Disease mortality in domesticated animals is predicted by host evolutionary relationships
Supplementary Information

Maxwell J. Farrell[1]*& T. Jonathan Davies[2,3]

[1]Department of Biology, McGill University
[2]Botany, Forest & Conservation Sciences, University of British Columbia
[3]African Centre for DNA Barcoding, University of Johannesburg, South Africa
*To whom correspondence should be addressed; E-mail: maxwell.farrell@mail.mcgill.ca
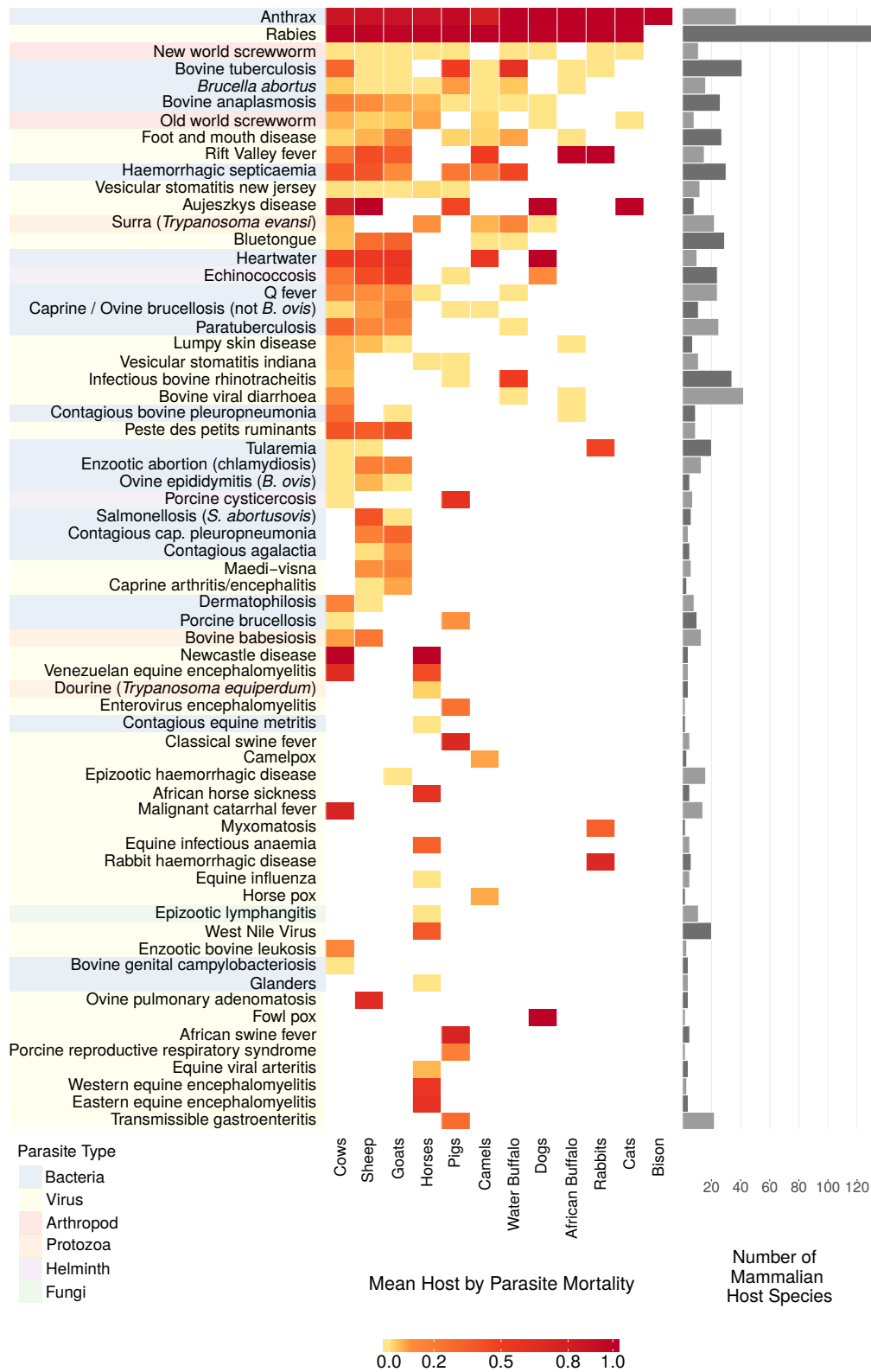
Extended Data Figure

