Predicting Ebola virus disease risk and the role of African bat birthing

C. Reed Hranac\textsuperscript{a,b,*}, Jonathan C. Marshall\textsuperscript{b}, Ara Monadjem\textsuperscript{c,d}, David T.S. Hayman\textsuperscript{b,*}

\textsuperscript{a} Molecular Epidemiology and Public Health Laboratory, Hopkirk Research Institute, Massey University, Private Bag, 11 222, Palmerston North 4442, New Zealand  
\textsuperscript{b} Institute of Fundamental Sciences, Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand  
\textsuperscript{c} Department of Biological Sciences, University of Eswatini, Private Bag 4, Kwaluseni, Eswatini  
\textsuperscript{d} Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria, Republic of South Africa

\begin{abstract}
Ebola virus disease (EVD) presents a threat to public health throughout equatorial Africa. Despite numerous ‘spillover’ events into humans and apes, the maintenance reservoirs and mechanism of spillover are poorly understood. Evidence suggests fruit bats play a role in both instances, yet data remain sparse and bats exhibit a wide range of life history traits. Here we pool sparse data and use a mechanistic approach to examine how birthing cycles of African fruit bats, molossid bats, and non-molossid microbats inform the spatio-temporal occurrence of EVD spillover. We create ensemble niche models to predict spatio-temporally varying bat birthing and model outbreaks as spatio-temporal Poisson point processes. We predict three distinct annual birthing patterns among African bats along a latitudinal gradient. Of the EVD spillover models tested, the best by quasi-Akaike information criterion (qAIC) and by out of sample prediction included significant bat birth-related terms. Temporal bat birthing terms fit in the best models for both human and animal outbreaks were consistent with hypothesized viral dynamics in bat populations, but purely spatial models also performed well. Our best model predicted risk of EVD spillover at locations of the two 2018 EVD outbreaks in the Democratic Republic of the Congo was within the top 12–35% and 0.1% of all 25 × 25 km spatial cells analyzed in sub-Saharan Africa. Results suggest that sparse data can be leveraged to help understand complex systems.
\end{abstract}

\section{Introduction}

Since the first reports of Ebola virus disease (EVD) in the late 1970s outbreaks have occurred sporadically (Pourrut et al., 2005), with 34 identified human index cases to date (Fig. 1). The 2013–2016 West African EVD epidemic caused the death of at least 11,310 people and was a Public Health Emergency of International Concern (Briand et al., 2014). Outbreaks have been unpredictable, sometimes including more than one primary transmission or ‘spillover’ event, and together with high case fatality rates cause significant strain on public health authorities in typically resource poor settings (Piggott et al., 2016). Additional conservation concerns have arisen such as the large scale great ape die-offs linked to EVD in central Africa (Leroy et al., 2004). Despite decades of research the ecology, sylvatic maintenance, and mechanisms of spillover of these viruses to humans, or non-humans, have remained understudied (Olival and Hayman, 2014).

The genus \textit{Ebolavirus} (family Filoviridae) is composed of five recognized viral species, four of which are agents of EVD (Olival and Hayman, 2014). An additional proposed sixth \textit{Ebolavirus} species with unknown ability to cause EVD was also described recently (Goldstein et al., 2018). For the four \textit{Ebolavirus} species known to cause EVD mounting evidence implicates African Old World fruit bat species (Order Chiroptera; family Pteropodidae) as potential reservoirs as defined by Haydon et al. (2002) in maintaining and propagating infection on the landscape (De Nys et al., 2018; Leroy et al., 2005; Han et al., 2016; Pourrut et al., 2009). However, different bats may play roles as both necessary and sufficient hosts in reservoir host communities (Nishiura et al., 2009). To date no full EVD-causing \textit{Ebolavirus} has been isolated from a bat host. Possible reasons for this, if they are true maintenance reservoir hosts, include: the remote geographic locations of the EVD outbreaks; nomadic or migratory nature of bats (Nowak, 1994); potentially acute viremic periods (Swanepoel et al., 1996); spatio-temporal extinction and recolonization dynamics, the unknown role of viral persistence and intermittent viral shedding (Schuh et al., 2017), and possibly biased samples unrepresentative of true bat diversity (Leendertz et al., 2016).

