

Human movement and transmission dynamics early in Ebola outbreaks

Alexandria Gonzalez¹, Behnam Nikparvar¹, M. Jeremiah Matson², Stephanie N. Seifert³, Heather D. Ross⁴, Vincent Munster⁵, Nita Bharti^{1*}

1 Biology Department and Center for Infectious Disease Dynamics, Penn State University, University Park, PA, 16802, USA

2 University of Utah Health, Department of Internal Medicine, Salt Lake City, UT 84132, USA

3 Paul G. Allen School for Global Health, Washington State University, Pullman, WA, 99164 USA

4 Donald W. Hamer Center for Maps and Geospatial Information, Penn State University Libraries, Penn State University, University Park, PA, 16802, USA

5 Division of Intramural Research, Laboratory of Virology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, Montana, 59840 USA

All authors declare no competing interests.

*Corresponding author: nita@psu.edu

Classification: Biological, Health, and Medical Sciences; Public Health and Epidemiology

Keywords: Movement, Spillovers, Ebola, Roads, Pathogen Transmission

Abstract

Human movement drives the transmission and spread of communicable pathogens. It is especially influential for emerging pathogens when population immunity is low and spillover events are rare. We digitized serial printed maps to measure transportation networks (roads and rivers) in Central and West Africa as proxies for population mobility to assess relationships between movement and Ebola transmission. We find that the lengths of roads and rivers in close proximity to spillover sites at or near the time of spillover events are significantly correlated with the number of EVD cases, particularly in the first 100 days of each outbreak. Early management and containment efforts along transportation networks may be beneficial in mitigation during the early days of transmission and spatial spread for Ebola outbreaks.

Significance Statement

This study links human movement and pathogen transmission across fifty years. While this relationship is well understood for modern outbreaks, it has not been characterized at local scales for historical outbreaks. We compared the number of cases and the spatial spread of each documented outbreak of Ebola (*Orthoebolavirus zairense*, EBOV) to the road and river networks surrounding each spillover location at the time of each spillover event. We measured the road and river networks by digitizing a series of paper maps that were printed during or near the year of each spillover. We show that the connectivity of spillover locations is consistently correlated to the severity of the outbreak over time and across all locations of EBOV spillover events.

Introduction

Understanding the relationship between human mobility and the spread of pathogens is crucial for mitigating outbreak events. This has been demonstrated for acute respiratory infections, such as influenza and daily commuter patterns^[1], for measles and large-scale seasonal movements^[2], and for HIV with long term changes in connectivity^[3]. Understanding this relationship can highlight spatial and temporal risks of transmission, which can inform public health preparedness, surveillance, and outbreak response efforts.

The zoonotic origins and high case fatality rate in humans (40-90%) of Ebola virus (EBOV, member of species *Orthoebolavirus zairense*, formerly known as Zaire ebolavirus) are particularly concerning as a disease-causing agent^[4,5]. Possible reservoir hosts include fruit bats while non-human primates and some antelope are known intermediate hosts^[4]. Initial symptoms of Ebola virus disease (EVD) include headaches, high fever, and muscle aches. Disease progression can lead to more severe symptoms, such as internal bleeding and organ failure, which can result in death^[4,6].

The first documented case of EVD caused by the EBOV occurred in Yambuku, DRC, 1976. From 1976-2020, a total of 18 human EVD spillovers were documented, all within Central and West Africa^[7]. A spillover is defined as a cross-species transmission event and often refers to pathogen transmission from wildlife to humans. Many EVD outbreaks are triggered by a single spillover event followed by human-to-human transmission between close contacts. Most spillover events have been traced to rural and heavily forested areas, where human-wildlife interactions could have occurred. Index cases are often linked to hunting activities, or individuals who work in forests or areas of land conversion (e.g., logging, mining, etc.) and may have had contact with bats or non-human primates^[8,9]. Infected humans can transmit the virus directly to other humans with the greatest risk to close contacts, including caregivers, who may be family members or medical professionals^[10].

To identify possible links between human movement and the spread of EBOV, we measured movement and transportation networks across time and compared them to concurrent EVD outbreaks. Specifically, we analyzed characteristics of the road and river networks surrounding the spillover locations near the time of each spillover event for all 18 documented EVD outbreaks from 1976-2020. We quantified the connectivity of spillover locations to surrounding areas. Roads vary by condition and surface (e.g., unpaved roads) and enable transportation by vehicles, animals, and on foot^[11,12]. Rivers are also commonly used for transit, particularly in Central Africa, and link large cities as well as smaller settlements^[12].

We first examined metrics for total outbreaks and next focused on the first 100 days of each outbreak to isolate the immediate importance of movement and connectivity. This temporal scope helped remove the downstream effects of transmission and long-term outbreak management. It also allowed us to isolate the early trajectories of two recent EVD outbreaks that were uncharacteristically large in number of cases, spatial extent, and total duration ^[10].

We also examined proxies for the impact of outbreak response readiness and prior experience by comparing the total outbreak size and duration for the first EBOV outbreaks in an area to subsequent outbreaks in the same location. Using the locations of index cases, we defined ‘subsequent outbreaks’ as EBOV spillover events that occurred in close spatial proximity to a previous spillover event, either in the same town or a neighboring town with prior evidence of connectivity due to movement (< 60 km). We identified three sets (two pairs and one triad) of first and subsequent outbreaks and compared their outbreak metrics.

We found that measures of transportation networks were correlated with several metrics used to measure total outbreaks. Measures of transportation networks included total road length, total combined road and river length, and total number of intersections (road-road, road-river, and river-river intersections). However, this relationship was even stronger for the first 100 days of outbreaks; transportation networks surrounding each EBOV spillover event were significantly positively correlated with the number of reported cases in the first 100 days of an outbreak. This suggests that population mobility is a contributing factor in EBOV transmission. We also found that subsequent outbreaks were always smaller than initial outbreaks, suggesting preparedness or response experience may improve outbreak management. With increasing human mobility and population connectivity, emerging zoonoses present a growing threat to populations and global health systems.

Methods

Data sources

We scanned and digitized printed versions of Michelin Road maps of Central and South Africa that were published in 1963, 1969, 1974, 1981, 1989, 2003, 2007, and 2019 and West Africa from 1975, 1989, 1991, 2003, and 2019 (**Table 1**) ^[13,14]. We used shapefiles of administrative areas from *DIVA-GIS* as reference maps for each country ^[15]. We included all countries with documented Ebola-Zaire virus spillovers or transmission events from 1976 to 2020: Central African Republic, The Democratic Republic of the Congo (DRC), Gabon, Guinea, Liberia, Republic of the Congo (RC), Sierra Leone, and Uganda (Fig. 1a). We used ArcMap *version 10.8.1* to georeference and digitize roads and rivers from each scanned map ^[16]. Maps were georeferenced using the WGS 1984 geographic coordinate system and projected to the Africa Sinusoidal projected coordinate system. Serial road and river networks from different years were registered together to minimize the errors associated with digitizing. Topology corrections were applied on the networks in GRASS GIS *version 7.8.7* ^[17].

We analyzed documented spillover events from the DRC that occurred in 1976, 1977, 1995, 2007, 2008, 2014, 2017, 2018, 2019, and 2020, Gabon in 1994, 1996, 1996(b), and 2001, RC in 2003, 2003(b), and 2005, and Guinea in 2013. We used the map year closest to each spillover event year to perform a detailed analysis on the road and river networks. The greatest time difference between a map year and the corresponding spillover year for an event was 7 years, and the average was 3 years. Table 1 includes spillover event details and the map year that was matched to each spillover event. We excluded outbreaks that were determined by genetic sequence analysis that were caused by transmission from a survivor of an outbreak in years prior; we included only outbreaks that were seeded by spillover events. We also excluded other orthoebolaviruses, which differ in incubation period, transmissibility, clinical presentation, and geographic range.