Figure SI 1: (A) Heatmap of mean host by parasite mortality derived from the OIE World Animal Health yearly reports from 2005 to 2011 [1, 2, 3, 4, 5, 6, 7]. Parasite names are represented by OIE disease names, and are colour coded by parasite type. (B) Bar plot of the number of documented mammal host species per parasite derived from the GMPD 2.0 [8] and the Enhanced Infectious Disease Database (EID2) [9]. The order of parasites matches the column order in A.

1 Materials & Methods

1.1 Case-fatality reports

Reports of number of cases and deaths due to infection were taken from published OIE year end reports for the years 2005-2011 [1, 2, 3, 4, 5, 6, 7]. Reported by individual countries, these include information per disease-host combination on the number of cases (infected individuals), deaths due to infection, individuals destroyed, and individuals slaughtered. We included only reports of diseases in mammal hosts. We excluded any observations in which host individuals were reported as destroyed or slaughtered as this would interfere with estimates of deaths due to infection. We also excluded the few instances where the reported number of deaths due to infection exceeded the number of reported cases.

1.2 Host and parasite Latin binomials

Reported host codes were assigned a Latin binomial based on a combination of geographic location, OIE reports, and classifications defined by Clutton-Brock[10] (Table SI 1). Reports that included OIE host codes "cer" (cervidae) and "o/c" (sheep or goats) could not be attributed to a single host species and were excluded.

OIE code	Location	Binomial
bov	Global	*Bos taurus*
buf	Sub-saharan Africa	*Syncerus caffer*
buf	North America	*Bison bison*
buf	Europe, Latin America, Asia, Caribbean, North Africa	*Bubalus bubalis*
can	Global	*Canis lupus*
cap	Global	*Capra hircus*
cml	Global	*Camelus dromedarius*
equ	Global	*Equus caballus*
fel	Global	*Felis silvestris*
lep	Global	*Oryctolagus cuniculus*
ovi	Global	*Ovis aries*
sui	Global	*Sus scrofa*

Table SI 1: Conversion table for OIE host codes to Latin binomials.

Reported disease names were assigned a parasite Latin binomial based on OIE publications (disease summaries from the OIE Terrestrial Manual [11] and OIE technical disease cards) (conversions listed in table SI 2). For diseases caused by a particular subspecies or strain, this subtype was kept in cases where susceptible host species was available (Equine Influenza being largely caused by strain H3N8, and Paratuberculosis caused by *Mycobacterium avium paratuberculosis*). Although there can be large variation in virulence and host specificity among viral subtypes, the OIE does not include information on subtypes in their annual disease reports. Therefore, for all other viruses Latin binomials were standardized to the 2015 version of the International Convention on Viral Taxonomy [12]. Diseases attributed to multiple species were removed (Atrophic rhinitis of swine, Equine piroplasmosis, Equine rhinopneumonitis, Horse mange, Leishmaniosis, Leptospirosis, Sheep and goat pox, Theileriosis, Trichinellosis, and Trypanosomosis), unless the likely causative species could be identified based on geography and/or reported host species (Bovine babesiosis in Europe caused by *Babesia divergens*, Malignant catarrhal fever in sheep worldwide largely caused by *Macavirus ovine herpesvirus 2*, and Malignant catarrhal fever in African cattle caused by *Macavirus alcelaphine herpesvirus 1*). Diseases caused by prions (Scrapie, Bovine Spongiform Encephalopathy) were excluded.