Despite these limitations, RNA viral fragments of ebolaviruses causing EVD have been recovered from three African fruit bat species: Franquet’s epauletted fruit bat (\textit{Epomops franqueti}), the Hammer-headed bat (\textit{Hypsipetes monstrosus}), and the Little collared fruit bat...
(Myonycteris torquata) (Leroy et al., 2005). Ebolavirus fragments were also detected in four rodent species: two Praomys sp., two Mus sp. and one shrew Sylvisorex ollula (family: Soricidae) (Morvan et al., 1999), although there is some contention whether these species could effectively act as reservoir hosts (Swanepoel et al., 1996) and whether these were in fact active infections (Taylor et al., 2010). Phylogenetic analyses have also suggested that fruit bats have played some role in the recent evolutionary history of Zaire ebolavirus (Walsh et al., 2005; Biek et al., 2006). Most recently 20% of an Ebolavirus has reportedly been sequenced from an insectivorous greater long-fingered bat (Miniopterus inflatus), captured in 2016 in Liberia (Kupferschmidt, 2019).

Much of what is postulated about the ecological niche and sylvatic maintenance of ebolaviruses is inferred from a member of its sister genus, Marburgvirus marburgvirus (MARV). MARV circulates within cave dwelling Egyptian fruit bat (Rousettus aegyptiacus) colonies with seasonal increases in viral shedding among juvenile bats and the high birthing synchrony (Amman et al., 2012), although the inclusion of persistent infection with intermittent viral shedding has still to be explored for ebolaviruses. Anti-marburgvirus antibodies have also been detected within co-roosting insectivorous bats, which to date play an unknown role in sylvatic maintenance. Machine learning and infection dynamic simulations have both identified biannual, synchronous birth-pulses as potentially important host life history traits for filovirus maintenance, all traits found within African bat species (Han et al., 2016; Hayman, 2015).

While there are many similarities between MARV and ebolavirus systems, the principal suspected ebolavirus reservoirs are primarily tree roosting bat species with far lower colony densities compared to the Egyptian fruit bat. The sporadic occurrence of outbreaks, lifestyles of putative reservoirs, viral dynamics and subsequent antibody production in bats (Swanepoel et al., 1996) suggest a potential reliance on metapopulations within a larger reservoir community to sustain sylvatic ebolavirus transmission (Brown et al., 2013; Grenfell et al., 2001; Han et al., 2016; Hassanin et al., 2016). Strong, spatially asynchronous seasonal forcing events, including birthing, likely add additional layers of complexity in disease outbreak events (Duke-Sylvester et al., 2011).

Spatial risk models have previously identified regions of realized and unrealized risk throughout equatorial sub-Saharan Africa, linked to a variety of drivers, including forest fragmentation and vegetation index (Rulli et al., 2017; Pigott et al., 2014, 2016). Additional temporal risk models have suggested that EVD spillover risk across the region is greatest during transitions between wet and dry seasons (Schmidt et al., 2017), periods also associated with African bat species birthing events (Cumming and Bernard, 1997). One hypothesis of filovirus spillover risk posits that synchronous shifts in population demographics may be responsible for periods of increased viral shedding similar to those witnessed in the MARV system (Hayman, 2015; Pourrut et al., 2009).

In this work we test the hypothesis that birthing among some African bat species is linked to EVD spillover events to humans and forest dwelling animals susceptible to EVD, hereafter referred to as "non-human spillover hosts." Due to the rarity of EVD outbreak events we were not able to model each EVD-causing Ebolavirus species...
separately, and therefore combined them all together to represent a single, general, “ebolavirus” capable of infecting and causing EVD in mammalian hosts. We created two general model frameworks: a model describing the birthing patterns of African bats; and another to describe the relative spillover risk. We assumed bats were reservoir hosts and that non-human spillover hosts have different habitat use, risk factors and contact patterns with the potential reservoir host(s) than humans. Subsequently we modeled spillover into non-human spillover hosts independently of humans. We then used the two most recent outbreaks of EVD in the Democratic Republic of the Congo (DRC) (Barry et al., 2018; WHO, 2018), which were excluded from our modeling procedures (but see methods), to test the predictive power of our generated models.