Data preparation

We measured outbreaks using the total number of cases recorded during the outbreak, the duration of each outbreak in days, and the number of cases in the first 100 days (Fig. 1b) for Yambuku ^[18,19], Tandala, Mekouka ^[20], Kikwit ^[20,21], Mayibout ^[20], Booue ^[20,22], Mekambo/Mbomo ^[23,24], Mbomo ^[25,26], Etoumbi ^[27], Luebo ^[28,29], Meliandou ^[30,31], Inkanamango ^[32], Likati ^[33,34], Bikoro ^[35,36], North Kivu ^[37,38], and Mbandaka ^[39]. We also examined the spatial spread of notified cases within the first 100 days of each outbreak from existing maps and surveillance reports in above cited references.

We performed a sensitivity analysis of correlation between transportation networks and outbreak measures across study areas with each spillover event at the center. We examined square study areas of 50 km x 50 km, 100 km x 100 km, 150 km x 150 km, 200 km x 200 km, and 300 km x 300 km. We assessed correlations between measures of mobility networks and outbreaks at 10 km increments for sensitivity analysis (Fig. S1).

We used estimates of walking distance and speed to determine the study area that would capture movement during the shortest incubation period for EVD. We apply the operational incubation period for EVD during outbreak response, which is 2-to-21-days, although 4-5 day minimum incubation periods are more typical than 2 days ^[40,41]. An average person is able to walk 5 km/hr along roads ^[42]. The estimated travel speed by boat is in the same range as walking speed ^[10]. Thus, at 7.5 hours of walking or traveling by boat per day, a person could potentially travel 75 km in 2 days. We concluded that a study area of 75 km around each spillover, or a 150 km x 150 km area centering each spillover, would adequately represent the minimum critical movement early in an epidemic and we focus our results accordingly (see SI for results from remaining study areas). We calculated the total length of the roads and rivers within each study area. Additionally, we quantified the number of total intersections for each study area where an intersection is defined as a point where at least two routes intersect (e.g road-road, road-river, river-river).

Subsequent Outbreaks

We defined subsequent outbreaks as EBOV spillover events that occurred in the same location or within close spatial proximity, no greater than 60 km apart, from a previous spillover event. These locations also exhibited evidence of connectivity, including pathogen transmission between them and direct links along transportation networks ^[10,43]. We identified three sets of subsequent outbreaks. First, an outbreak in Mekouka, Gabon in December 1994 preceded an outbreak in Mayibout, Gabon in February of 1996. Second, an outbreak in Mbomo, RC in February 2003 preceded both outbreaks in Mbomo in November 2003 and Etoumbi, RC in May 2005. Third, an outbreak in Luebo, DRC in September 2007 preceded an outbreak in Luebo in December 2008 (Fig. 1a).

Statistical Analysis

We calculated Pearson correlation values to assess the relationships between outbreak measures and transportation network characteristics. We used linear regressions and multiple regression models to assess the relationship between each outbreak measure and the transportation network characteristics described. We also added the total road networks and the total river networks separately to a multiple regression analysis against the total cases in the first 100 days of each outbreak for the 150 km x 150 km study areas. We reported R-squared and p-values for each analysis. We used R studio, version 2022.07.1+554 for all statistical analyses ^[44].

Results

We mapped each EBOV spillover event (Fig. 1a). Between 1976 and 2020, 10 EBOV spillover events occurred in the DRC, 3 in the ROC, 4 in Gabon, and 1 in Guinea. The total number of cases for each outbreak ranged from 1 to 28,652 (Table 1) ^[5,7]. The Guinea outbreak of 2013 reported 28,652 cases and a duration of 888 days; this outbreak represents the largest number of total cases and the longest duration of any EVD outbreak, while the Tandala outbreak in 1977 reported the fewest total cases (1 case) and the shortest duration (42 days). We also mapped the corresponding estimated geospatial spread for the first 100 days of each outbreak ^[43]. The total cases in the first 100 days of each outbreak ranged from 1 to 285; the Kikwit outbreak of 1995 reported the most cases in the first 100 days with 285 cases.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

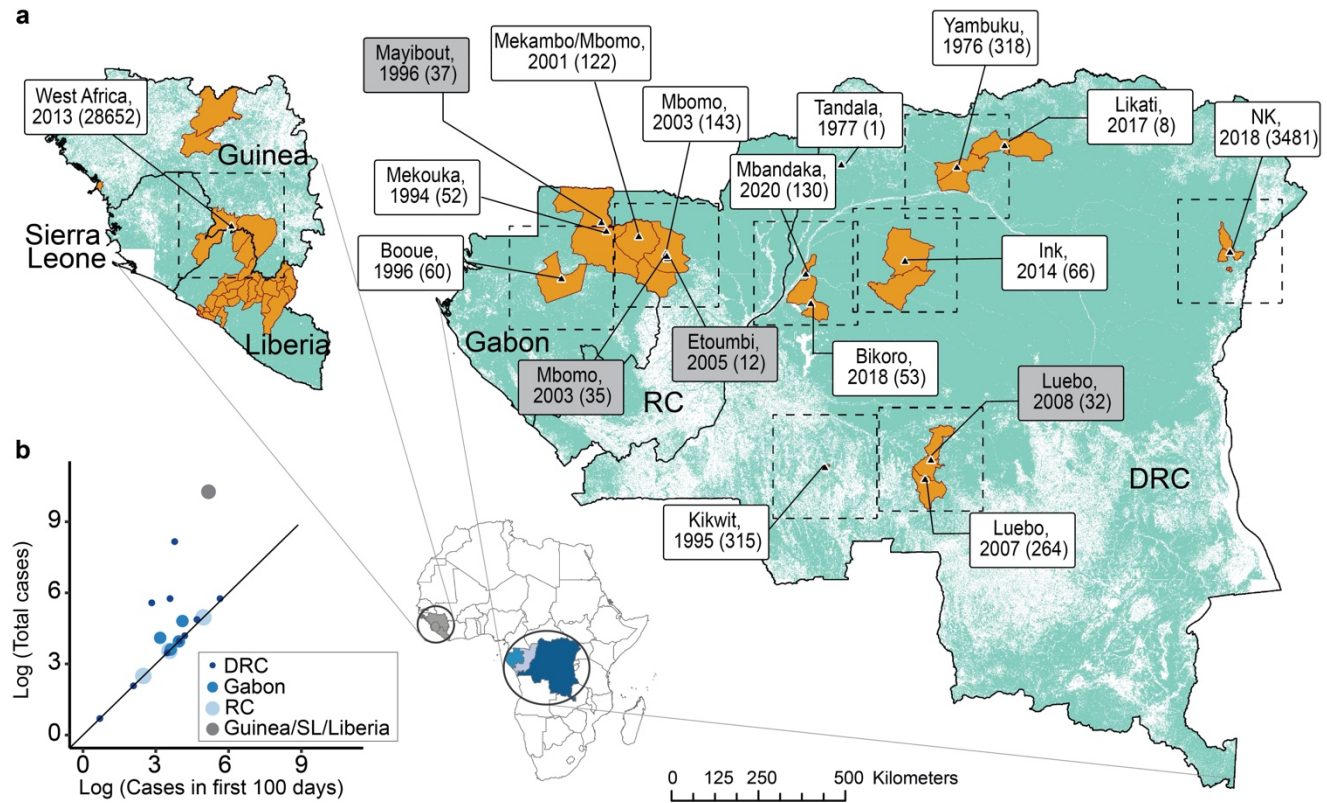


Figure 1. Spillover Events

(a) Locations of EBOV spillovers. Black triangles indicate locations of spillover events to humans, labeled with the location, year, and total cases in parentheses. Labels for subsequent outbreaks are shaded in gray. Orange filled polygons represent the estimated geospatial spread of each outbreak in the first 100 days of the outbreak; the West African outbreak shows the only non-contiguous area. **(b)** Outbreak measures: The total cases and the total cases in the first 100 days of each outbreak.

