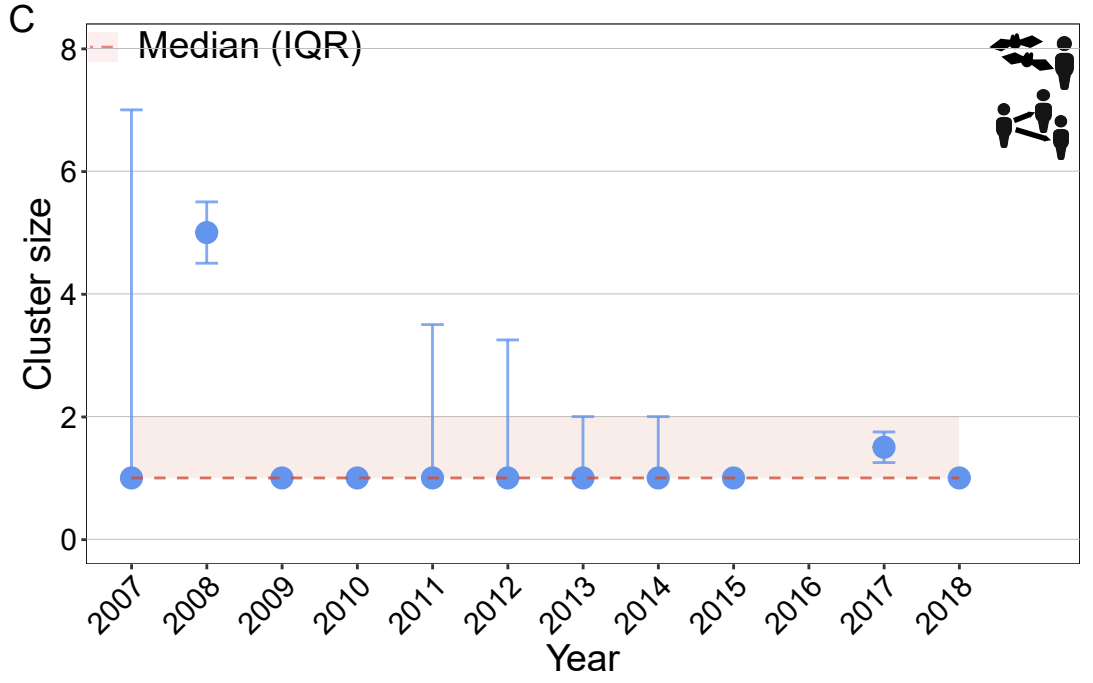
A


Mean (95%CI)
Number of spillovers
Year

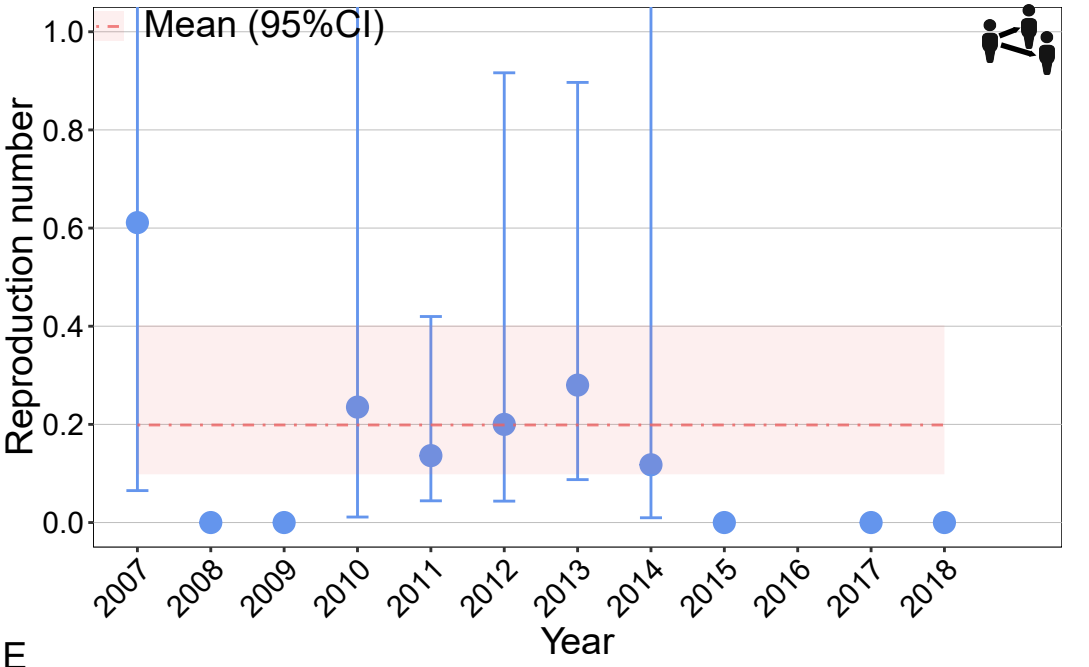
B


Number of districts
Year

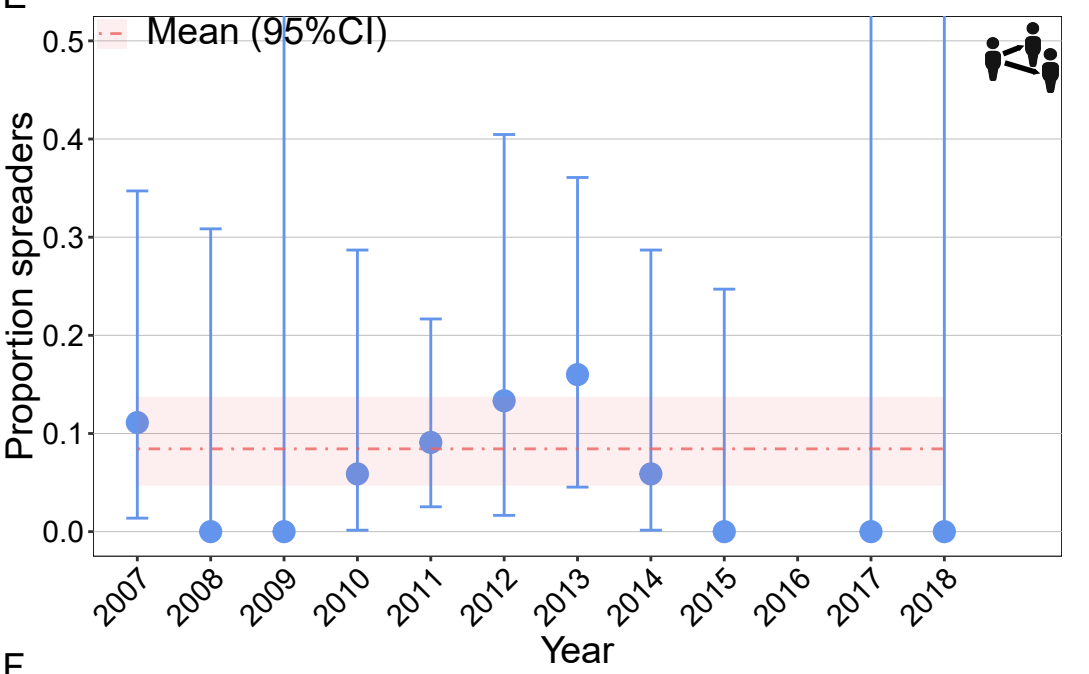
C


Median (IQR)
Cluster size
Year

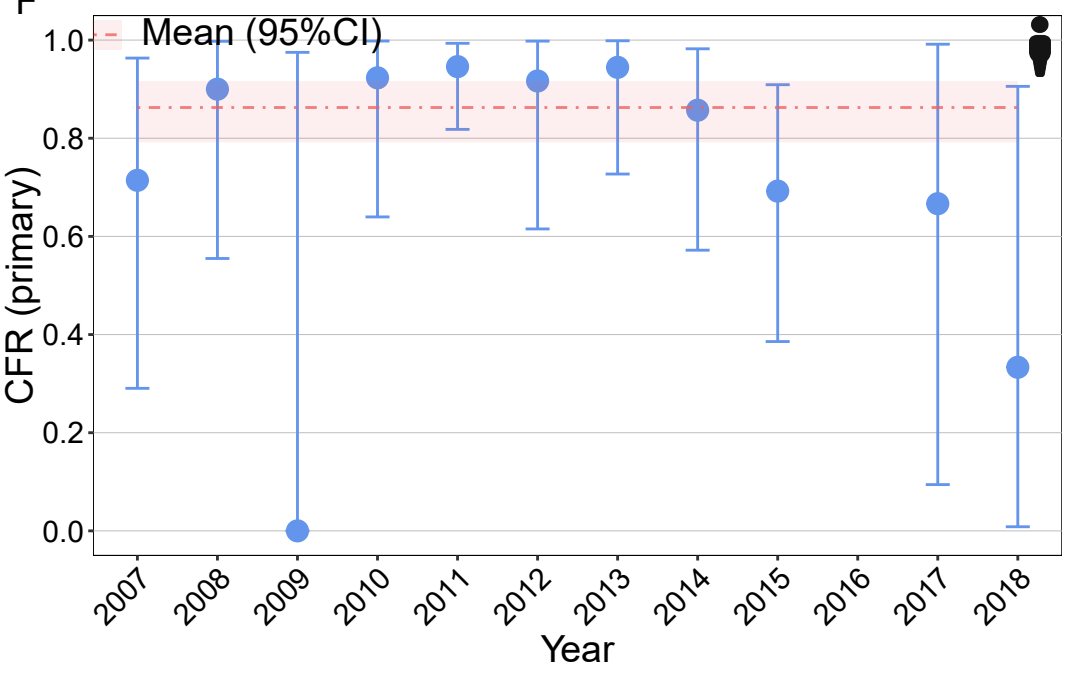
D


Mean (95%CI)
Reproduction number
Year

E


Mean (95%CI)
Proportion spreaders
Year

F


Mean (95%CI)
CFR (primary)
Year

Supplementary Material

A framework to monitor changes in transmission and epidemiology of emerging pathogens: Lessons from Nipah virus

Table of Contents

1. Probability of no spreaders and superspreaders 17

2. Importance of systematic surveillance for baseline establishment 17

3. Supplementary Figures 18

Erreur ! Signet non défini.

1. Probability of no spreaders and superspreaders

The probability of not observing any NiV spreading events by increasing number of cases is shown in Figure S2. For example, the probability of not observing any spreading event is 41% among 10 NiV cases and 17% among 20 cases.

The probability of observing at least one, at least two, or at least three superspreading events (i.e. a NiV case infecting ≥5 individuals) is shown in Figure S3. Among 10 cases, the probability of observing at least one superspreading event is 10%; however as low as 0.4% and 0.01% for observing at least 2 or 3 superspreading events.

2. Importance of systematic surveillance for baseline establishment

We also estimated the indicators for each year during 2001-2006, a time period before routine hospital-based surveillance was implemented and cases were not systematically detected. During 2001-2006, we observe a much higher variability in some of these key measures: the observed annual spillover frequency ranged from 0 to 14, annual median outbreak sizes from 1 to 13, and the annual estimated reproduction number from 0 to 0.85 (Figure S5). This demonstrates the difficulties for establishing some of these baseline measures in the absence of stable surveillance.

3. Supplementary Figures

Figure S1. Location of districts reporting Nipah cases within moving 3-year time intervals during 2007-2018.