2. Materials and methods

Evolutionary, environmental, and phenological processes undoubtedly drive the timing and location of bat breeding and birthing, yet these intricacies have yet to be fully described. Sparse data exist for the timing bat births, especially in mainland tropical Africa (Cumming and Bernard, 1997; Monadjem et al., 2010). To circumvent these issues we compiled the better part of 50 years of ecological research into a single, recirculating year of spatio-temporally defined birthing events (full data mining methods in SI Methods). Bat species were subset into three groups: African fruit bats (family Pteropodidae), mole-sossid bats (family Molossidae), and non-molossid microbats, as described by Cumming and Bernard (1997). These taxonomic clusters tend to share key life history traits such as diet and reproductive strategies making them useful subgroups for this analysis. Combining these occurrence data with static and temporally explicit geographic covariate layers (see SI Table 2) allowed us to construct temporally defined ecological niche models (ENMs) of bat birthing across a subset of sub-Saharan Africa. This subset was restricted to locations in mainland Africa receiving more than 500 mL of precipitation annually, as suggested by Schmidt et al. (2017). While outbreaks have not been observed in much of this region and hypothesized reservoir species use subsets of these regions making them a potentially unrealized viral niche. Unlike most ENMs that correlate a probably of occurrence with environmental conditions at any given location, we modified our procedure to correlate dynamic phenological signals focused on changing environments with bat birthing cycles (see full details for covariate selection and processing in SI methods and SI Table 2). We used a sliding two month observation window to estimate the probability of wild species giving birth across space and time for all pairs of consecutive months where sufficient data existed (see full ENM description in SI Methods).

The ENM results were used and combined with bat diversity in a number of different ways. First and most intuitively, the probability of birthing for the three different taxonomic class were used as covariates where probability of birth occurrence is \( P_{T,i,t} \) for each taxonomic group \( T \), at location \( i \) and time \( t \). Next the ENM results were combined with local species diversity estimates to create a “force of birthing” \( \lambda_{T,i,t} \) metric, representing the relative number of taxonomic member species giving birth at any time and space, as defined by:

\[
\lambda_{T,i,t} = P_{T,i,t} D_{T,i}
\]

where the number of taxonomic members is given by \( D_{T,i} \). The abbreviations \( afb \), \( mic \) and \( mol \) were used to represent African fruit bats, non-molossid microbats, and molossid bats respectively. Third, the total conditional probability \( P_{\text{total},i,t} \) of any bat birthing from any of the three taxonomic groups at location \( i \) and time \( t \), assuming independence, defined by:

\[
R_{\text{total},i,t} = 1 - \prod_{T=1}^{n} (1 - P_{T,i,t})
\]

A total “force of birthing” \( \lambda_{\text{total},i,t} \) metric representing the number of bats of all taxa giving birth at any time and space as defined by:

\[
\lambda_{\text{total},i,t} = P_{\text{total},i,t} \sum_{T=1}^{n} D_{T,i}
\]

Finally, a modified version of the probability of birth was made for the African fruit bats in which we accounted for variation within taxonomic classes in the first annual birthing pulse of the year. Previous work has indicated that the bi-annually polyestrous breeding African fruit bats will have a birth pulse proceeding the start of the short dry period in addition to giving birth following the long dry season like all monestrous species. Historically phenomena was only seen above \( \sim 18^\circ \text{S} \), south of which all fruit bat populations only had a single annual birth associated with the long dry-season (Cumming and Bernard, 1997). However, our modeling results suggested a threshold of \( \sim 10^\circ \text{S} \) and therefore we used this value. We accounted for this by down-weighting the probability of birth on the first six months of the calendar year by a ratio of biannual genera:total genera \( \left( \frac{1}{2} \right) \), as determined by our data-mining results. This information was not available on a comparable scale for any other taxonomic group. This modification allowed us to further probe the involvement of African fruit bat births on ebolavirus spillover events and is denoted throughout with the * symbol.

To test the hypothesized link between spillover events and bat birthing we collapsed the 42 year history of EVD outbreaks into the same recirculating year used in the birthing ENM. We then treated outbreaks as a spatio-temporal Poisson point process, with rate (herein, risk) estimated using a spatially weighted, over-dispersed general linear model (spatGLM). To assess the hypothesized temporal lags between birth pulses and outbreak events due to shifting population demographics we included lag terms for bat birthing terms. In the model with most parameters, the risk of EVD spillover into non-human spillover hosts was defined as:

\[
\log(R_{\text{risk},i,t}) \sim \beta_1 P_{\text{total},i,t} + \beta_2 P_{\text{mac},i,t} + \beta_3 P_{\text{mol},i,t} + \beta_4 D_{\text{fl} i} + \beta_5 D_{\text{fl} i} + \beta_6 D_{\text{fl} i} + \beta_7 D_{\text{fl} i} + \beta_8 D_{\text{fl} i} + \beta_9 \text{NBMdiv}_{i} + \beta_{10} \text{PopDen}_{i} + \beta_{11} \text{fragIndex}_{i} + \beta_{12} \text{BVD}_{i}
\]