1.3 Host specificity

The suite of mammalian host species infected by each parasite was taken from the Global Mammal Parasite Database 2.0 [8] and a static version of the Enhanced Infectious Disease Database (EID2) database [9]. Host species for *Influenza A H3N8* and *Mycobacterium avium paratuberculosis* are not included in the static version

of the EID2 database, so were instead taken from EID2 online (`eid2.liverpool.ac.uk`) on June 14th 2017. We also included the host species reported as infected by each parasite in the OIE report data used in the analysis. Host Latin binomials were standardized to the 2005 Wilson & Reeder taxonomy [13] using Wilson & Reeder online (`www.departments.bucknell.edu/biology/resources/msw3`) and the Wilson & Reeder 1993-2005 binomial synonym table included in PanTHERIA [14]. Hosts reported to subspecies were collapsed to the parent binomial, and hosts not reported to species level were removed. *Homo sapiens* were excluded. Host species richness was then calculated as the number of unique host Latin binomials associated with each parasite. For each combination of host and parasite reported in the OIE data, mean phylogenetic distances from all known hosts to the infected OIE host was calculated using the Fritz et al. mammal supertree [15] and the R package ape version 3.4 [16].

1.4 Parasite traits

Transmission mode is often listed as a key factor linked to virulence [17, 18, 19, 20]. Here we include whether a parasite is transmitted by an arthropod vector, is transmitted as a function of reproduction (either vertically transmitted, sexually transmitted, or passed from mother to offspring via ingestion of milk or colostrum), and whether it has a resting stage capable of persisting for long periods of time in the environment (typically months to years). Binary parasite traits coding primary modes of transmission and the use of avian species as reservoir hosts were taken from OIE publications (disease summaries from the OIE Terrestial Manual [11] and OIE technical disease cards), and from Lèfevre et al. [21]. Parasite-level effects were modelled as a function of these covariates plus hierarchical effects of parasite type (arthropod, bacteria, fungi, helminth, protozoa, and virus), to account for phylogenetic non-independence and capture additional parasite traits not measured directly. To further evaluate model sensitivity to phylogenetic non-independence among parasites, we construct an alternative model with parasite taxonomic family in place of parasite type (See SI Appendix 2.2.4), and show that the estimates of continuous and binary predictors remain qualitatively unchanged (in direction and relative magnitude) to our main model (Table SI 7).

1.5 Country-level covariates

Host mortality is also likely influenced by local environmental conditions. In our data, these may include socioeconomic factors such as the ability of local peoples to maintain animal health, effects of ambient temperature on parasite growth rate, or co-infection with other parasites. While the scale of reporting does not allow us to investigate these factors directly, we include two country-level predictors: 1) per capita Gross Domestic Product (GDP) to model economic abilities to reduce host mortality, and 2) latitude as a proxy for temperature and biodiversity gradients that may reflect environmental conditions determining the strength of species interactions [22], in addition to modelling country-level variation. To include country-level covariates from the World Bank World Development Indicators API, we standardized country names to those used in the WDI R package version 2.4 [23]. For each country we extracted mid-country latitude and per capita GDP in current US dollars (WDI code "NY.GDP.PCAP.CD") using the WDI package. Countries that did not have reported GPD per capita from the WDI were supplemented with information from the United Nations Data Retreival System (`data.un.org`) so that there was at least one estimate of per capita GDP for the period of 2005-2011. Mean gross domestic product per capita per country was then calculated across all years. We excluded records from countries with no iso3 code or for which no latitude was reported.

1.6 Model

Using a hierarchical Bayesian binomial-logit model, we model deaths ($deaths_i$) as following a binomial distribution determined by sample size per observation ($cases_i$) and a probability parameter p_i. The higher-level structure of the model is as follows:

$$deaths_i \sim Bin(cases_i, p_i) \tag{1}$$

Where p_i is modeled with β_0 as the grand mean plus the effects of mean phylogenetic distance from all known hosts to the species infected ($EvoIso_i$), the number of cases per observation ($cases_i$), and partially-pooled hierarchical effects for parasites (μ_{para}), hosts (μ_{host}), countries ($\mu_{country}$), and years (μ_{year}):

$$logit(p_i) = \beta_0 + \beta_1 * EvoIso_i + \beta_2 * log(cases_i) + \mu_{para} + \mu_{host} + \mu_{country} + \mu_{year} \tag{2}$$

