Published in final edited form as:

J Vector Ecol. 2013 December ; 38(2): 251–259. doi:10.1111/j.1948-7134.2013.12038.x.

A statistical model of Rift Valley fever activity in Egypt

John M. Drake^{1,2}, Ali N. Hassan³, and John C. Beier⁴

John M. Drake: jdrake@uga.edu

¹Eugene P. Odum School of Ecology, University of Georgia, Athens, GA 30602, U.S.A.

²Department of Zoology, Oxford University, Oxford, UK

³Institute of Environmental Studies and Research, Ain Shams University, Cairo, Egypt

⁴Department of Epidemiology and Public Health, University of Miami Miller School of Medicine, Miami, FL, U.S.A.

Abstract

Rift Valley fever (RVF) is a viral disease of animals and humans and a global public health concern due to its ecological plasticity, adaptivity, and potential for spread to countries with a temperate climate. In many places, outbreaks are episodic and linked to climatic, hydrologic, and socioeconomic factors. Although outbreaks of RVF have occurred in Egypt since 1977, attempts to identify risk factors have been limited. Using a statistical learning approach (lasso-regularized generalized linear model), we tested the hypotheses that outbreaks in Egypt are linked to (1) River Nile conditions that create a mosquito vector habitat, (2) entomologic conditions favorable to transmission, (3) socio-economic factors (Islamic festival of Greater Bairam), and (4) recent history of transmission activity. Evidence was found for effects of rainfall and river discharge and recent history of transmission activity. There was no evidence for an effect of Greater Bairam. The model predicted RVF activity correctly in 351 of 358 months (98.0%). This is the first study to statistically identify risk factors for RVF outbreaks in a region of unstable transmission.

Keyword Index

Rift valley fever; forecast; Egypt; regularized regression

Introduction

Rift Valley fever (RVF) is a vector-borne viral zoonosis that infects livestock and humans (Flick and Bouloy 2005). Economic losses in endemic regions can be great due to death of the host and/or abortion in livestock, particularly cattle, sheep, camels, and goats. Recent epidemics have resulted in tens to hundreds of thousands of human cases, of which <1% are fatal (Flick and Bouloy 2005). Transmission of RVF virus (RVFV) is typically through the bite of an infected mosquito or direct contact with body fluids of infected animals during slaughter. The ecology of RVFV varies greatly among regions. In areas of endemism, like sub-Saharan Africa, primary vectors are biting mosquitoes in the genus *Aedes*, which may also transmit the virus vertically, enabling persistence during periods of drought (Linthicum et al. 1999, Favier et al. 2006, Métras et al. 2011). In the Horn of Africa, outbreaks are associated with extreme climatic conditions, particularly above average rainfall associated with the warm phase of the El Niño/Southern Oscillation (Linthicum et al. 1999, Linthicum et al. 2007, Anyamba et al. 2009). Ecologic forecasting of outbreaks in endemic regions has

been well studied and is widely considered to be effective (Anyamba et al. 2009). By contrast, in North Africa and the Arabian peninsula, including Egypt, *Culex* spp. are the dominant vectors (Hoogstraal et al. 1979, Meegan et al. 1980, Hanafi et al. 2011). *Culex* do not transmit transovarially however, and persistence requires a permissive vertebrate reservoir. In these regions, outbreaks are more likely to be associated with importation of new infections and secondary transmission than with amplification of viruses already circulating (Abdo-Salem et al. 2011). A reliable model to forecast the origin and duration of outbreaks of Rift Valley fever in such regions would be of considerable value to public health planning, the targeting of vector control operations, and reducing economic losses. The development of such models is an active area of research.

Although Rift Valley fever has historically been limited to sub-Saharan Africa, introduction to other parts of the world is a concern (EFSA 2005, Linthicum et al. 2007, Hartley et al. 2011). Particularly, sporadic epidemics/epizootics have occurred in Egypt since 1977, when RVFV appeared in southern Egypt and then spread to the Nile Delta causing a severe outbreak in humans (~200,000 human infections and ~600 deaths) and livestock (Meegan 1979, Darwish and Hoogstraal 1981). Currently, Egypt represents the northernmost extent of RVFV. Unsurprisingly, transmission of RVFV in Egypt is unstable, giving rise to sporadic outbreaks which presumably occur only when the pathogen is introduced under the right entomologic and hydrologic conditions. Although there have been five major outbreaks of Rift Valley fever in Egypt since 1977 (Table 1), attempts to discern risk factors associated with outbreaks have been limited (Métras et al. 2011). Hypothesized risk factors include hydrologic conditions favorable to the creation of mosquito vector habitat and socio-economic factors favorable to the importation of the pathogen (Hoogstraal et al. 1979, Sellers et al. 1982, Gad et al. 1986), both of which have been identified as risk factors in other regions (Métras et al. 2011, Abdo-Salem et al. 2011). Particularly, since the first two Rift Valley fever epidemics in Egypt coincided with the timing of the Greater Bairam, an Islamic religious festival during which animal sacrifice and feasting are customary, we hypothesized that the timing of the Greater Bairam, when large numbers of livestock are imported from Sudan and sub-Saharan Africa, might be a risk factor (Abdo-Salem et al. 2011). This hypothesis is consistent with a recent study in which chance of introduction of the virus to Yemen was considered greater during festival periods (Abdo-Salem et al. 2011). Since importation of livestock for the Greater Bairam occurs in advance of the festival, we supposed that introduction of the virus might lead the Greater Bairam holiday by as much as two months. The presence of entomologic conditions favorable to the transmission of the virus, previously reported to correlate with RVF transmission activity (Kenawy et al. 1987, Gad et al. 1999, Hanafi et al. 2011), was also included as a covariate. Finally, if the virus is either temporarily or permanently persistent, then recent transmission activity should be a good predictor of current transmission intensity. Lagged RVF status was therefore also considered an important candidate predictor.