where \( R_{\text{risk},i,t} \) represented the risk of EVD spillover at spatial location \( i \) and month \( t \). Potential sources of infection on the landscape were defined as times and locations where RNA fragments had been recovered from wild bats, our hypothesized reservoir hosts, and were included by the term BVD\(_i\). Seropositive animals were not included because the timing of infection is unknown. Non-bat mammalian diversity and a forest fragmentation index (Wilkinson et al., 2018) were also included as terms NBMdiv\(_i\) and fragIndex\(_i\). Human population density, Pop\(_{\text{Den}}\), was included to represent the human-wildlife interface (Plowright et al., 2015). The transformation \( \log(x + 1) \) was used to treat all variables (with the exception of BVD\(_i\)) on the log scale while allowing that they might be zero.

Human outbreak events are not exclusively linked to direct exposure to proposed bat reservoir host(s) (Pourrut et al., 2005) and so outbreaks in non-humans were included within the human spillover model to capture these transmission chains. A lag term of a single month was included for non-human outbreaks to describe the transmission chain and incubation periods of non-human spillover hosts acting as amplifying hosts and transmitting virus to humans (Swaneepoel et al., 1996). The risk of EVD spillover into humans (\( R_{\text{human}} \)) was therefore described as:

\[
\lambda_{\text{total},i,t} = P_{\text{total},i,t} \sum_{T=1}^{n} D_{T,i}
\]
log(\(R_{\text{hum},i,t}\)) \sim \beta_1 P_{\text{hum},i,t} + \beta_2 P_{\text{mic},i,t} + \beta_3 P_{\text{mol},i,t} + \\
\beta_4 P_{\text{hum},i,t-2} + \beta_5 P_{\text{mic},i,t-2} + \beta_6 P_{\text{mol},i,t-2} + \\
\beta_7 P_{\text{hum},i,t-4} + \beta_8 P_{\text{mic},i,t-4} + \beta_9 P_{\text{mol},i,t-4} + \\
\beta_{10} P_{\text{hum},i,t-6} + \beta_{11} P_{\text{mic},i,t-6} + \beta_{12} P_{\text{mol},i,t-6} + \\
\beta_{13} D_{\text{hum},i} + \beta_{14} D_{\text{mic},i} + \beta_{15} D_{\text{mol},i} + \\
\beta_{16} \text{NBMdiv} + \beta_{17} \text{PopDen} + \beta_{18} \text{fragIndex} + \\
\beta_{19} \text{BVD}_{i,t} + \beta_{20} \text{OB}_{\text{mic},i,t} + \beta_{21} \text{OB}_{\text{mol},i,t-1} \tag{4}

where OB\(_{\text{mic},i,t}\) is the occurrence of an EVD spillover into a non-human spillover host at spatial location \(i\), and month \(t\).

Within alternative models, bat diversity terms were tested as both combined (read: any bat species), and separated by taxon to test the relative contribution of each tax group to overall spillover risk. “Null” models in which all bat birth associated terms (\(P_{\text{hum},i,t}\), \(P_{\text{mic},i,t}\), \(P_{\text{mol},i,t}\)) and \(\lambda_{\text{total},i,t}\) were excluded were also computed and model comparison and selection were completed though qAIC (see SI Methods for details).

Additional validation was provided by testing our model predictions using the two most recent EVD outbreaks in the DRC, which were not included in the analysis. To do so, we estimated relative risk as \(R_{\text{null},i,t}\), where \(g_{\text{mean}}(R_{\text{null},i,t})\), which was calculated for all cells across all 12 months. Values were extracted for all cells covering the health districts where the spillover cases were purported to have occurred (Barry et al., 2018; WHO, 2018) and visually compared. Because the outbreak numbers are low, we also checked the model sensitivity to the inclusion of the two data points, and performed the same analyses as above.

Model forms for both human, and non-humans spillover host spatGLMs are available from SI Tables 5 and 6. All work was completed using custom code written for R (version 3.4.3 R Core Team, 2017) and is available at github.com/cReedHranac/FS.