Spillover year	Map year selected	Spillover event	Spillover location coordinates	Outbreak measures		
				Cases in first 100 days	Total cases	Total duration (in days)
1976	1974	DRC (Yambuku)	2.82535, 22.22567	36	318	112
1977	1974	DRC (Tandala)	2.88874, 19.22384	1	1	42
1994	1989	Gabon (Mekouka) ¹	1.372356, 12.983872	52	52	113
1995	1989	DRC (Kikwit)	-5.039533, 18.80675	285	315	233
1996	1989	Gabon (Mayibout) ¹	1.145606, 13.111622	37	37	90
1996b	1989	Gabon (Booue)	-0.094905, 11.949045	24	60	231
2001	2003	Gabon/RC (Mekambo/Mbomo)	1.011731, 13.952944	60	122	197
2003	2003	RC (Mbomo) ²	0.560156, 14.657393	143	143	162
2003b	2003	RC (Mbomo) ²	0.432044, 14.696622	35	35	94
2005	2003	RC (Etoumbi) ²	0.494444, 14.678611	12	12	81
2007	2007	DRC (Luebo) ³	-5.350627, 21.416457	17	264	174
2008	2007	DRC (Luebo) ³	-4.85, 21.566667	32	32	81
2013	2019	Guinea (Meliandou)	8.616067, -10.061178	176	28652	888
2014	2019	DRC (Inkanamango)	0.366878, 20.870139	66	66	118
2017	2019	DRC (Likati)	3.39, 23.46754	8	8	71
2018	2019	DRC (Bikoro)	-0.747083, 18.429114	53	53	111
2018	2019	DRC (North Kivu)	0.7918, 29.046	44	3481	694
2020	2019	DRC (Mbandaka)	0.0266, 18.2971	110	130	184

Table 1. Summary of EBOV spillover events from 1976–2020 including the year, location, number of cases in the first 100 days, total cases, and total duration of each spillover. The map year selected represents the publication year of the Michelin map that was temporally closest to each spillover event. The three sets (two pairs and one triad) of initial and subsequent outbreaks are indicated with matching numerical superscripts.

Network characteristics

We measured the lengths of the road and river networks around each reported spillover location from the Central/South Africa and North/West Africa Michelin maps from 1963-2019 and 1975-2019, respectively. Across the 150 km x 150 km study areas in Central Africa, from 1963 to 2019, we observed a maximum increase in road network length of 363 km and a maximum decrease in road network length of 120 km for 17 spillover events. The average road network length change was an increase of 53 km (see Table S1 for the net changes in road lengths in a 150 km x 150 km study area surrounding each spillover). The change in road network length from 1975 to 2019 in the 150 km x 150 km study area surrounding the West Africa spillover event was an increase of 133 km. River lengths were largely stable over time within each study area.

We calculated the ratio of river network length to road network length in a 150 km x 150 km study area surrounding each spillover (Fig. 2). The study area surrounding the 1994 outbreak in Mekouka, Gabon contained the highest ratio of river to road length, at 853 km of rivers to 61 km of roads, while the study area around the 1977 spillover event in Tandala contained the lowest ratio of rivers to roads, at 260 km of rivers to 710 km of roads. The Kikwit outbreak event of 1995 had the most road length with 1196 km. There was no temporal relationship between the year of spillovers and the ratio of river to road network length. At larger spatial scales, such as 300 km x 300 km, the ratio of river length to road length was approximately 1:1 for most outbreaks.

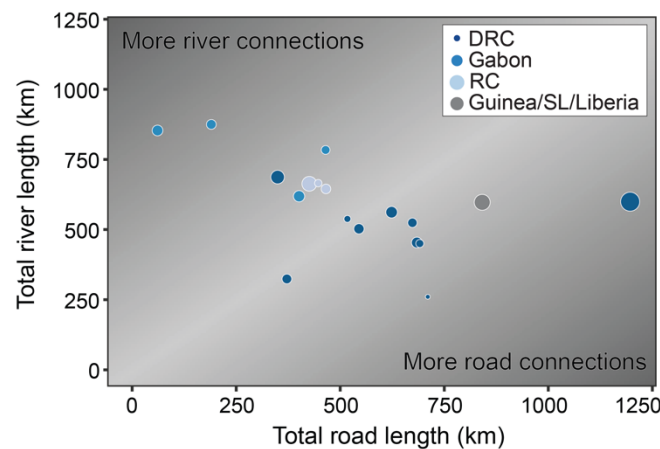


Figure 2. Total road and river network lengths for each spillover event at the time of the spillover event in a 150x150 study area surrounding the spillover. The 1:1 line is shaded in the lightest gray in the background, size of points corresponds to number of cases in the first 100 days of each outbreak, and color corresponds to location of spillover.

Total cases and outbreak duration

We examined the relationship between the transportation networks and outbreak metrics, specifically the total number of cases and total duration in days for each outbreak, using linear regression models. We conducted two separate analyses: one including all spillover events and one excluding the large Guinea and North Kivu spillover events.

All spillover events

In 150 km x 150 km study areas, the R-squared values indicated that total river network lengths, total road network lengths, total combined road and river network lengths, and total number of intersections each explained between 0 - 12% ($p > 0.05$) of the variation in the total cases and total duration in days of the outbreaks. Within the 100 km x 100 km study area, the R-squared values indicated that the total combined road and river network lengths explained 28% ($p = 0.0236$) of the variation in total cases and 35% ($p = 0.0101$) of the variation in total duration in days (Fig. S2). Within the 300 km x 300 km study area, the R-squared values indicated that the total combined road and river network lengths explained 23% ($p = 0.0434$) of

the variation in total cases. We found no significant relationships between transportation networks and outbreak metrics for the remaining study areas (see Table S2 for complete results).

Excluding Guinea and North Kivu spillover events

We found that the two largest and longest EBOV, the Guinea (2013; 28652 cases in 888 days) and North Kivu (2018; 3481 cases in 694 days) events (Table 1), occasionally obscured the relationship between transportation networks and both the total cases and total duration in days. When excluding these outbreaks in 150 km x 150 km study areas, the R-squared values indicated that total river network lengths, total road network lengths, total combined road and river network lengths, and total number of intersections each explained between 4 – 25% ($p > 0.05$) of the variation in the total cases and total duration in days of the outbreaks. Within the 200 km x 200 km study areas, the R-squared values indicated that the total combined road and river network lengths and the total road network lengths each explained 29% ($p = 0.0320$) of the variation in the total cases. Additionally, the total combined road and river network lengths explained 28% ($p = 0.0354$) of the variation in outbreak duration within this study area. When excluding the two largest outbreaks in our analyses of 300 km x 300 km study areas, the R-squared values indicated that the total number of intersections explained 27% ($p = 0.0380$) of the variation in the total cases and the total river length explained 57% ($p = 0.0007$) of the variation in the total duration. See Table S3 for complete R-squared values.