To test these hypotheses and construct a statistical forecasting system, we trained a statistical model to make one-month-ahead predictions of the status of Rift Valley fever in Egypt. We constructed a data set of 24 potential predictors representing a range of social, hydrologic, and entomologic factors. As the number of candidate predictors is of the same order as the number of months during which RVF was active, the associated regression model is poorly conditioned. Therefore, we adopted a regularization procedure (lasso-regularized generalized linear model; Park and Hastie 2007, Friedman et al. 2010) commonly used in statistical learning to select the best subset of predictors. The objectives of our model were (1) to predict the continuation (alternately, termination) of outbreaks, and (2) to predict the onset of new outbreaks. The model was optimized to provide the maximum performance in leave-one-out cross-validation. Although this model generally failed to predict the starting point of an epidemic, the ongoing status of infection-free vs infected

state was predicted with 98% accuracy; the month of the termination of outbreaks was predicted correctly in three out of the five outbreaks that occurred during the 30-year study period.

Materials and Methods

Response variables

Two response variables were considered in this study. First, as an aid to short-term forecasting (continuation or termination of outbreak), all months between August, 1975 and July, 2005 ($n=360$) were scored as RVF+ (evidence of local transmission of Rift Valley fever to either animals or humans as indicated by official authorities and/or published work, usually including Ministry of Health and/or Ministry of Agriculture officials as authors, particularly references (Hoogstraal 1978, Hoogstraal et al. 1979, Laughlin et al. 1979, Meegan 1979, Meegan et al. 1980, Darwish and Hoogstraal 1981, Sellers et al. 1982, Arthur et al. 1993, Abd el-Rahim et al. 1997, Gad et al. 1999, Okda et al. 2006, Kamal 2011, Hanafi et al. 2011) or RVF- (absence of evidence of local transmission Rift Valley fever). Second, as an aid to anticipating new outbreaks (onset), the time series of RVF activity was broken into intervals of consecutive activity, i.e., runs of RVF+ or RVF-. From these intervals it was determined if each month was the start of a new outbreak of Rift Valley fever, i.e., the first month in a series scored RVF+.

Predictor variables

Hydrologic data were obtained from records of the Ministry of Water Resources and Irrigation for a range of locations upstream and downstream of Aswan High Dam in Egypt, where conditions might contain the signal of future outbreak conditions (Figure 1):

Mosquito activity: An indicator variable for the seasonal fluctuations in abundance of *Cx. pipiens* and *Cx. antennatus* mosquitoes (December-April=0, May-November=1)

Greater Bairam: An indicator variable for the celebration of Greater Bairam

Rift Valley fever: An indicator variable for the detection of RVF transmission in Egypt (used later to construct lagged variables that can be used in prediction)

RVF × Mosquito: An indicator variable taking the value 1 if both *Mosquito Activity* and *Rift Valley fever* are 1; otherwise 0

Greater Bairam × Mosquito: An indicator variable taking the value 1 if both *Greater Bairam* and *Mosquito* are 1; otherwise 0

Monthly rainfall at Gambeila: Continuous variable in units of mm precipitation month⁻¹ (mean: 125.1; sd: 97.6)

Monthly river discharge at Malakal station: Continuous variable in millions of m³ month⁻¹ (mean: 2584.4; sd: 720.1)

Monthly river discharge at Dongola station: Continuous variable in millions of m³ month⁻¹ (mean: 5775.0; sd: 5719.7)

Natural inflow at Aswan Dam: Continuous variable in millions of m³ month⁻¹ (mean: 6991.9; sd: 5741.3)

Total water arriving at Aswan Dam: Continuous variable in millions of m³ month⁻¹ (mean: 5957.5; sd: 5597.6)

Monthly river discharge at Gaafra: Continuous variable in millions of m³ month⁻¹ (mean: 4821.9; sd: 1460.1)

Monthly river discharge at Damietta: Continuous variable in millions of m³ month⁻¹ (mean: 816.6; sd: 395.2)