Parasite level effects, μ_{para}, are defined by a normal distribution as follows:

$$\mu_{para} \sim \mathcal{N}(\beta_3 * SR_{para} + \beta_4 * vect_{para} + \beta_5 * repro_{para} + \beta_6 * envRest_{para} + \mu_{type}, \sigma_P^2) \tag{3}$$

Where the difference from the grand mean (β_0) for each parasite ($para$) is determined by host species richness (SR_{para}), transmission modes ($vect_{para}$, $repro_{para}$, $envRes_{para}$), and a hierarchical effect of the parasite type (μ_{type}), and variance parameter (σ_P^2).

Parasite taxonomic type, μ_{type}, is modelled following a normal distribution with mean of zero and variance parameter (σ_T^2) as follows:

$$\mu_{type} \sim \mathcal{N}(0, \sigma_T^2) \tag{4}$$

Host level effects, μ_{host}, are modelled following a normal distribution with mean determined by a hierarchical effect of the host taxonomic order (μ_{order}) and variance parameter (σ_H^2) as follows:

$$\mu_{host} \sim \mathcal{N}(\mu_{order}, \sigma_H^2) \tag{5}$$

Host taxonomic order level effects, μ_{order}, are modelled following a normal distribution with mean of zero and variance parameter (σ_O^2) as follows:

$$\mu_{order} \sim \mathcal{N}(0, \sigma_O^2) \tag{6}$$

Country level effects, $\mu_{country}$, are modelled following a normal distribution with mean determined by gross domestic product per capita (GDP_c) and latitude ($latitude_c$), and variance parameter (σ_C^2) as follows:

$$\mu_{country} \sim \mathcal{N}(\beta_7 * GDP_c + \beta_8 * latitude_c, \sigma_C^2) \tag{7}$$

Year level effects, μ_{year}, are modelled following a normal distribution with mean of zero and variance parameter (σ_Y^2) as follows:

$$\mu_{year} \sim \mathcal{N}(0, \sigma_Y^2) \tag{8}$$

1.7 Priors & Data transformations

Following the recommendations of Gelman et al. [24], continuous predictors were normalized to mean of zero and standard deviation of 0.5. Estimated parameters were modelled using weakly informative priors as recommended by [25] and the Stan development team:

$$\beta_{0-8} \sim Student\ t(4, 0, 1) \tag{9}$$

$$\sigma_{P,H,O,C,Y}^2 \sim Half\ Student\ t(4, 0, 1) \tag{10}$$

1.8 Sampling and Convergence Diagnostics

Models were fit in Stan [26, 27] via R 3.2.3 [28] with rstan version 2.14.2 [29] using 4 chains with 30,000 iterations per chain. The first 15,000 iterations per chain were used for warm-up and discarded. The remaining posterior was thinned to retain every 10th iteration, resulting in a total of 6,000 posterior draws. Convergence was diagnosed by observation of Rhat values equal to 1 (see Table SI 3) and explored with shinystan version 2.4.0 [30]. Posterior predictive checks were performed to ensure model validity and fit to the data. The main model was also fit with simulated data (sim_data.R) to ensure the model performs as expected and is able to recover simulated parameters.

Parasite Latin binomials

Table SI 2: Conversion table for OIE diseases to Latin binomials.