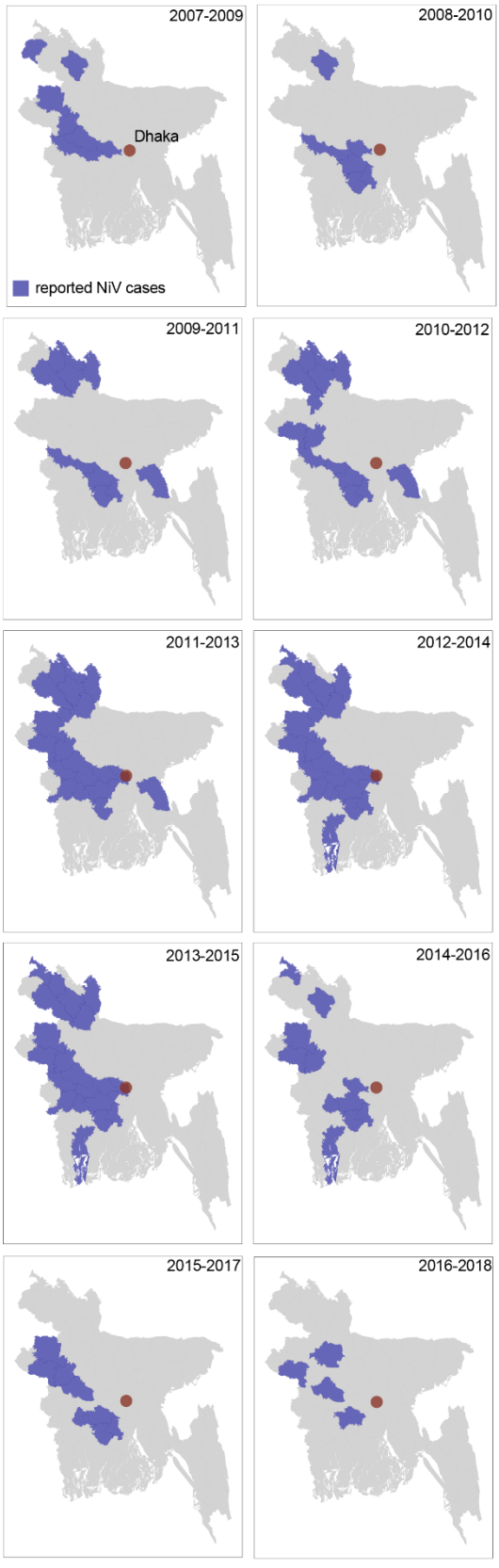


Figure S2. Probability of observing zero spreading events by number of cases.

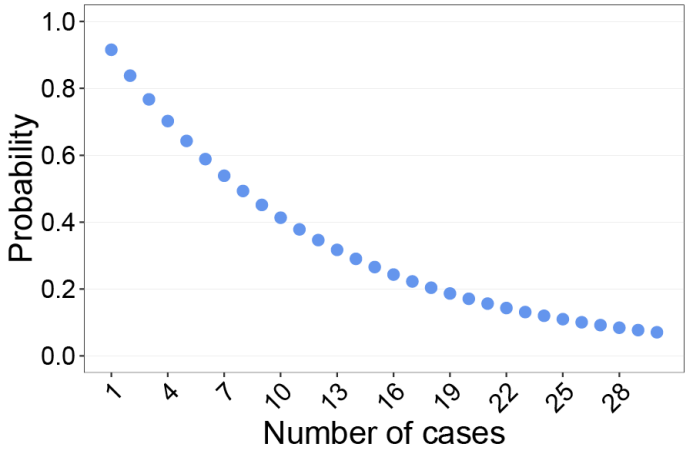


Figure S3. Probability of observing at least 1, at least 2, and at least 3 superspreading events by number of cases.