Monthly river discharge at Rosetta: Continuous variable in millions of m³ month⁻¹ (mean: 525.9; sd: 515.9)

Average water level at Malakal station: Continuous variable in m (mean: 11.8; sd: 0.8)

Average water level at Dongola station: Continuous variable in m (mean: 10.6; sd: 1.6)

Maximum monthly water level at Lake Nasser: Continuous variable in m (mean: 172.2; sd: 6.2)

Average water level at Gaafra: Continuous variable in m (mean: 82.7; sd: 0.9)

Average water level at Damietta: Continuous variable in m (mean: 13.3; sd: 0.3)

Average water level at Rosetta: Continuous variable in m (mean: 13.1; sd: 0.3)

One record was missing for *Monthly Rainfall at Gambeila* (July, 2004) and was replaced by the average July rainfall calculated over all other years. To enable evaluation of the relative importance of selected variables, continuous variables were rescaled by dividing each observation by the variable standard deviation. Thus, both continuous and categorical variables were confined to a comparable range. Additional variables were then constructed from this list, including:

RVF-Lag1: A one-month lagged variable for RVF status and mosquito density

GB-Lag1 and *GB-Lag2*: One- and two-month lagged variables for Greater Bairam holidays

RVF-Lag1 $\times$ *Mosquito-Lag0*: A constructed feature positive for positive RVF status at lag one month and high mosquito density at current time

RVF-Lag1 $\times$ *Mosquito-Lag1*: A constructed feature positive for RVF status at lag one month and high mosquito density at lag one

Statistical Modeling

We initially studied models estimated using linear discriminant analysis, recursive partitioning, and generalized linear models. In these models, variable selection and model structure were guided by iterative fitting, inspection of coefficients, and significance tests. Ultimately, this approach proved too unwieldy. We therefore turned to a systematic search via lasso-regularized generalized linear modeling (Park and Hastie 2007, Friedman et al. 2010). One interpretation of this estimation procedure is that it is as a penalization or “shrinkage” method that trades bias for reduced variance, i.e., the unbiasedness of the maximum likelihood estimates is lost, but precision on most or all estimated coefficients is improved (Hastie et al. 2009). A desirable consequence of this estimation scheme is that it favors sparse solutions to overdetermined regression problems, i.e., coefficients on many variables shrink to zero (Hastie et al. 2009). Since a coefficient of zero implies no effect, this feature may be exploited for variable selection in addition to improving model identifiability.

Recognizing that hypothesized interactions are already coded in the variables *RVF* $\times$ *Mosquito* and *Greater Bairam* $\times$ *Mosquito*, we fit a binomial model with first order terms only to the combined data set consisting of the original and constructed variables. Lasso regularization requires setting a tuning parameter (λ) that governs the severity of the penalty function (Hastie et al. 2009). To optimize the model, we searched for the value of λ that minimized average residual deviance (equivalently, minimized classification error) on the

withheld observation by iteratively fitting the model to all points but one (leave-one-out cross-validation; Hastie et al. 2009). Because we explored covariates with time lags up to two months, the number of records available for model fitting was reduced from $n=360$ to $n=258$. As predictor variables include lagged covariates, it was impossible to avoid information leakage from test data. Given this and the additional constraints imposed by the relatively small number of observations, we determined that model performance should be reported as the fit of the model to all the training data after optimization using leave-one-out cross-validation rather than by designating a test data set withheld from model fitting. We used leave-one-out cross-validation and our estimates of model performance were calculated for the entire data set, but in each case were calculated for records excluded from model fitting. For comparison, we also estimated null accuracy, which we define as the accuracy the model would have achieved if we had simply chosen to classify every point according to the majority class (in our case RVF-). Fitting was performed with the 'glmnet' package in R (R Development Core Team 2008). Data and source code to perform the analysis may be obtained from the lead author of this article.

Results

Performance of models

Although the best fit model for predicting the continuation of RVF activity was obtained at $\lambda=0.00729$, the range of values for λ that result in nearly equivalent models is broad (Figure 2). This suggests that many variables are nearly equivalent in their predictive ability. The model that minimized average deviance on withheld observations was selected for further analysis. Variables in this model included *Monthly rainfall at Gambeila*, *Monthly discharge at Gaafra*, *RVF-Lag1* $\times$ *Mosquito-Lag0*, and *RVF-Lag1* $\times$ *Mosquito-Lag1*. Performance of this model was high overall, but selective. Specifically, the model predicted RVF activity correctly in 351 of 358 months (98.0% accuracy), achieved with a model specificity of 99.4% and model sensitivity of 84.4%. This performance was achieved mostly by correct prediction of the continuation of outbreaks and inter-outbreak intervals. To illustrate, Figure 3 shows rainfall at Gambeila (the most important hydrologic indicator detected in this analysis) plotted through time with outbreak intervals designated in red. Misclassified points are circled, showing that the start of new outbreaks is considerably more difficult to predict than the continuation or termination of an outbreak. Since both categorical and continuous observations are scaled to a common range, variable importance may be evaluated by inspection of the magnitudes of the fit coefficients. Table 2 shows that importance of *RVF-Lag1* $\times$ *Mosquito-Lag1* exceeds that of the next most important variable (*Monthly rainfall at Gambeila*) by more than eight-fold. No variables could be identified that predicted the onset of the epidemic, either when all RVF active months were used or when the model was trained to classify only onset months.