OIE disease name	Latin binomial
African horse sickness	*Orbivirus african horse sickness virus*
African swine fever	*Asfivirus african swine fever virus*
Anthrax	*Bacillus anthracis*
Aujeszkys disease	*Varicellovirus suid herpesvirus 1*
Bluetongue	*Orbivirus bluetongue virus*
Bovine anaplasmosis	*Anaplasma marginale*
Bovine babesiosis	
Bovine babesiosis (in Europe)	*Babesia divergens*
Bovine genital campylobacteriosis	*Campylobacter fetus*
Bovine tuberculosis	*Mycobacterium bovis*
Bovine viral diarrhoea	*Pestivirus bovine viral diarrhea virus*
Brucella abortus	*Brucella abortus*
Brucella melitensis	*Brucella melitensis*
Brucella suis	*Brucella suis*
Camelpox	*Orthopoxvirus camelpox virus*
Cap/ovi brucel. (not B. ovis)	*Brucella melitensis*
Caprine arthritis/encephalitis	*Lentivirus caprine arthriti*s encephalitis virus
Classical swine fever	*Pestivirus classical swine fever virus*
Contagious agalactia	*Mycoplasma agalactiae*
Contagious bov. pleuropneumonia	*Mycoplasma mycoides*
Contagious cap. pleuropneumonia	*Mycoplasma capricolum*
Contagious equine metritis	*Taylorella equigenitalis*
Dermatophilosis	*Dermatophilus congolensis*
Dourine	*Trypanosoma equiperdum*
Echinococcosis	*Echinococcus granulosus*
Enterovirus encephalomyelitis	*Teschovirus porcine teschovirus*
Enzootic abortion (chlamydiosis)	*Chlamydophila abortus*
Enzootic bovine leukosis	*Deltaretrovirus bovine leukemia virus*
Epizootic haemorrhagic disease	*Orbivirus epizootic hemorrhag*ic disease virus
Epizootic lymphangitis	*Histoplasma capsulatum*
Equine encephalomyelitis Eastern	*Alphavirus eastern equi*ne encephalitis virus
Equine encephalomyelitis Western	*Alphavirus western equi*ne encephalitis virus
Equine infectious anaemia	*Lentivirus equine infectious anemia virus*
Equine influenza	*Influenzavirus A H3N8*
Equine viral arteritis	*Arterivirus equine arteritis virus*
Foot and mouth disease	*Aphthovirus foot-and-mouth disease virus*
Fowl cholera	*Pasteurella multocida*
Fowl pox	*Avipoxvirus fowlpox virus*
Glanders	*Burkholderia mallei*
Haemorrhagic septicaemia	*Pasteurella multocida*
Heartwater	*Ehrlichia ruminantium*
Horse pox	*horsepox virus*
Inf.bov.rhinotracheit. (IBR/IPV)	*Varicellovirus bovine herpesvirus 1*
Lumpy skin disease	*Capripoxvirus lumpy skin disease virus*
Maedi-visna	*Lentivirus visna/maedi virus*
Malignant catarrhal fever	
Malignant catarrhal fever (Cattle in Africa)	*Macavirus alcelaphine herpesvirus 1*
Malignant catarrhal fever (Sheep Worldwide)	*Macavirus ovine herpesvirus 2*
Myxomatosis	*Leporipoxvirus myxoma virus*

N. w. screwworm (C. hominivorax)	*Cochliomyia hominivorax*
Newcastle disease	*Avulavirus newcastle disease virus*
O. w. screwworm (C. bezziana)	*Chrysomya bezziana*
Ovine epididymitis (B. ovis)	*Brucella ovis*
Ovine pulmonary adenomatosis	*Betaretrovirus jaagsiekte sheep retrovirus*
Paratuberculosis	*Mycobacterium avium paratuberculosis*
Peste des petits ruminants	*Morbillivirus peste-des-petits-ruminants virus*
Porcine brucellosis	*Brucella suis*
Porcine cysticercosis	*Taenia solium*
Porcine reproductive respiratory syndrome	*Arterivirus porcine reproductive and respiratory syndrome virus*
Q fever	*Coxiella burnetii*
Rabbit haemorrhagic disease	*Lagovirus rabbit hemorrhagic disease virus*
Rabies	*Lyssavirus rabies virus*
Rift Valley fever	*Phlebovirus rift valley fever virus*
Salmonellosis (S. abortusovis)	*Salmonella abortusovis*
Surra (Trypanosoma evansi)	*Trypanosoma evansi*
Transmissible gastroenteritis	*Alphacoronavirus alphacoronavirus 1*
Tularemia	*Francisella tularensis*
Venezuelan equ.encephalomyelitis	*Alphavirus venezuelan equine encephalitis virus*
Vesicular stomatitis indiana	*Vesiculovirus vesicular stomatitis indiana virus*
Vesicular stomatitis new jersey	*Vesiculovirus vesicular stomatitis new jersey virus*
West Nile Fever	*Flavivirus west nile virus*