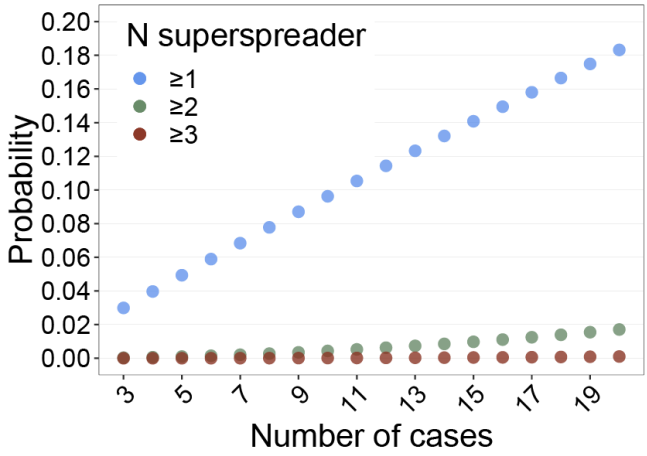


Figure S4. Baseline estimates and historical changes of key indicators to monitor Nipah virus epidemiology and transmission- additional measures. CFR among secondary cases, i.e. cases infected by interhuman transmission (binomial exact 95%CI); no secondary cases occurred during 2015-2018 (A), proportion of cases with difficulty breathing (binomial exact 95%CI) (B). The year represents a transmission season, which lasts from December of the previous year until May of the indicated year.

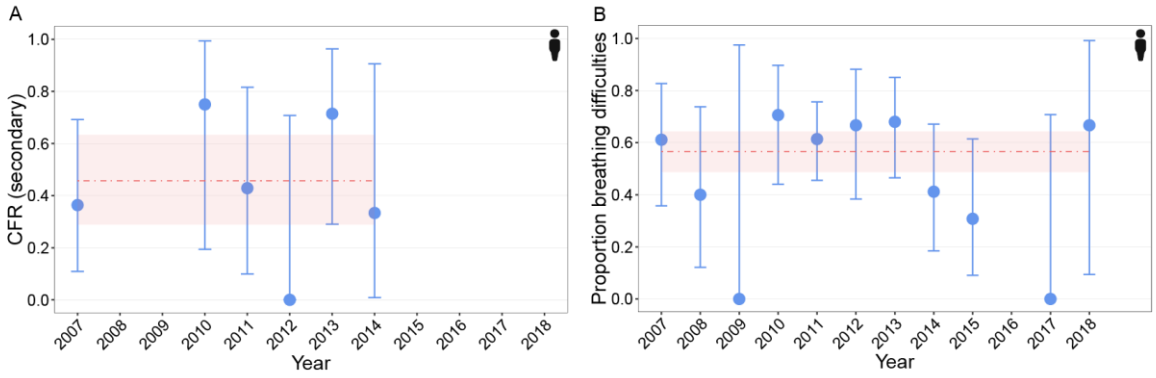


Figure S5. Comparing key indicators before (2001-2006) and after implementation of systematic hospital-based surveillance (2007- 2018). Average number of spillovers (Poisson exact 95%CI) (A), number of districts reporting NiV cases within 3 year window (time windows are indicated as grey bars) (B), median cluster size (interquartile range) (C), reproduction number (negative binomial 95%CI) (D), proportion spreaders (binomial exact 95%CI) (E), CFR among primary cases, i.e. cases infected by the reservoir (binomial exact 95%CI) (F). The year represents a transmission season, which lasts from December of the previous year until May of the indicated year.