Comparison with alternative models

The null accuracy, i.e., the accuracy the model would have achieved if we had simply chosen to classify every point according to the majority class, for this model was 91% (sensitivity: 0%, specificity: 100%). Further insight into the parsimony of the best fit model is obtained by comparing using the simple rule, "If *RVF-Lag1*=1 then RVF+, otherwise RVF-". Such a rule-of-thumb, although superior to the null model, only achieves accuracy of 97.2% (sensitivity: 84.4%, specificity: 98.5%). These comparisons show the statistically optimized model to be superior to these naïve alternatives.

Discussion

The development of statistical systems for predicting emerging and re-emerging infectious diseases is an active area of research. Here we have shown that a statistical model can predict the continuation and termination of outbreaks of Rift Valley fever in Egypt, a region where transmission is unstable but periodic introduction and subsequent outbreaks have caused a significant public health and economic burden. Specifically, months were misclassified only seven times in nearly thirty years of forecasts. These findings provide evidence that a statistical system for Rift Valley fever may be capable of accurate forecasting of the duration of outbreaks after they have begun.

We hypothesize that RVF emerges in Egypt when a constellation of interconnected hydrologic, entomologic, and social conditions are met. Due to the different periods of the hydrologic cycle (which determine mosquito abundance) and the Islamic calendar (which determines the dates of the Greater Bairam, the putative condition for importation of the virus, and possibly also increases the population of susceptible livestock thereby boosting amplification), the circumstances enabling the emergence of RVF are infrequently and irregularly met. The predictors identified by our model relate to hydrologic conditions occurring both upstream (rainfall at Gambeila) and downstream (discharge at Gaaфра) of Aswan High Dam and their effects on vectors and disease prevalence. Previous studies indicate that the sudden appearance or a surge in quantity of environmental water is associated with outbreaks or resurgence through increasing vector populations (Digoutte and Peters 1989, Mondet et al. 2005). Causes of such increases in water quantities include, *inter alia*, unusually heavy rainfall (Linthicum 1999), flooding following dam construction (Digoutte and Peters 1989), and rainfall patterns typical of the rainy season (Mondet et al. 2005). Our finding that the continuation of RVF outbreaks is well predicted by river discharge is not surprising as the water budget for much of Egypt, which is heavily irrigated, is dependent on rainfall at upper Nile reaches including Gambeila (Said 1993). Once in Egypt, water discharges downstream to the Nile Valley and Delta depend on normal requirements as well as storage capacity of the Aswan Dam. In cases where water storage capacity reaches critical levels, more water is released to reduce physical pressure on the body of the Dam (Aziz and Sadek 2003). Thus, with above normal rainfall in Africa, Egypt receives higher budgets allowing for more surface water availability downstream. Linkages between water availability, creation of mosquito breeding habitats, and increased RVF vector species abundance are well documented (Davies et al. 1985, Bicout and Sabaïter 2004, Elfadil et al. 2006, Abdo-Salem et al. 2011).

Our demonstration of a strong statistical association between RVF activity and both upstream and downstream hydrological factors in Egypt highlights the importance of hydrology in RVF epidemiology (and probably for other vector-borne diseases, too). RVF virus is present in a variety of ecosystem types, including arid areas (Northern Senegal, Yemen), irrigated areas (Egypt/Aswan, Delta of Senegal River), forest (Madagascar), and wetlands called dambos (Kenya, South Africa; Métras et al. 2011). In Egypt, the 1977 and 1993 RVF outbreaks occurred at the time of the first and second filling of the Aswan High Dam to spillway elevation following high floods in Africa (Jobin 1999). A series of successive high floods occurred in the years 1999-2000, 2000-2001, and 2001-2002 (Amer 2003) which may have contributed to the 2003 RVF epidemic. In all studies analyzed by Métras et al. (2011), extrinsic factors associated with RVF or RVF virus exposures were found to be water-related. However, the mechanism(s) through which hydrological factors influence vector populations and disease events vary considerably among regions because the ecological habits of the vector species, hydrological systems, and water management practices differ. For example, rainfall in Africa is linked to RVF epidemics through inundation of dambos and hatching of vertically infected *Aedes* spp. females (Linthicum et

al. 1999). Here we show that rainfall in Africa is also linked to RVF activities in Egypt, in areas located many miles away through the River Nile hydrological system. Such remote hydrological linkages (ca. 5,500-6,500 km) help to explain how large-scale climate-related events like ENSO lead to vector-borne disease outbreaks (Linthicum et al. 1999, Anyamba et al. 2009).