The first 100 days of each outbreak

We examined the relationship between the transportation networks and the total cases in the first 100 days of each outbreak (Table 1 and Table S2-5). The total combined road and river network lengths were strongly correlated with the total number of intersections for all of the study areas examined: 50 km x 50 km, 100 km x 100 km, 150 km x 150 km, 200 km x 200 km, and 300 km x 300 km ($r = 0.81 - 0.94$). We found that the total combined road and river network lengths, explained between 35 – 56% of the variance of the total cases in the first 100 days across all outbreaks for each of the study areas examined: 50 km x 50 km, 100 km x 100 km, 150 km x 150 km, 200 km x 200 km, and 300 km x 300 km (Table S2). We observed similar results with the number of intersections and the total cases in the first 100 days for each of the study areas. The total road network lengths explained between 27 - 36% of the variance of the total cases in the first 100 days of an outbreak for the following study areas: 150 km x 150 km, 200 km x 200 km, and 300 km x 300 km. The total river network lengths explained between 2 - 22% ($p > 0.05$) of the variance of the total cases in the first 100 days of an outbreak for each of the study areas examined. The strongest relationship was between the total combined road and river network lengths and the total cases in the first 100 days for each of the study areas examined.

In 150 km x 150 km study areas, the R-squared values indicated that the total combined road and river network lengths and total road length explained 54% ($p = 0.0005$) and 30% ($p = 0.0183$) of the variation in the total cases in the first 100 days of each outbreak, respectively (Fig. 3). The total number of intersections against the total cases in the first 100 days is shown in Fig. S3. The total river network length in this study area explained less than 3% of the variation in the number of cases during the first 100 days of each outbreak. See Table S2 for complete R-squared values.

Multiple regression analysis

We found that the total road network lengths and the total river network lengths added separately to a multiple regression model explained 54% ($p = 0.0030$) of the variation in the number of cases during the first 100 days of each outbreak. These results are similar to those that we observed when we analyzed the relationship between total combined road and river network lengths and total cases in the first 100 days using a linear regression model, as described above. The R-squared values also indicated that the total road network length alone explained 30% ($p = 0.0183$) of the variation in the total cases in the first 100 days of each outbreak.

Complete R-squared values for the transportation networks against outbreak measures including total cases, total duration in days, and total cases in the first 100 days at each study area is found in Table S2. Additionally, Fig. S4a-d shows the total combined road and river network lengths against the total cases in the first 100 days for the following study areas: 50 km x 50 km, 100 km x 100 km, 200 km x 200 km, and 300 km x 300 km.

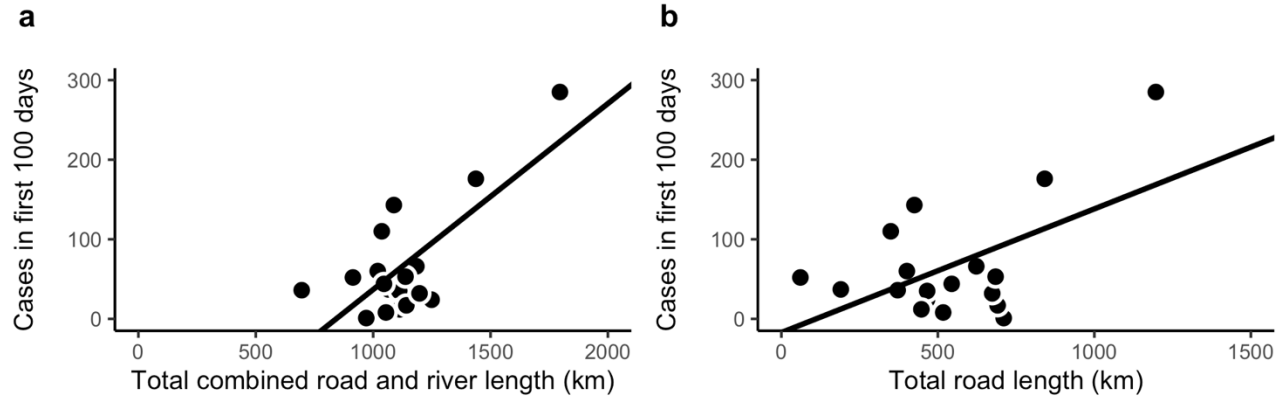


Figure 3. Linear Regression model on 150 km x150 km study area. **(a)** Total cases in the first 100 days of each outbreak against the total combined road and river length in km (R -squared = 0.5377 , p = 0.0005) **(b)** Total cases in the first 100 days of each outbreak against the total road length in km (R -squared = 0.3012 , p = 0.0183).

Subsequent outbreaks

Each of the 4 outbreaks that were classified as ‘subsequent outbreaks’ reported fewer total cases and a shorter duration in days than the preceding outbreak that occurred in the same or similar location. Each subsequent outbreak occurred within 2 years and < 60 km of the first outbreak. The shortest elapsed time between an initial and subsequent spillover event was 268 days while the longest elapsed time was 471 days. The subsequent outbreaks reported between 15 and 232 fewer cases than their preceding outbreaks. The subsequent outbreaks were between 15 and 108 days shorter in total duration than their associated initial outbreaks. Following the initial Mbomo outbreak in February 2003, the subsequent outbreaks in Mbomo in November 2003 and Etoumbi in May 2005, respectively, reported approximately 75% and 90% fewer cases compared to the total cases reported in the first outbreak. Likewise, the total duration of these subsequent outbreaks decreased by over 40% and 50%, respectively, compared to the duration of the initial spillover event (Table 1, Table S4).

Discussion

This study provides a novel data resource and novel analyses of transportation infrastructure surrounding all documented EBOV spillover events through 2020. All 18 of these events occurred within Central and West Africa. We examined the relationship between transportation infrastructure, which we used as a proxy for population mobility and connectivity, and outbreak measures. For outbreak measures, we considered the total number of cases in each outbreak, the total duration of each outbreak measured in days, and the number of cases reported in the first 100 days of each outbreak. We observed that the transportation networks had a stronger relationship with the number of cases in the first 100 days of an outbreak compared to the outbreak totals. Transportation network characteristics did not consistently show a significant relationship with the total cases or the total duration in days for each outbreak, likely because additional factors, including outbreak management, become determining factors in outbreak size after 100 days. These results suggest that transportation networks may play an important role in the transmission dynamics of EVD in the early stages of an outbreak.

Transportation networks connect people across locations. These networks both influence and are influenced by human mobility. Some EBOV-infected patients travel to healthcare centers and unknowingly transmit the virus along the way while others transmit the virus to healthcare personnel or visitors at the hospital^[45] while others believe they are uninfected and travel to escape an outbreak. The geographic spread of EVD following a spillover event is determined by the hosts’ ability to travel. This is consistent with the outbreaks that reported the greatest numbers of cases (e.g., West Africa) and large geographic spread^[46].

The influence of road networks of river networks on movement depends on various factors. Previous studies found that small villages are especially subject to poor road upkeep which limits the utilization of vehicles and hinders efficient travel [47]. Individuals in these areas will instead use animals like donkeys or walking as means of travel along roads [11]. River networks are also commonly used for transportation. Rivers connect major cities, for example Kinshasa and Kisangani in DRC, and Brazzaville and Kinshasa between DRC and RoC and also provide critical links between many smaller settlements [10,12,48].