2 Model Results

2.1 Main model

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.57	0.74	-2.06	-1.03	-0.57	-0.08	0.87	1338	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.34	-1.33	-1.32	-1.30	6000	1.00
	Evolutionary Isolation	1.69	0.12	1.45	1.61	1.70	1.78	1.93	6000	1.00
	Host Species Richness	0.90	0.71	-0.42	0.41	0.88	1.37	2.36	3335	1.00
Parasite	Vectored	-0.27	0.65	-1.60	-0.69	-0.26	0.15	1.00	3631	1.00
	Reproduction	0.12	0.66	-1.19	-0.32	0.13	0.56	1.43	4317	1.00
	Environmental Resting Stage	0.22	0.96	-1.65	-0.40	0.20	0.83	2.19	6000	1.00
Country	Latitude	-0.21	0.42	-1.04	-0.49	-0.21	0.08	0.61	3796	1.00
	GDP per capita	-0.56	0.37	-1.29	-0.81	-0.56	-0.31	0.18	6000	1.00

Table SI 3: Summary of model output for continuous predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).
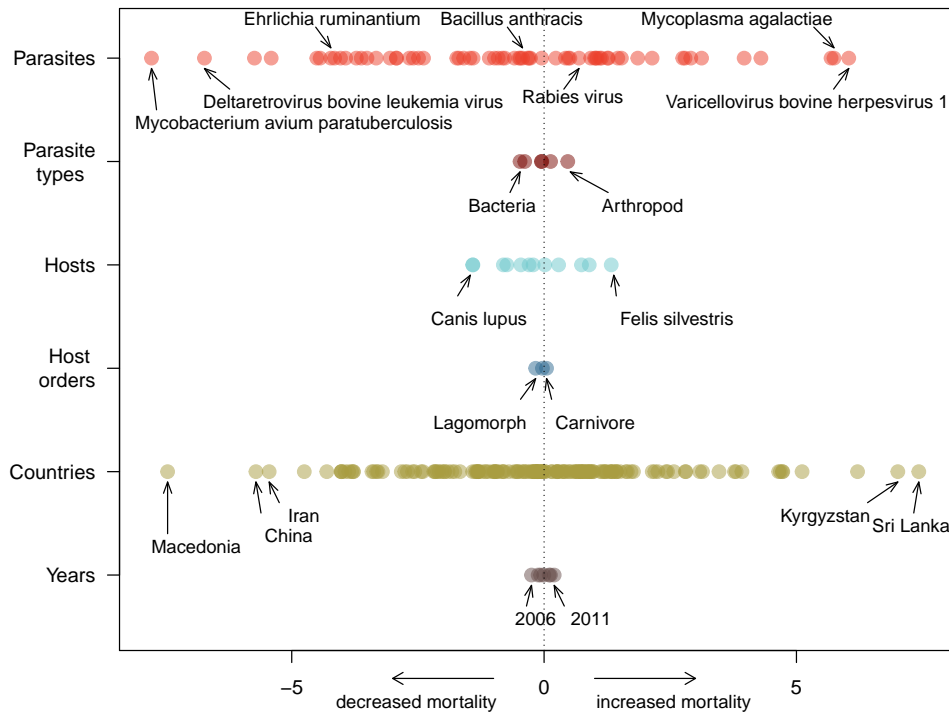