On the vector side, ongoing work by one of the authors (ANH) demonstrates that the hydrologic control system of Aswan High Dam may contribute to temporal abundance of mosquito vectors. The operation rules governing the discharge of water into Egypt allow the flow of calculated quantities that vary throughout the year according to water requirements and storage capacity. Interestingly, the water discharged downstream increases gradually starting from the months of March-April each year and peaks in September. Maximum density of *Culex pipiens* and *Cx. antennatus*, two known vectors of RVF (see Hoogstraal et al. (1979) and Hanafi et al. (2011), respectively), were found to follow such a pattern of increase in water levels by a one-month lag period (ANH, unpublished data). Hence, hydrological monitoring may continue to provide a useful tool for vector and disease surveillance and prediction.

Recent RVF activity is another significant factor identified in our study. Maintenance of the virus during inter-epidemic periods is still an epidemiological mystery, not only in Egypt but also in other areas with a similar epidemiological situation. In areas where floodwater mosquitoes of the genus *Aedes* exist, vertical transmission through drought-resistant eggs is responsible for both maintenance and emergence of the virus in the system. Our laboratory observations on *Aedes (Ochlerotatus)* spp. in Egypt indicate that their eggs may resist desiccation and hatch when soaked in water (ANH, unpublished observations). Under laboratory conditions, *Ae. caspius* has been shown to be capable of transmitting the virus (Turell et al. 1996, Gad et al. 1987). Under field conditions, this species was suggested to play a role in RVF transmission as a bridge vector between humans and bovines (Gad et al. 1999). The ability of Aedine species to vertically transmit RVF virus in Egypt is an important topic for future research.

Although Greater Bairam was not among the significant RVF predictors in this study, we suggest that the contribution of festive occasions to RVF emergence and resurgence should nevertheless remain an open topic for study (compare with Gad et al. (1986), Jobin (1999), and Davies (2006)). Particularly, animal importation may be associated with virus introduction/re-introduction, increased virus amplification in sheep populations, increased vector foraging behavior related to increased host availability, and/or higher human exposure. Kamal (2011) suggested that importation of animals infected with RVF from Sudan is among the main causes of resurgence of RVF epizootics in Egypt and that non-compliance with rules that imported live animals were to be slaughtered upon arrival leads to such animals being mixed with native breeds, creating conditions for spread throughout the country. This provides support to earlier studies implicating animal movement between Sudan and Egypt (whether legal or illegal) in RVF virus introduction (Gad et al. 1986). In fact, the soon-to-be-operational land transportation route between the two countries would facilitate animal movement; hence increasing the risk of outbreaks.

The onset of RVF outbreaks in areas of unstable transmission is not only influenced by natural and socioeconomic phenomena but is also complicated by a country's political agenda and vigilance of health systems. Due to fears of potential economic losses, countries may prevent or delay the official announcement of an outbreak due to potential negative impacts on the economy (Kamal 2011). This situation limits the availability of precise information about outbreak events. For example, Egypt has not officially recognized the RVF epidemic of 2003 (Hanafi et al. 2011, Kamal 2011). Additionally, the onset of some

outbreak events in Egypt have been attributed to live animal vaccine (Kamal 2011), although not the first and largest outbreak, which occurred prior to the use of live animal vaccine. It follows that modeling capability to predict the onset of RVF events in Egypt would improve with proper epidemiological information and with alternative modeling approaches. We hypothesize that Rift Valley fever occurs in Egypt when a constellation of entomologic, hydrologic, and social conditions are simultaneously met. Due to the different periods of the hydrologic cycle (which determines mosquito abundance) and the Islamic calendar (which determines the dates of the Greater Bairam, the putative condition for importation of the virus in infected livestock, as well as large numbers of livestock that may accelerate virus amplification), the circumstances enabling the emergence of Rift Valley fever are infrequently and irregularly met. A better understanding of this interaction will be crucial for developing more effective forecasting systems in non-endemic regions.

The failure of this model to predict the onset of outbreaks does not diminish its potential use in anticipating the continuation and termination of outbreaks. We propose that potential improvements to the model might be obtained through the development of mechanistic models for transmission (Favier et al. 2006, Gaff et al. 2007). Specifically, since the hypothesized emergence scenario is specific, one promising approach is to parametrically encode the timing of introduction into a model for transmission, which could then be fit to data such as have been reported here. Such a model would benefit greatly, however, from quantitative information about the relative prevalence of Rift Valley fever or transmission activity.