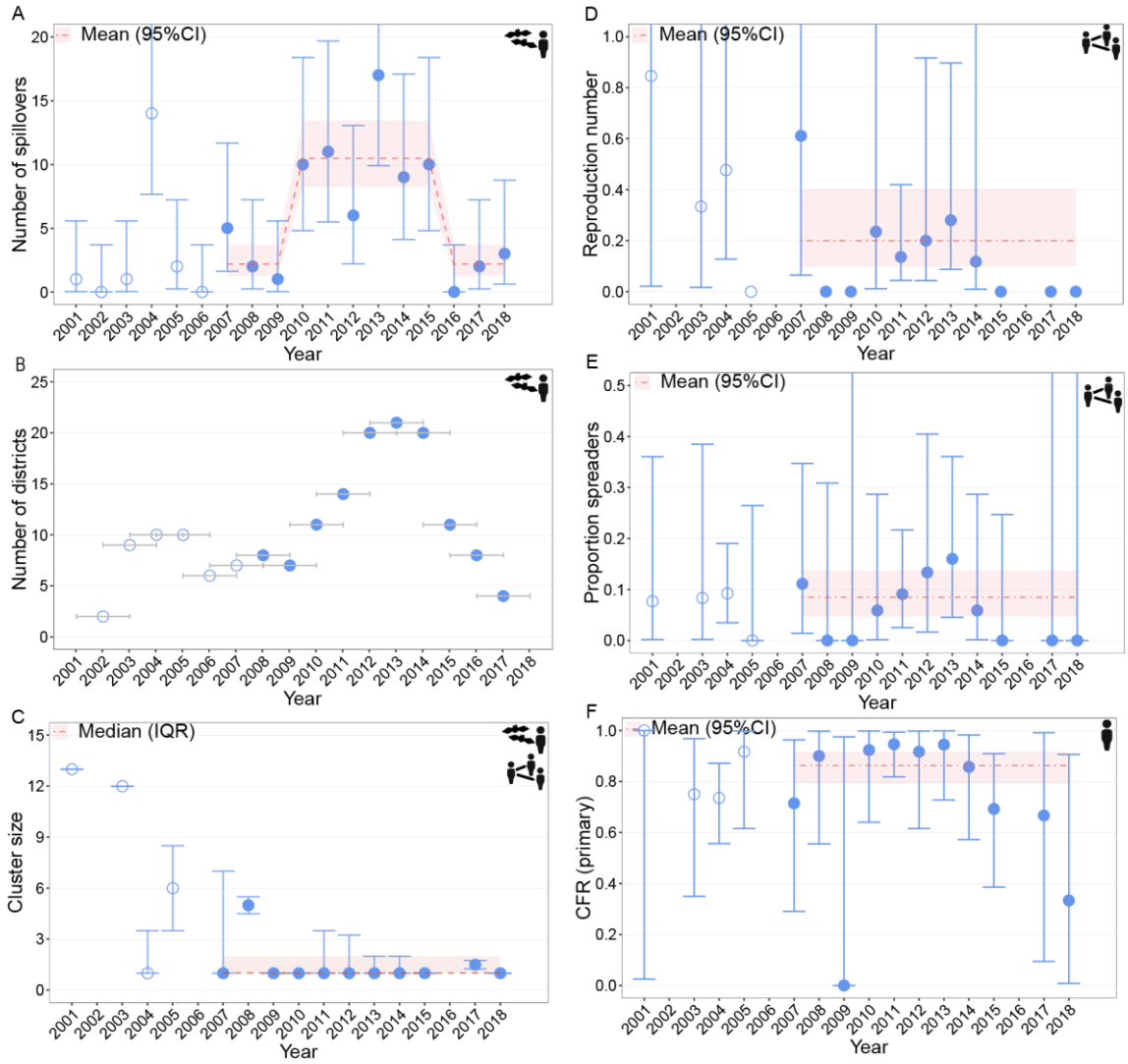


Figure S6. Detecting outliers in Nipah virus transmission and epidemiology using baseline estimates. Distribution of annual number of outbreaks (A). Empirical distribution of observed outbreak sizes (B). Distribution of number of individuals infected by a case (C). Distribution of number of spreaders among 20 cases (D). Distribution of number of deaths among 20 primary cases (E). Distribution of number of cases with difficulty breathing among 20 cases (F). The dotted lines indicate the 50th percentile; the dashed lines indicate the 2.5th and 97.5th percentile.