To examine the immediate importance of movement and connectivity as it relates to outbreak severity, we conducted a targeted analysis on the first 100 days of each outbreak. This approach provided insight into the short-term effects of transmission and early outbreak management rather than the downstream effects of prolonged exponential outbreak growth and long-term outbreak management. This approach highlighted the potential importance of early interventions in outbreak containment. In addition, the largest outbreaks (West Africa 2013 and North Kivu 2018) were no longer outliers when using this approach (Fig. 1b). This analysis allowed for a comprehensive examination of the early mechanisms that underlie each outbreak despite the sizable differences in the total number of cases and duration in days across outbreaks. Due to the very high proportion of susceptible individuals to EBOV infection in most populations, movement patterns early in an outbreak are extremely important in determining outbreak transmission dynamics. We find strong positive correlations between the transportation networks and EVD incidence in the first 100 days of each outbreak at various study areas. These findings provide insight into the relationship between human mobility and the spread of Ebola virus and further suggest that rapid response measures along transportation networks surrounding spillover events during the initial stages of an outbreak have the potential to reduce outbreak size.

Previous research identified an association between delayed recognition of EVD and longer, larger outbreaks. We observed similar results in the context of a small number of subsequent outbreaks which occurred in the ROC, Gabon, and the DRC (Fig. 1a). We found that even when spillover events occur in close spatial proximity and close together in time, and therefore rely on very similar or identical road and river networks, the subsequent outbreak resulted in less severe outbreaks in terms of total cases, total duration (in days), and spatial spread. This was most strongly demonstrated by the two subsequent spillover events in Mbomo in November 2003 and Etoumbi in May 2005 that followed the initial Mbomo outbreak in February 2003. With each of these subsequent spillover events, we observe fewer total cases and shorter duration, indicating a favorable trend towards efficient containment and control of the outbreaks. Excluding subsequent events from the regression analysis strengthened the positive association between transportation networks and both total cases and total outbreak duration for many scenarios (Table S2 and S4). The responses to subsequent spillover events, especially those that occur shortly after the first spillover, may have benefited from experiential knowledge and outbreak response infrastructure developments, such as improved surveillance and diagnostics. This suggests that preparedness and rapid response, including early outbreak recognition, maybe be able to overcome the impact of movement in outbreak trajectory. We performed an additional analysis between transportation networks and total cases and total outbreak duration in which we excluded both subsequent events and outlier events. We found that the positive association between transportation networks and total cases and total duration was often further increased when excluding both subsequent and outlier events (Table S2 and S5).

To control future outbreaks, management plans can benefit from the inclusion of human movement and population connectivity, particularly in the early stages of outbreaks. Land use change (e.g., deforestation due to logging, mining, and development) has drastically increased pressures on ecosystems in parts of Central and West Africa. One study found that the index cases in humans of EVD outbreaks (2004-2014) occurred mainly in forest fragmented hotspots [49]. As population mobility and human settlements expand due to an increased demand for services and resources [50], forest fragmentation and wildlife disruption are an imminent concern with potentially deadly consequences.

In addition to enabling the geographic spread of pathogen transmission following a spillover event, movement may also assist in the occurrence of spillover events. Areas along rivers can have an abundance of fruit trees, which are common feeding sites for migrating fruit bats. For example, during the 2007 EBOV spillover event in Luebo, DRC, fruit bats suspected of being EBOV reservoirs (*H. monstrosus* and *E. franqueti*) migrated along the Lulua River near Luebo, stopping

at settlements along the river, where residents handled them^[51]. The influence of transportation networks and population mobility on the emergence and transmission of EBOV and other transmissible pathogens is wide-reaching and would benefit from further exploration.

Control efforts, particularly in the early days of an outbreak are critical in infectious disease outbreak management, including EBOV. These efforts include strengthening infection prevention & control (IPC) programs (this may include testing and contact tracing along roads and rivers), strengthening disease surveillance systems, improving health infrastructure, and developing accessible public health communications^[52]. Additional local factors that must be considered in control efforts include political and economic instability and conflicts, displaced populations, and large-scale population movements^[50]. Experiential knowledge gained from managing spillover events of other zoonotic pathogens can also serve to help guide control efforts within the early stages of an outbreak in the face of future threats.

Acknowledgments

We thank Kelsee Baranowski, Christina Faust, and Ephraim Hanks for their valuable feedback and insight in the development of this manuscript.

Contributions

NB and VM were responsible for the concept of the study. NB, AG, and BN completed the analyses. VM, MJM, and SNS provided relevant insight, data, and feedback. HDR provided support for map discovery, archiving, and digitization. NB and AG prepared the first draft of the manuscript. All authors reviewed and approved the manuscript before submission.

Funding

This study was supported by the joint National Institutes of Health (NIH) - National Science Foundation (NSF) - National Institute of Food and Agriculture (NIFA) Ecology and Evolution of Infectious Disease (award R01TW012434 to NB), NSF RAPID (award 2202872 to NB), and the Intramural Research Program of the National Institute of Allergy and Infectious Diseases (NIAID), NIH (1ZIAAI001179-01 to VM). SNS was partly supported by funding to Verena (viralemergence.org) from the NSF (award BII 2021909 and BII 2213854). Funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability. Hard copies of all Michelin maps used in this study are available at the Donald W. Hamer Center for Maps and Geospatial Information in the Penn State University Pattee Library. Digital copies and code used for analyses are available at <https://github.com/bhartilab/EbolaMaps>.

References

1. Viboud, C. *et al.* Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451 (2006).
2. Bharti, N. *et al.* Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science* **334**, 1424–1427 (2011).
3. Faria, N. *et al.* HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* **346**, 56–61 (2014).
4. Murray, M. J. Ebola Virus Disease: A Review of Its Past and Present. *Anesth Analg* **121**, 798–809 (2015).
5. Rosello, A. *et al.* Ebola virus disease in the Democratic Republic of the Congo, 1976-2014. *eLife* **4**, e09015 (2015).
6. Malvy, D., McElroy, A. K., de Clerck, H., Günther, S. & van Griensven, J. Ebola virus disease. *Lancet* **393**, 936–948 (2019).
7. Matson, M. J., Chertow, D. S. & Munster, V. J. Delayed recognition of Ebola virus disease is associated with longer and larger outbreaks. *Emerging Microbes & Infections* **9**, 291–301 (2020).
8. Lacroix, A. *et al.* Investigating the Circulation of Ebola Viruses in Bats during the Ebola Virus Disease Outbreaks in the Equateur and North Kivu Provinces of the Democratic Republic of Congo from 2018. *Pathogens* **10**, 557 (2021).
9. Wolfe, N. D., Daszak, P., Kilpatrick, A. M. & Burke, D. S. Bushmeat Hunting, Deforestation, and Prediction of Zoonotic Disease. *Emerg Infect Dis* **11**, 1822–1827 (2005).
10. Georges, A. J. *et al.* Ebola hemorrhagic fever outbreaks in Gabon, 1994-1997: epidemiologic and health control issues. *J Infect Dis* **179 Suppl 1**, S65-75 (1999).