3. Results

A total of 63 articles provided information on 14 African fruit bat species, 7 molossid species, and 43 non-molossid microbats at 75 unique locations, totaling 164 event records (Fig. 2, SI Table S1). Of these, 89% were retained with 8 records discarded as outside of our study extent within mainland Africa. EMNs were generated for all three bat taxonomic groups for 34 of 36 possible month pairs, and full model results are summarize in SI Table S2 and visualized in SI Figures 1–3.

At equatorial latitudes (± 10°) some African fruit bats and non-molossid microbats were predicted to experience conditions suitable for biannual birthing seasons, while molossid bat species had suitable conditions for birth nearly continuously throughout the year (Fig. 3). Above ~10°N the signal of habitat suitability was reduced for the non-molossid microbats and fruit bat species to two periods and one more broad period of habitat suitable for birth respectively. Results for molossid bats still suggested the suitability for two birth pulses although the signal for the period following the long dry season was less intense than at equatorial latitudes. At more southern latitudes (> ~10°S) molossid species results suggested suitability for biannual birthing patterns with a high degree of synchronicity in following to the long dry season. Both fruit bats and the non-molossid microbats were predicted to reduce births to a monoestrous reproductive cycle with births prior to the long dry season at southern latitudes. These results are largely consistent with previous work (Bernard and Cumming, 1997; Happold and Happold, 1990, 2013; Monadjem et al., 2010), but critically provide a continuous metric describing both the spatial and temporal aspects of the phenomenon.

The inclusion of bat terms improved the qAIC and predictive ability of our spatGLM models compared to the “null” models without bat terms for all human and the majority of animal models (Fig. 4, SI Tables 5 and 6). The top animal spillover model by qAIC included product terms of the conditional probability of bat birth (including the modified fruit bat term) and total bat diversity. The African fruit bat terms describing bat birthing events 5–6 months prior and both the molossid and fruit bat terms the month before and the month of spillover events were both significant and positively related to spillover risk to other mammals. Diversity of non-volant mammals is identified as a significant predictor and human population density had the only inverse significant relationship identified with spillover risk (see SI Table 5). The spatial and temporal locations of viral RNA detections in bats was also found to be a significant term providing credence to some of our

![Fig. 2. Spatial and temporal occurrence of bat birthing records obtained from the literature.](Image)
Fig. 3. Probability of birth of African bats. Top row: African fruit bats. Middle row: Molossid bats. Bottom row: Non-molossid microbats. Left: Mean annual probability of a bat giving birth spatially. Grey represents low values while more intense colors denote regions where environmental conditions correlated with bat birth events are met. White regions were excluded from this study. Right: Probability of birth by latitude and month. Height and color intensity are proportional to relative number of species giving birth within the latitudinal window. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
hypotheses. Risk of EVD spillover to non-human spillover hosts was largely focused on Central Africa with high annual probability of infection across Gabon and the Republic of the Congo (RoC), the Eastern region of the Central African Republic (CAR), Northeastern DRC and included pockets in Western Africa covering Liberia, northern Côte d’Ivoire, and Ghana (SI Figure S4). Additional regions of potential risk were predicted throughout South Sudan, Uganda, and parts of Kenya. At equatorial latitudes risk was constantly elevated over the rest of the study area with seasonal peaks centered around May and September. Another region of elevated annual risk was predicted at 25° S covering Angola, Zambia, eastern Mozambique and Northern South Africa despite no known occurrences of ebolaviruses south of 10° S. Seasonal fluctuations in risk peaked between May to July in a single annual pulse rather than the bi-modal pattern observed further North.

The top model for human spillover risk by qAIC used the taxonomically separated probability of birth terms and included the down-weighted fruit bat modification (*). Significant terms identified included EVD spillover into a non-human host in the month of and month prior to an outbreak, suggesting the model’s sensitivity to the stuttering infection chains within the system and supporting observed patterns of infection transmission. The down-weighted probability of birthing term for the African fruit bat group 5–6 months prior, and the month of and month proceeding spillover events for both the fruit bats and non-mollusid microbats were significant to predictions and top two models are summarized in Table 1 (See SI Table 6 for full alternative model results). Diversity of non-bat mammals was again identified as a significant predictor although neither viral detection in bats, nor the effect of forest fragmentation were identified as significant predictors of ebolavirus spillover to humans within this analysis. Risk of EVD spillover to humans was somewhat different to that of spillover into the non-human spillover hosts, as annual risk was more diffuse across Central Africa and more evenly distributed across areas of Western Africa (Fig. 5). Similar to the non-human spillover hosts, the highest risk zones occurred around the border of Gabon and the RoC, although Cameroon was included and the high risk of spillover within the CAR were missing from the human model results. The associated temporal analysis (Fig. 5 right) suggest seasonality within the system with two distinct periods of increased risk centered around the months of June and November just north of the equator. Acknowledging possible surveillance bias, the difference in peak relative risk between the human and non-human spillover host models suggest potential differences in spillover mechanisms between human and non-reservoir mammal host systems.