One potential explanation for the failure of our model to anticipate the onset of outbreaks is that they were caused by live animal vaccine (Kamal 2011), not introduction in livestock, although clearly this explanation would not apply to outbreaks that occurred before the use of vaccines. This would explain why outbreaks do not occur annually when entomologic conditions are met. Of course, importation of susceptible (non-vaccinated) livestock hosts might still be a necessary condition for amplification, in which case a threefold set of conditions must be met for an outbreak to begin (ecologic and conditions favorable to the vector, expectation of Greater Bairam building up levels of susceptible hosts, and reversion of attenuated vaccine-derived virus to an infectious state to initiate the outbreak). Since live-virus vaccines are no longer in use, it is anticipated that future outbreaks are more likely to be sparked by introductions. The very good performance of our model during epidemic periods underscores its usefulness regardless of the mechanism of introduction, an important consideration in regions where transmission is unstable.

Acknowledgments

The authors thank Chris Cosner and Kenneth Linthicum for comments on an earlier draft of this paper. This study was supported by a grant from the National Institutes of Health (R01GM093345).

References Cited

- Abd el-Rahim, Abd el-Hakim U, Hussein M. An epizootic of Rift Valley fever in Egypt in 1997. *Rev Sci Tech OIE*. 1997; 18:741–748.
- Abdo-Salem S, Waret-Szkuta A, Roger F, Olive MM, Saeed K, Chevalier V. Risk assessment of the introduction of Rift Valley fever from the Horn of Africa to Yemen via legal trade of small ruminants. *Trop Anim Health Prod*. 2011; 43:471–480. [PubMed: 20967567]
- Amer, MH. General Authority for Aswan High Dam and Reservoir, Ministry of Water Resources and Irrigation. Egypt: 2003. Toshka spillway and depression and its role in flood control (Arabic); p. 167

- Anyamba A, Chretien JP, Small J, Tucker CJ, Formenty PB, Richardson JH, Britch SC, Schnabel DC, Erickson RL, Linthicum KJ. Prediction of a Rift Valley fever outbreak. *Proc Natl Acad Sci U S A*. 2009; 106:955–959. [PubMed: 19144928]
- Arthur RR, el-Sharkawy MS, Cope SE, Botros BA, Oun S, Morrill JC, Shope RE, Hibbs RG, Darwish MA, Imam IZ. Recurrence of Rift Valley fever in Egypt. *Lancet*. 1993; 342:1149–1150. [PubMed: 7901480]
- Aziz M, Sadek N. The proposed operation for the new Toshka barrage. *Engineer Res J*. 2003; 89:85–101.
- Bicout DJ, Sabaiter P. Mapping Rift Valley fever vectors and prevalence using rainfall variations. *Vector-Borne Zoonot Dis*. 2004; 4:33–42.
- Dar O, McIntyre S, Hogarth S, Heymann D. Rift Valley fever and a new paradigm of research and development for zoonotic disease control. *Emerg Infect Dis*. 2013; 19:189–193. [PubMed: 23347653]
- Darwish M, Hoogstraal H. Arboviruses infecting humans and lower animals in Egypt: A review of thirty years of research. *J Egypt Publ Hlth Assoc*. 1981; 56:1–112.
- Davies FG. Risk of Rift Valley fever epidemic at the haj in Mecca, Saudi Arabia. *Rev Sci Tech OIE*. 2006; 25:137–147.
- Davies FG, Linthicum KJ, James AD. Rainfall and epizootic Rift Valley fever. *Bull Wld Hlth Org*. 1985; 63:941–943.
- Digoutte JP, Peters CJ. General aspects of the 1987 Rift Valley fever epidemic in Mauritania. *Res Virol*. 1989; 140:27–30. [PubMed: 2711042]
- EFSA. The risk of a Rift Valley fever incursion and its persistence within the community. *EFSA J*. 2005; 238:1–128.
- Elfadil AA, Hasab-Allah KA, Dafa-Allah OM. Factors associated with Rift Valley fever in south-west Saudi Arabia. *Rev Sci Tech OIE*. 2006; 25:1137–1145.
- Favier C, Chalvet-Monfray K, Sabatier P, Lancelot R, Fontenille D, Dubois MA. Rift Valley fever in West Africa: the role of space in endemicity. *Trop Med Int Hlth*. 2006; 11:1878–1888.
- Flick R, Bouloy M. Rift Valley fever virus. *Curr Mol Med*. 2005; 5:827–834. [PubMed: 16375716]
- Fontenille D, Taore-Lamizana M, Daillo M, Thonnon J, Digoutte JP, Zeller HG. New vectors of Rift Valley fever in West Africa. *Emerg Infect Dis*. 1998; 4:289–293. [PubMed: 9621201]
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010; 33:1–22. [PubMed: 20808728]
- Gad AM, Farid HA, Ramzy RR, Riad MB, Presley SM, Cope SE, Hassan MM, Hassan AN. Host feeding of mosquitoes (Diptera: Culicidae) associated with the recurrence of Rift Valley fever in Egypt. *J Med Entomol*. 1999; 36:709–714. [PubMed: 10593070]
- Gad AM, Feinsod FM, Allam IH, Eisa M, Hassan AN, Soliman BA, el Said S, Saah AJ. A possible route for the introduction of Rift Valley fever virus into Egypt during 1977. *J Trop Med Hyg*. 1986; 89:233–236. [PubMed: 3795324]
- Gad AM, Hassan MM, Sharif ES, Mahmoud IM, Owen LW. Rift Valley fever virus transmission by different Egyptian mosquito species. *Trans R Soc Trop Med Hyg*. 1987; 81:694–698. [PubMed: 2895516]
- Gaff HD, Hartley DM, Leahy NP. An epidemiological model of Rift Valley fever. *Electron J Diff Equat*. 2007; 2007:1–12.
- Gerdes GH. Rift Valley fever. *Vet Clin Food Anim*. 2002; 18:549–555.
- Gerdes GH. Rift Valley fever. *Rev Sci Tech*. 2004; 23:613–623. [PubMed: 15702723]
- Grobbelaar AA, Weyer J, Leman PA, Kemp A, Paweska JT, Swanepoel R. Molecular epidemiology of Rift Valley fever virus. *Emerg Infect Dis*. 2011; 17:2270–2276. [PubMed: 22172568]
- Hanafi HA, Fryauff DJ, Saad MD, Soliman AK, Mohareb EW, Medhat I, Zayed AB, Szumlas DE, Earhart KC. Virus isolations and high population density implicate *Culex antennatus* (Becker) (Diptera: Culicidae) as a vector of Rift Valley Fever virus during an outbreak in the Nile Delta of Egypt. *Acta Trop*. 2011; 119:119–124. [PubMed: 21570939]
- Hartley DM, Rinderknecht JL, Nipp TL, Clarke NP, Snowden GD. Potential effects of Rift Valley fever in the United States. *Emerg Infect Dis*. 2011; 17:e1. [PubMed: 21801607]

- Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Precision. 2nd. Springer, NY: 2009.
- Hoogstraal H. Rift Valley fever: An historical perspective. *J Egypt Publ Hlth Assoc.* 1978; 53:129–135.
- Hoogstraal H, Meegan JM, Khalil GM, Adham FK. The Rift Valley fever epizootic in Egypt 1977–1978 2. Ecological and entomological studies. *Trans R Soc Trop Med Hyg.* 1979; 73:624–629. [PubMed: 44038]
- Jobin, WR. Dams and Disease: Ecological Design and Health Images of large Dams, Canals, and Irrigation Systems. CRC Press; 1999. p. 600
- Kamal SA. Observations on Rift Valley fever virus and vaccines in Egypt. *Virol J.* 2011; 8:532. [PubMed: 22152149]
- Kenawy MA, Beier JC, Zimmerman JH, Said SE, Abbassy MM. Host-feeding patterns of the mosquito community (Diptera: Culicidae) in Aswan Governorate, Egypt. *J Med Entomol.* 1987; 24:35–39. [PubMed: 3820238]
- Laughlin LW, Meegan JM, Strausbaugh LJ, Morens DM, Watten RH. Epidemic Rift Valley fever in Egypt: observations of the spectrum of human illness. *Trans R Soc Trop Med Hyg.* 1979; 73:630–633. [PubMed: 575446]
- Linthicum KJ, Anyamba A, Tucker CJ, Kelley PW, Myers MF, Peters CJ. Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya. *Science.* 1999; 285:397–400. [PubMed: 10411500]
- Linthicum KJ, Anyamba A, Britch SC, Chretien JP, Erickson RL, Small J, Tucker CJ, Bennett KE, Mayer RT, Schmidtman ET, Andreadis TG, Anderson JF, Wilson WC, Freier JE, James AM, Miller RS, Drolet BS, Miller SN, Tedrow CA, Bailey CL, Strickman DA, Barnard DR, Clark GG, Zou L. A Rift Valley fever risk surveillance system for Africa using remotely sensed data: potential for use on other continents. *Vet Ital.* 2007; 43:663–674. [PubMed: 20422546]
- Meegan JM. The Rift Valley fever epizootic in Egypt 1977–78. 1. Description of the epizootic and virological studies. *Trans R Soc Trop Med Hyg.* 1979; 73:618–623. [PubMed: 538803]
- Meegan JM, Khalil GM, Hoogstraal H, Adham FK. Experimental transmission and field isolation studies implicating *Culex pipiens* as a vector of Rift Valley fever virus in Egypt. *Am J Trop Med Hyg.* 1980; 29:1405–1410. [PubMed: 7446827]
- Metras R, Collins LM, White RG, Alonso S, Chevalier V, Thuramira-McKeever C, Pfeiffer DU. Rift Valley fever epidemiology, surveillance, and control: what have models contributed? *Vector Borne Zoonotic Dis.* 2011; 11:761–771. [PubMed: 21548763]
- Mondet B, Diaité A, Fall AG, Chevalier V. Relations between rain patterns and risk of virus transmission by mosquitoes: Rift Valley fever (RVF) virus in the Ferlo region of Senegal. *Environnement, Risques Santé.* 2005; 4:125–129.
- Okda M, Badier S, Ashour I, Abo Gabal A. An Epidemic of Rift Valley encephalitis in the region of Khafer Al-Sheikh between the period of October 2003 to April 2004. *Egypt J Neurol Psychiat Neurosurg.* 2006; 43:399–407.
- Park MY, Hastie T. L1-regularization path algorithms for generalized linear models. *J R Stat Soc B.* 2007; 69:659–677.
- Said, R. The River Nile: Geology, Hydrology and Utilization. Pergamon Press Ltd; Oxford, UK: 1993.
- Sellers RF, Pedgley DE, Tucker MR. Rift Valley fever, Egypt 1977: disease spread by windborne insect vectors? *Vet Rec.* 1982; 11:73–77. [PubMed: 7064329]
- Swanepoel, R.; Coetzer, JA. Rift Valley fever. In: Coetzer, JA.; Tustin, RC., editors. *Infectious Diseases of Livestock.* Oxford Univ. Press; Oxford: 2004. p. 1037-1070.
- Thonnon J, Picquet M, Thiongane Y, Lo M, Sylla R, Vercruysse J. Rift Valley fever surveillance in the lower Senegal river basin: update 10 years after the epidemic. *Trop Med Int Hlth.* 1999; 4:580–585.
- Turell MJ, Presley SM, Gad AM, Cope SE, Dohm DJ, Morrill JC, Arthur RR. Vector competence of Egyptian mosquitoes for Rift Valley fever virus. *Am J Trop Med Hyg.* 1996; 54:136–139. [PubMed: 8619436]