11. Blanford, J. I., Kumar, S., Luo, W. & MacEachren, A. M. It's a long, long walk: accessibility to hospitals, maternity and integrated health centers in Niger. *Int J Health Geogr* **11**, 24 (2012).
12. Foster, V. & Sr, D. The Democratic Republic of Congo's Infrastructure: A Continental Perspective. (2011).
13. Michelin. Michelin Africa Central South and Madagascar Map 746. Penn State Library. Michelin Travel Publications.
14. Michelin. Michelin Africa North & West Map 741. Penn State Library. Michelin Travel Publications.
15. Hijmans, R. DIVA-GIS | free, simple & effective. *DIVA-GIS* <https://www.diva-gis.org/>.
16. ArcGIS Desktop | Desktop GIS Software Suite. <https://www.esri.com/en-us/arcgis/products/arcgis-desktop/overview>.
17. GRASS GIS - Bringing advanced geospatial technologies to the world. <https://grass.osgeo.org/> (2023).
18. Report of a WHO/International Study Team. Ebola haemorrhagic fever in Sudan, 1976. *Bulletin of the World Health Organization* **56**, 247 (1978).
19. Camacho, A. *et al.* Potential for large outbreaks of Ebola virus disease. *Epidemics* **9**, 70–78 (2014).
20. Georges, A.-J. *et al.* Ebola hemorrhagic fever outbreaks in Gabon, 1994–1997: epidemiologic and health control issues. *The Journal of infectious diseases* **179**, S65–S75 (1999).
21. Khan, A. S. *et al.* The reemergence of Ebola hemorrhagic fever, Democratic Republic of the Congo, 1995. *The Journal of infectious diseases* **179**, S76–S86 (1999).
22. 1996 - Ebola haemorrhagic fever in Gabon (new outbreak) - Update 2. https://www.who.int/emergencies/disease-outbreak-news/item/1996_10_18-en (1996).
23. 2001 - Ebola haemorrhagic fever in Gabon - Update. https://www.who.int/emergencies/disease-outbreak-news/item/2001_12_11-en (2001).
24. Gabon: Ebola outbreak appeal No. 04/2002 operations update No. 01 - Congo | ReliefWeb. <https://reliefweb.int/report/congo/gabon-ebola-outbreak-appeal-no-042002-operations-update-no-01> (2002).
25. Formenty, P. *et al.* Outbreak of Ebola hemorrhagic fever in the Republic of the Congo, 2003: a new strategy? *Med Trop (Mars)* **63**, 291–295 (2003).
26. Ebola haemorrhagic fever in the Republic of the Congo - Update 5. https://www.who.int/emergencies/disease-outbreak-news/item/2003_02_26-en (2003).
27. Nkoghe, D., Kone, M. L., Yada, A. & Leroy, E. A limited outbreak of Ebola haemorrhagic fever in Etoumbi, Republic of Congo, 2005. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **105**, 466–472 (2011).
28. Ebola haemorrhagic fever in the Democratic Republic of the Congo - update 3. https://www.who.int/emergencies/disease-outbreak-news/item/2007_09_27-en (2007).
29. Grard, G. *et al.* Emergence of Divergent Zaire Ebola Virus Strains in Democratic Republic of the Congo in 2007 and 2008. *J Infect Dis* **204**, S776–S784 (2011).
30. Sub-national time series data on Ebola cases and deaths in Guinea, Liberia, Sierra Leone, Nigeria, Senegal and Mali since March 2014 - Humanitarian Data Exchange. (2015).
31. WHO Ebola Response Team. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. *New England Journal of Medicine* **371**, 1481–1495 (2014).
32. Maganga, G. D. *et al.* Ebola virus disease in the Democratic Republic of Congo. *New England Journal of Medicine* **371**, 2083–2091 (2014).
33. Grimes, K. E. *et al.* Contextual, Social and Epidemiological Characteristics of the Ebola Virus Disease Outbreak in Likati Health Zone, Democratic Republic of the Congo, 2017. *Frontiers in Public Health* **8**, 349 (2020).
34. Hemingway-Foday, J. J. *et al.* Lessons learned from reinforcing epidemiologic surveillance during the 2017 Ebola outbreak in the Likati district, Democratic Republic of the Congo. *Health security* **18**, S-81 (2020).
35. Democratic Republic of Congo: Ebola Virus Disease - External Situation Report 15 - Democratic Republic of the Congo. <https://reliefweb.int/report/democratic-republic-congo/democratic-republic-congo-ebola-virus-disease-external-situation-13> (2018).
36. Ebola Virus Disease Democratic Republic of Congo. 6. <https://apps.who.int/iris/bitstream/handle/10665/272509/SITREP-EVD-DRC-20180511.pdf?ua=1> (2018).
37. Vossler, H. *et al.* Analysis of individual-level data from 2018–2020 Ebola outbreak in Democratic Republic of the Congo. *Sci Rep* **12**, 5534 (2022).
38. Ebola Cases and Deaths in the North Kivu Ebola Outbreak in the Democratic Republic of the Congo (DRC) - Humanitarian Data Exchange.
39. Ebola virus disease – Democratic Republic of the Congo. <https://www.who.int/emergencies/disease-outbreak-news/item/ebola-virus-disease-democratic-republic-of-the-congo> (2020).
40. Hasan, S., Ahmad, S. A., Masood, R. & Saeed, S. Ebola virus: A global public health menace: A narrative review. *J Family Med Prim Care* **8**, 2189–2201 (2019).
41. Ebola Disease. WHO | Regional Office for Africa <https://www.afro.who.int/health-topics/ebola-disease> (2023).
42. Pu, Q., Yoo, E.-H., Rothstein, D. H., Cairo, S. & Malemo, L. Improving the spatial accessibility of healthcare in North Kivu, Democratic Republic of Congo. *Applied Geography* **121**, 102262 (2020).

43. Mylne, A. *et al.* A comprehensive database of the geographic spread of past human Ebola outbreaks. *Sci Data* **1**, 140042 (2014).
44. R: The R Project for Statistical Computing. <https://www.r-project.org/>.
45. Breman, J. G. *et al.* The epidemiology of Ebola haemorrhagic fever in Zaire 1976. in *Proceedings of an International Colloquium on Ebola Virus Infection and Other Haemorrhagic Fevers held in Antwerp, Belgium* 103–124 (Elsevier / North-Holland Biomedical Press, 1978).
46. Coltart, C. E. M., Lindsey, B., Ghinai, I., Johnson, A. M. & Heymann, D. L. The Ebola outbreak, 2013–2016: old lessons for new epidemics. *Philos Trans R Soc Lond B Biol Sci* **372**, 20160297 (2017).
47. Maganga, G. D. *et al.* Ebola virus disease in the Democratic Republic of Congo. *N Engl J Med* **371**, 2083–2091 (2014).
48. Grimes, K. E. L. *et al.* Contextual, Social and Epidemiological Characteristics of the Ebola Virus Disease Outbreak in Likati Health Zone, Democratic Republic of the Congo, 2017. *Front Public Health* **8**, 349–349 (2020).
49. Rulli, M. C., Santini, M., Hayman, D. T. S. & D’Odorico, P. The nexus between forest fragmentation in Africa and Ebola virus disease outbreaks. *Sci Rep* **7**, 41613 (2017).
50. Democratic Republic of Congo Urbanization Review. 89.
51. Leroy, E. M. *et al.* Human Ebola Outbreak Resulting from Direct Exposure to Fruit Bats in Luebo, Democratic Republic of Congo, 2007. *Vector-Borne and Zoonotic Diseases* **9**, 723–728 (2009).
52. Kuisma, E. *et al.* Long-term wildlife mortality surveillance in northern Congo: a model for the detection of Ebola virus disease epizootics. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**, 20180339 (2019).