The two most recent EVD outbreaks in the DRC provided a test of the predictive power of our model (Fig. 6). The EVD outbreak announced by the WHO in July likely started in April within the Bikoro health district in the northeastern part of the country close to the border with RoC (Barry et al., 2018). This health district showed seasonality with peaks and troughs in the relative risk of viral spillover into humans
Fig. 5. Risk of Ebola virus disease outbreaks among people. Left: Mean annual risk of EVD emergence in humans across the study area. Low values are in yellow, high values in red, and grey zones were not included in the analysis. Right: Risk of EVD aggregated by latitude and month. Height and color intensity are proportional to the risk of EVD within the latitudinal window. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Model results in the context of the most recent EVD outbreaks. In 2018 there were two recognized EVD outbreaks in the Democratic Republic of the Congo (DRC). The district of Bikoro where the July 2018 outbreak occurred is highlighted in blue, and highlighted in orange is the district of Beni where the outbreak announced in August 2018 started. The index case for the Bikoro epidemic was believed to be in April, while the Beni outbreak is believed to have begun in July. Upper: Percent risk ranking against all cells and all months (not just the DRC). Lower: Relative risk of viral spillover. Left: Annual average risk by district health zone. Right: Risk with in the health districts throughout the year. Each line represents a cell within the district and the median value is bolded. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
(Fig. 6). Of the 16 cells that covered the Bikoro health district there was a median percent rank scores of 88.3% compared to all cells in all months. Relative risk of ebolavirus spillover fluctuated between minimum, median, and high of 14.9, 21.3 and 50.4 respectively. The more recent EVD outbreak occurred in the district of Beni, a persistently highly ranked region. The district in northwestern DRC near the border with Uganda announced the outbreak on first of August with the index case was presumably in July or June (WHO, 2018). The smaller region was covered by only nine cells and risk percent rank values extracted from the model covering July were among the highest ranking of all cells with a median of 99.6%, minimum of 98.0% and maximum of 99.8%. Relative risk for the region was considerable with minimal values of 93.0, a median of 195.0 and a high of 273 across the year. Percent rank and relative risk scores were very similar between $D + B_{tax}$ and $D + B_{tax}'$ models with the later preforming slightly better in both metrics at the Bikoro site. We ultimately chose the $D + B_{tax}$ model as our top model due to it's inclusion of more precise information related to our *a priori* hypotheses, and its comparable predictive power over all other alternative models.

4. Discussion

The spatial and temporal occurrence of ebolavirus on the landscape is complex, with seasonal drivers likely modifying the behaviors of both the (yet unidentified) maintenance reservoir host(s) and susceptible non-maintenance host populations. The inability of outbreak investigation teams to determine a definitive source of infection in all but a small number of cases has severely hindered the ability of public health officials to generate actionable measures against future EVD outbreak events. Infections from non-human spillover hosts to humans have only been confirmed from great apes and duikers, despite ebolavirus sequence detection in some Rodentia species (Morvan et al., 1999). Chiropteran species and the wide diversity of susceptible species within the biome (Leroy et al., 2004; Swanepoel et al., 1996). Here we predict the continental birthing patterns of African bats and find some statistical support for bats births, and African fruit bats in particular, being correlated with EVD outbreaks. Together we use these to predict the spatio-temporal risk of EVD outbreaks in non-human mammals and humans across the African continent.