Figure 1.
Location of sampling stations upstream and downstream of Aswan High Dam. Average annual precipitation is shown in mm/y.

Figure 2.

Model selection by lasso. The penalization parameter ($\lambda=0.00729$) was chosen to minimize the average deviance of the withheld observation in leave-one-out cross-validation. Red points show the average error. Error bars are standard deviations across cross-validation folds. Dashed vertical lines show (A) the value of λ at which average cross-validation deviance is minimized and (B) the largest value of λ at which the average cross-validation deviance is within one standard error of the minimum.

Figure 3.

Rift Valley fever epidemics, lagged transmission activity and mosquito abundance, and rainfall at Gambeila. To illustrate the data used in this study, monthly rainfall at Gambeila is plotted over the study period. To show how this relates to RVF transmission activity, epidemic periods are colored red and non-epidemic periods are colored grey. Points misclassified by the fit model are circled, showing poor ability to anticipate the origination of a new outbreak, but satisfactory performance predicting duration (three out of five epidemics terminated at the predicted time). The most important predictor identified by this model was the joint realization of RVF transmission activity at lag 1 and mosquito presence at the current time ($RVF-Lag1 \times Mosquito-Lag0$). Intervals during which this categorical variable is true are signified by dark bars across the top of the plot.

Table 1

Summary of outbreaks of Rift Valley fever in Egypt, 1977-2011.

Outbreak year	Outbreak duration (months)	Human cases	Primary Epidemiological References	Vaccination ¹	Other African nations reporting RVF in the same year ⁵
1977-1978 ²	24	~200,000	Meegan (1979) Darwish & Hoogstraal (1981)	None	Zimbabwe ⁶
1993-1994	20 ³	600-1,500	Arthur et al. (1993)	Live attenuated (Smithburn strain)	Zimbabwe Senegal
1997	na	None ⁴	Abd el-Rahim (1997)	Live attenuated (Smithburn strain)	Somalia Kenya Somalia Tanzania
2003	5	375	Okda et al. (2006) Hanafi et al. (2011)	Live attenuated (Smithburn strain)	Mauritania

¹ Source: Kamal (2011)

² The first outbreak occurred from July-December 1977. A second outbreak occurred from July-December 1978. Phylogenetic evidence suggests that these two outbreaks represented a single introduction event (Grobelaar et al. 2011).

³ Two peaks of transmission to humans reported (see Figure 3). Outbreak duration calculated as time elapsed between first and final months of transmission activity.

⁴ Outbreak confined to domestic ruminants.

⁵ Sources: Swanepoel & Coetzer (1994), Fontenille et al. (1998), Thonnon et al. (1999), Dar et al. (2013), Gerdes (2004)

⁶ RVF was also reported from Sudan in 1976, the year prior to the Egypt outbreak.

Table 2

Variables included in the best fit model in declining order of importance. No other covariates were detected to have any effect. The effect of the regularization parameter on the number of selected coefficients is shown in Figure 2. The theory of glmnet does not provide a method for calculating standard errors on the coefficients.

Variable	Coefficient
<i>RVF-Lag1</i> × <i>Mosquito-Lag0</i>	6.142
<i>Monthly rainfall at Gambeila</i>	0.708
<i>RVF-Lag1</i> × <i>Mosquito-Lag1</i>	0.544
<i>Monthly river discharge at Gaafra</i>	0.073