Supplementary Information

Human movement and transmission dynamics early in Ebola outbreaks

Alexandria Gonzalez¹, Behnam Nikparvar¹, M. Jeremiah Matson², Stephanie N. Seifert³, Heather D. Ross⁴, Vincent Munster⁵, Nita Bharti^{1*}

1 Biology Department and Center for Infectious Disease Dynamics, Penn State University, University Park, PA, 16802, USA

2 University of Utah Health, Department of Internal Medicine, Salt Lake City, UT 84132, USA

3 Paul G. Allen School for Global Health, Washington State University, Pullman, WA, 99164 USA

4 Donald W. Hamer Center for Maps and Geospatial Information, Penn State University Libraries, Penn State University, University Park, PA, 16802, USA

5 Division of Intramural Research, Laboratory of Virology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, Montana, 59840 USA

*Corresponding author: nita@psu.edu

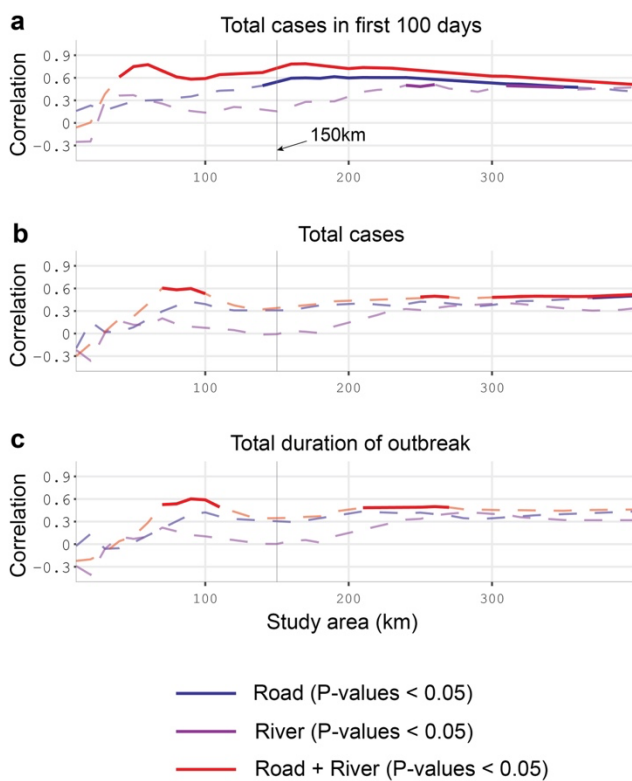


Figure S1. Sensitivity analysis of study area for correlations between transportation networks and outbreak measures: (A) cases in first 100 days, (B) total cases, and (C) total duration of outbreak. Dashed lines show p-values > 0.05; solid lines show p-values < 0.05; colors indicate transportation network.

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Spillover location	Net change in road length (km)
Yambuku	+ 362.99
Tandala	+ 105.67
Mekouka	+ 9.48
Kikwit	+ 184.77
Mayibout	+ 0.77
Booue	+ 52.49
Mekambo/Mbomo	+ 57.48
Mbomo1	+ 13.78
Mbomo2	+ 0.93
Etoumbi	+ 14.42
Luebo07	+ 24.31
Luebo08	+ 15.59
Inkanamango	- 5.25
Likati	- 37.86
Bikoro	+ 162.38
North Kivu	- 119.55
Mbandaka	+ 13.83

Table S1. Net change in road length from 1963 to 2019 in the 150 km x 150 km study area surrounding the 17 spillover events that occurred in Central Africa. (+) Gain of road length, (-) loss of road length.

Study area	Outbreak measure											
	Cases in first 100 days				Total cases				Total duration (in days)			
	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section
50 x 50	0.1365	0.0827	0.5580 ^{***S4A}	0.5589 ^{***}	0.0137	0.0070	0.0519	0.0071	0.0050	0.0005	0.0114	0.0040
100 x 100	0.0201	0.1565	0.3540 ^{**S4B}	0.3940 ^{**}	0.0065	0.1529	0.2812 [*]	0.1367	0.0110	0.1780	0.3470 ^{*S2}	0.1143
150 x 150	0.0259	0.3012[*]	0.5377^{***}	0.4330 ^{***S3}	0.0000	0.0939	0.1146	0.0219	0.0000	0.0900	0.1150	0.0178
200 x 200	0.1276	0.3558 ^{**}	0.5250 ^{***S4C}	0.4910 ^{**}	0.0200	0.1531	0.1906	0.0735	0.0227	0.1717	0.2143	0.1091
300 x 300	0.2168	0.2694 [*]	0.3890 ^{***S4D}	0.4306 ^{**}	0.1626	0.1441	0.2311 [*]	0.1494	0.1700	0.1184	0.2061	0.1970

Table S2. R-squared values for the transportation networks including total river length, total road length, total combined road and river length, and total intersections against outbreak measures: cases in first 100 days, total cases, and total duration in days for each study area examined (*p < 0.05, **p<0.01, ***p<0.001). All spillover events included. Upper case superscripts refer to figures that illustrate each relationship. Bold values correspond to Fig. 3A and Fig. 3B of the main paper, which illustrate these relationships.

Study area	Outbreak measure											
	Cases in first 100 days				Total cases				Total duration (in days)			
	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section
50 x 50	0.1170	0.0738	0.5580 ^{***}	0.6020^{***}	0.0370	0.0119	0.1352	0.2620[*]	0.1190	0.0003	0.1630	0.1804
100 x 100	0.0123	0.100	0.2710 [*]	0.3160 [*]	0.0182	0.0223	0.0010	0.0333	0.1130	0.0114	0.2670 [*]	0.1800
150 x 150	0.0267	0.2450	0.4790 ^{**}	0.4280 ^{**}	0.0690	0.1236	0.0411	0.0855	0.0915	0.0538	0.2460	0.1804
200 x 200	0.1030	0.2890 [*]	0.4560 ^{**}	0.4590 ^{**}	0.0002	0.2890[*]	0.2880[*]	0.2330	0.2110	0.1060	0.2790[*]	0.2060
300 x 300	0.130	0.1970	0.2950 [*]	0.3740 [*]	0.0364	0.2130	0.2420	0.2740[*]	0.5710^{***}	0.0245	0.1960	0.1850

Table S3. R-squared values for the transportation networks including total river length, total road length, total combined road and river length, and total intersections against outbreak measures: cases in first 100 days, total cases, and total duration in days for each study area examined (*p < 0.05, **p<0.01, ***p<0.001). Outlier events

It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

excluded. The outlier events are as follows: Meliandou Guinea in 2013 and North Kivu DRC in 2019. Upper case superscripts refer to figures that illustrate each relationship. Bold values are greater than when all spillover events are included.

Study area	Outbreak measure											
	Cases in first 100 days				Total cases				Total duration (in days)			
	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section
50 x 50	0.1780	0.1340	0.6140^{***}	0.6110^{***}	0.0180	0.0090	0.0520	0.0073	0.0070	0.0004	0.0110	0.0043
100 x 100	0.0613	0.1560	0.3800[*]	0.3720 [*]	0.0202	0.1620	0.2910[*]	0.1250	0.0452	0.1760	0.3770[*]	0.0916
150 x 150	0.0831	0.3160[*]	0.5910^{**}	0.4300 [*]	0.0020	0.0920	0.1180	0.0162	0.0111	0.0760	0.1260	0.0085
200 x 200	0.2030	0.3540 [*]	0.5200 ^{**}	0.4650 ^{**}	0.0340	0.1490	0.1820	0.0600	0.0482	0.1470	0.1920	0.0760
300 x 300	0.2430	0.3260[*]	0.4280[*]	0.4490^{**}	0.1770	0.1620	0.2390	0.1470	0.1930	0.1400	0.2230	0.1930

Table S4. R-squared values for the transportation networks including total river length, total road length, total combined road and river length, and total intersections against outbreak measures: cases in first 100 days, total cases, and total duration in days for each study area examined (*p < 0.05, **p<0.01, ***p<0.001). Subsequent events excluded. The subsequent events are as follows: Mayibout Gabon in February 1996, Mbomo RC in November 2003, Etoumbi RC in May 2005, and Luebo DRC in September 2007. Upper case superscripts refer to figures that illustrate each relationship. Bold values are greater than when all spillover events are included.