The natural history, distributions and diversity of African bat species are broad yet through this analysis we are able to generate the first comprehensive spatio-temporal predictions for bat parturition events. Data availability ultimately required the use of several simplifying assumptions, most notably the consolidation of species into the taxonomic groups. This simplification creates scenarios in which within a single location two taxonomic members can give birth at different, non-overlapping times during a single birth pulse, or that a single species at two separate locations can give birth at different times again within a broader birth pulse. Despite this, the patterns predicted were largely consistent with previous work (Cumming and Bernard, 1997) with only minor deviations. The reduction in predictive signal from a biannual, or continual birthing above 10° is novel, although potentially an artefact of scare data availability at the northern most latitudes of our study extent. Our analysis also refined the latitudinal threshold from polyestrous to monoestrous from the 18°S to 10°S and was supported by more observations than the northern phenomenon. Data mining and collection highlighted the spatial and taxonomic bias of observations with extremely limited data availability for some of the largest, densest forests in continental Africa, and low representation of the most diverse taxonomic groups. However, models with terms describing the birth pulses of bats preformed highly for both the human and non-reservoir spillover host spatGLM. Furthermore the inclusion of the proportionally down-weighted birth term for the African fruit bats species improved predictive ability in most cases.

The discovery of statistical support for some African fruit bats species being correlated with ebolavirus outbreaks supports a number of previously disparate hypotheses (Hayman, 2015; Walsh et al., 2005; Han et al., 2016). The historical association of EVD spillover events and African fruit bats has been more anecdotal than factual in most cases, and to date no hunters targeting fruit bats have been confirmed index cases despite the widespread practice (Kamins et al., 2011; Mickleburgh et al., 2009; Peel et al., 2017). This may be because the species most hunted, Eidolon helvum (Kamins et al., 2011; Mickleburgh et al., 2009), is likely not a reservoir species, possibly due in part to its single annual birth pulse (in contrast to some other African fruit bats species) (Hayman et al., 2012, 2010) and unique cellular mechanisms (Ng et al., 2015). The large geographic spread by viral subtypes, such as that observed in the 2014 outbreak in West Africa, could be facilitated by other broad ranging African fruit bats such as the hammer-headed bat, whose population spans both Central to West African regions with little measurable population genetic structure (Hasanin et al., 2016). Until recently the only potential reservoir species from which EVD-causing Ebola virus viral fragments have been recovered were Pteropodidae members (Leroy et al., 2005; Olival and Hayman, 2014). Serological surveys have implicated molossid and non-molossid species, however serological data are difficult to interpret (Gilbert et al., 2013). Our analysis also identified molossid birthing events as a potential contributor to human risk of EVD spillover and serological findings are consistent with the patterns observed in MARV infected R. aegyptiacus (Pourrut et al., 2009; Amman et al., 2012). Recent detections of ebolavirus fragments in molossid species and our analyses suggest the role of molossid bats in ebolavirus sylvatic maintenance is an important area for potential research (Kupferschmidt, 2019; Goldstein et al., 2018).

Seasonality of EVD events has been noted (Schmidt et al., 2017), however the abiotic conditions (such as rainfall) used within these analyses are likely not directly driving spillover, especially for unstable RNA viruses in tropical environments (Fischer et al., 2015). We hypothesized viral persistence within bat populations themselves, meaning viral sylvatic dynamics, are also unlikely to be directly attributable to exogenous pressures. By exchanging the traditional abiotic metrics of seasonality with a derived metric such as seasonality of birth we are able to explore a mechanistically plausible driver of seasonality within the sylvatic EVD system (Altizer et al., 2006) and aim to understand both necessary and sufficient causes of EVD emergence and contributions of bats as reservoirs (Nishiura et al., 2009; Rothman, 1976). Temporally varying bat birth terms identified by both our non-human spillover host and human risk models support the hypothesis that birthing may facilitate transmission into non-human mammals (Pourrut et al., 2009; Schuh et al., 2017). In our best models for ebolavirus spillover the birthing term for both fruit bats and non-molossid microbats were significant. The later of the lag terms identified for the African fruit bats is consistent with the maternal derived anti-body (MDA) hypothesis of Pourrut et al. (2009) and Peel et al. (2018), dynamics predicted by generic filovirus models (Hayman, 2015), and observed in empirical MARV data (Amman et al., 2012).