Study area	Outbreak measure											
	Cases in first 100 days				Total cases				Total duration (in days)			
	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section	River	Road	Road & river	Inter-section
50 x 50	0.1490	0.1210	0.6070^{**}	0.6460^{**}	0.0514	0.0203	0.1550	0.3090	0.1800	0.0006	0.1920	0.2210
100 x 100	0.0418	0.1070	0.3210	0.2970	0.0043	0.0124	0.0043	0.0145	0.3470[*]	0.0033	0.3860[*]	0.1590
150 x 150	0.0782	0.2620	0.5390[*]	0.4200 [*]	0.0334	0.1150	0.0568	0.0607	0.2980	0.0399	0.3360[*]	0.1640
200 x 200	0.1650	0.2940	0.4570 [*]	0.4370 [*]	0.0050	0.2800	0.2800	0.1870	0.4380[*]	0.0776	0.2820	0.1620
300 x 300	0.1510	0.2580	0.3450 [*]	0.4050 [*]	0.0484	0.3150	0.3220	0.3100	0.7930^{***}	0.0460	0.2740	0.2150

Table S5. R-squared values for the transportation networks including total river length, total road length, total combined road and river length, and total intersections against outbreak measures: cases in first 100 days, total cases, and total duration in days for each study area examined (*p < 0.05, **p<0.01, ***p<0.001). Subsequent and outlier events excluded. The subsequent events are as follows: Mayibout Gabon in February 1996, Mbomo RC in November 2003, Etoumbi RC in May 2005, and Luebo DRC in September 2007. The outlier events are as follows: Meliandou Guinea in 2013 and North Kivu DRC in 2019. Upper case superscripts refer to figures that illustrate each relationship. Bold values are greater than when all spillover events are included.

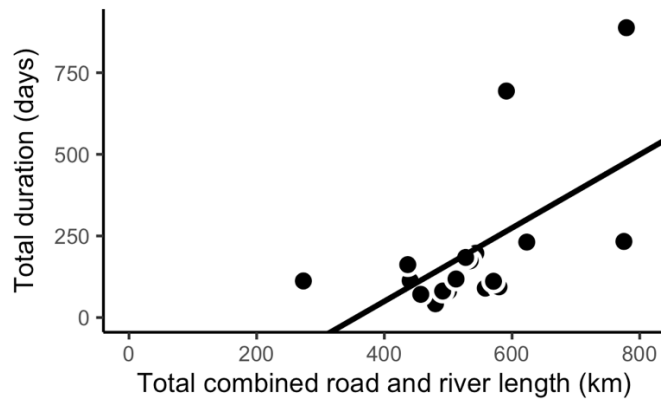


Figure S2. Linear Regression model of the total duration in days of each outbreak and the total combined road and river length in km in each 100x100 study area surrounding the spillovers (R-squared = 0.3470 , p = 0.0101).

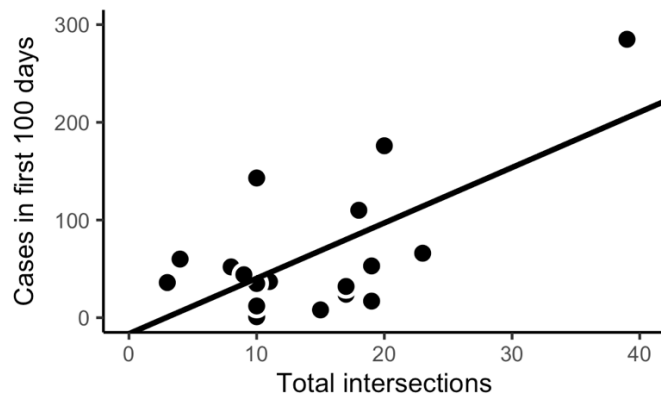


Figure S3. Linear Regression model of the total cases in the first 100 days of each outbreak and the total number of intersections in a 150 km x 150 km study area surrounding the spillovers (R-squared = 0.4338, p = 0.0030).

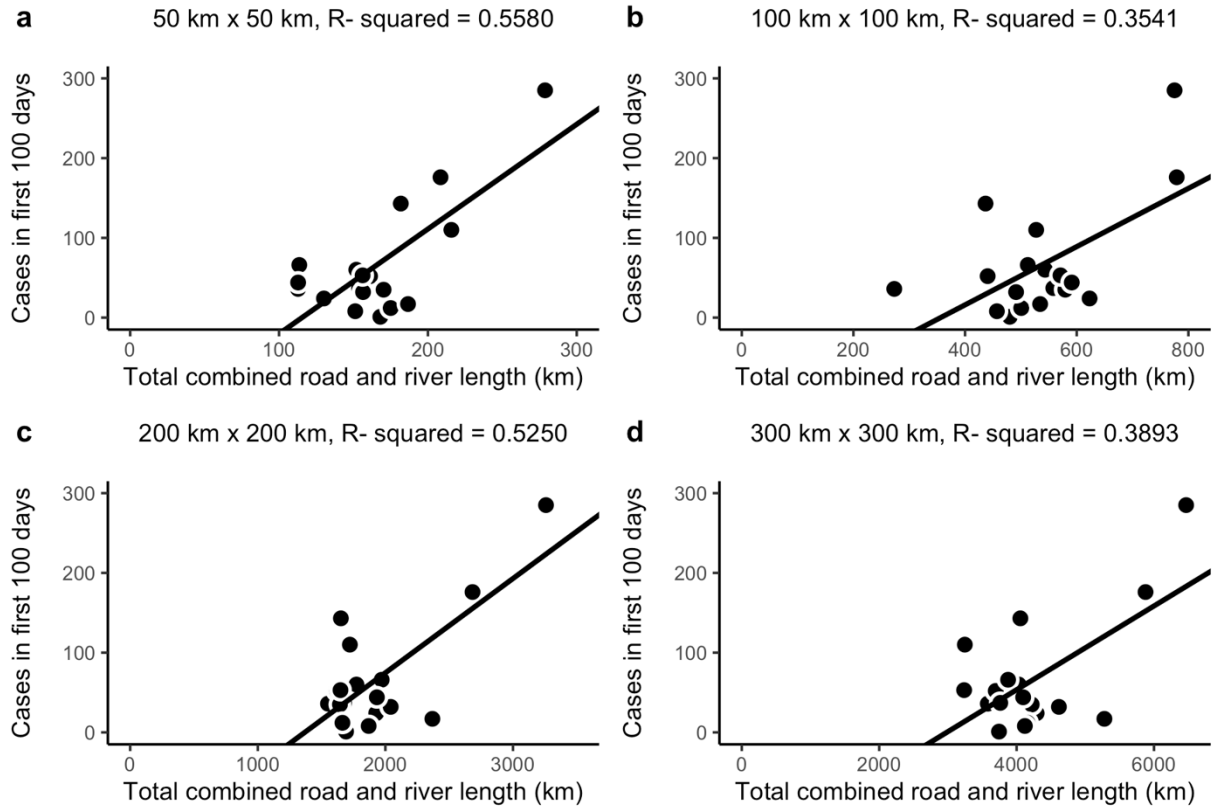


Figure S4. Linear Regression models of the total cases in the first 100 days and the total combined road and river length in km in a **(a)** 50 km x 50 km study area (R-squared = 0.5580, $p = 0.0004$), **(b)** 100 km x 100 km study area (R-squared = 0.3541, $p = 0.0092$), **(c)** 200 km x 200 km study area (R-squared = 0.5250, $p = 0.0007$), **(d)** 300 km x 300 km study area (R-squared = 0.3893, $p = 0.0057$).