Risk of spillover to non-human spillover hosts was strongly driven by human population density and responsible for much of the patchy geographic risk distribution (see SI Figure 4). This negative human population density-EVD risk relationship was still significant, but the effect size was much smaller, within the human system. This may be expected given many outbreaks in animals are identified in forest-dwelling species, such as gorillas, and concurrently with human cases where index cases are reported to have had contact with other species such as gorillas (Leroy et al., 2004). The correlation of non-bat mammalian diversity in both spillover models could indicate simply that both the bat hosts and the non-reservoir spillover hosts live in the biodiverse tropical forests. Temporally lagged terms that describe chains of infection, both from bats to non-humans spillover hosts and from them to humans, were significant in our models as expected. In contrast to other work, however, we did not find habitat fragmentation to be predictive of EVD spillover risk (Rulli et al., 2017), perhaps because we are not modeling changes in forest through time or because...
we are modeling risk at a much larger scale than those studies (but see also Wilkinson et al. (2018)) and thus Simpson’s paradox applies, using slightly different methods, or because the other terms in our models are more important. It is difficult to consider how instances of Simpson’s paradox may be influencing our results, but clearly the interactions between habitat change, bat birthing, ebolavirus emergence and disease emergence in general are areas for future research (Rulli et al., 2017; Wilkinson et al., 2018; Han et al., 2016). Data sets to test these relationships, such as one populated with animal infections independent of human outbreak events, however, present serious hurdles in these regards.

The DRC has seen more EVD outbreak events than any other country, and the two most recent outbreaks exemplify the country’s ongoing struggle to combat a complex system. While the seasonality of relative risk is largely similar between the two locations, there are distinct and potentially important local peaks and troughs within the local seasonality of risk for each location and both outbreaks occurred at, or close to, local maximal values of relative risk. These seasonal patterns may be linked to the biannual birth pulses of the local fruit bat or molossid species, but are also likely tied to extremely local shifts in behavioral, phenological, and climatic factors that modify both the reservoir host(s) and susceptible populations.

Analyses of EVD emergence, including ours, are clearly limited by sparse data. This includes information on other possible host species such as apes, swine and other small mammals, which may warrant further investigation if data become available. Most recent serological analyses of African primates, however, support our assumptions that they are unlikely to be maintenance hosts (Ayouba et al., 2019; Hayman, 2019). The lack of detailed species specific bat birthing data hampered our ability to test some hypotheses and necessitated the aggregation of data by host taxonomic traits rather than a species, or genus level. While these groupings have proved useful (Cumming and Bernard, 1997), we do acknowledge that the patterns discussed may not be fully representative for all members (i.e. E. helvum) and was our primary motivation for creating down-weighted variation. However, despite individual bat variations, our model still broadly captures when species such as E. helvum give birth in space and time, even if the particular species only gives birth annually (Peel et al., 2017). Despite the significance of only fruit bat associated diversity terms, non-molossid microbats are under sampled virologically (Leendertz et al., 2016) and ecologically (this study) for such a diverse taxonomic group. Increased data may change our understanding of their role in EVD, especially in light of the recent news of ebolavirus viral fragments in the Miniopterus inflatus microbat and isolation of a potentially new Ebolavirus species by Goldstein et al. (2018) from a molossid species, demonstrate the undiscovered viral diversity within the system. We were also unable to include bat abundance terms, and clearly host abundance will affect the likelihood of viral persistence in populations (Hayman, 2015; Pourrut et al., 2009). These ecological gaps, however, can be filled without additional virological work, although contemporary ecological and virological data collection will provide advantages for data integration and improved system understanding (Rulli et al., 2017; Pourrut et al., 2009; Amman et al., 2012; Restif et al., 2012). Because of the limited number of viral detections, all Ebolavirus species were combined into a generic definition of ebolavirus despite the probably that each species has its own intricacies of sylvatic persistence, distribution, reservoir host and spillover. Further, the reliance on symptomatic EVD has increased these biases if viruses had different pathological consequences in man. Only increased virological studies in hosts and unfortunately outbreaks in people will allow us to disentangle these species level intricacies. Despite these weaknesses, we believe we have successfully captured the overall patterns of birth pulses of African bats and provide potentially useful insights in the ecology of ebolaviruses and how seasonal changes in EVD risk may be driven by bat reservoir host ecology.

Author contributions

D.T.S.H. conceived the study; D.T.S.H., C.R.H., J.C.M. developed the methods; C.R.H. & A.M. compiled the data; C.R.H. & J.C.M. wrote the code; C.R.H. ran the analysis; A.M. provided critiques; C.R.H. wrote the first draft with all authors editing and reviewing the manuscript.

Funding

D.T.S.H would like to acknowledge funding by Rutherford Discovery Fellowship RDF-MAU1701 and Marsden MAU1503.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

We would like to acknowledge Dr. David Wilkinson for all his help, along with the many reviewers who helped.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.epidem.2019.100366.

References

Grenfell, B.T., Bjørnstad, O.N., Kaperj, J., 2001. Travelling waves and spatial hierarchies...