Figure SI 2: Mean estimated effects for individual hierarchical terms (parasites, parasite types, hosts, host orders, countries, and years). These effects include the level-specific predictors and the estimated variation among factor levels (as described in SM 1.6). The hierarchical effects in our model represent offsets relative to the rest of the model fit. For example, the model aside from the parasite level captures the mean mortality of *Rabies virus* well, but *Varicellovirus bovine herpesvirus 1* is more deadly than we might expect given the reported case numbers, host evolutionary isolation, particular host species infected, and countries and years from which the case-fatality data was reported. Plotted estimates have been set to 50% transparency to visualize overlapping points, and examples of levels from each group have been identified.
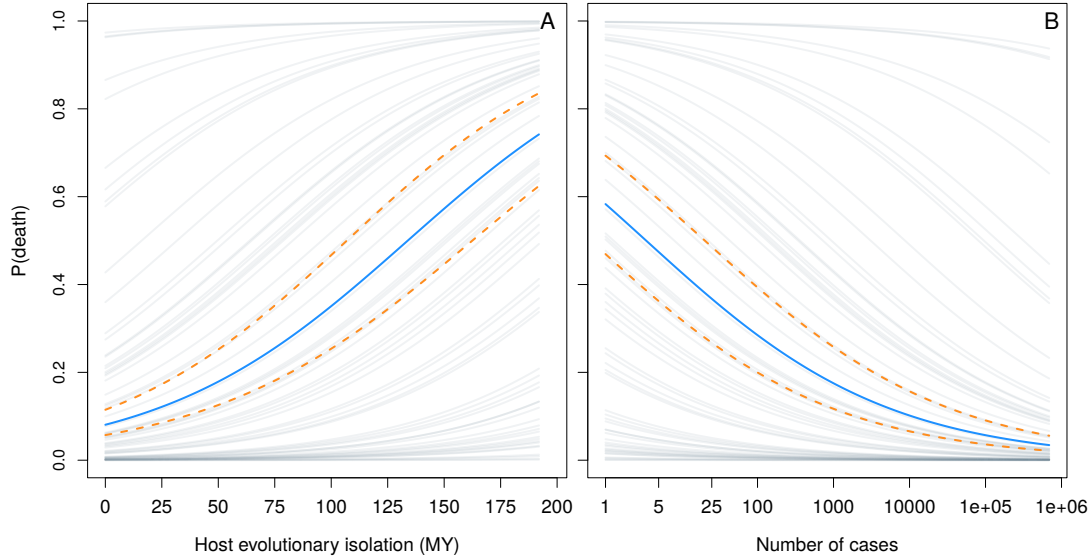
Figure SI 3: Posterior predictions of the probability of death as a function of A) host evolutionary isolation (in millions of years), and B) the number of cases. Solid blue lines represent the mean logistic curve, dashed yellow lines represent the upper and lower bounds of the 50% credible interval. Grey lines depict equivalent mean curves offset by the posterior mean effects for each parasite.

2.2 Sensitivity Analyses and Alternative Models

2.2.1 Excluding single-host parasites

As selective pressures driving virulence evolution are likely to differ among single and multi-host parasites, we re-fit the main model again after removing single-host parasites from the data.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.83	0.80	-2.50	-1.32	-0.81	-0.30	0.70	1075	1.00
	log (Cases)	-1.30	0.01	-1.33	-1.31	-1.30	-1.29	-1.28	6000	1.00
	Evolutionary Isolation	1.74	0.11	1.52	1.66	1.74	1.81	1.96	5847	1.00
	Host Species Richness	0.55	0.75	-0.86	0.06	0.53	1.04	2.09	3668	1.00
Parasite	Vectored	-1.05	0.76	-2.61	-1.53	-1.02	-0.54	0.36	2760	1.00
	Reproduction	0.47	0.71	-0.87	-0.02	0.44	0.92	1.92	3564	1.00
	Environmental Resting Stage	0.26	0.96	-1.60	-0.34	0.24	0.83	2.26	6000	1.00
Country	Latitude	-0.25	0.42	-1.08	-0.52	-0.24	0.04	0.57	4573	1.00
	GDP per capita	-0.58	0.37	-1.32	-0.83	-0.58	-0.33	0.13	5092	1.00

Table SI 4: Summary of main model excluding single-host parasites for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

2.2.2 Host taxonomic diversity

Due to incomplete sampling, the host species reported in the GMPD and EID2 databases are unlikely to include the complete set of susceptible hosts for each parasite. As a sensitivity analysis, host species richness (SR_p) was replaced by a measure of taxonomic diversity using data reported by Lèfevre et al. [21] and the OIE documentation. Host taxonomic diversity varies from 1-6 corresponding to whether parasites infect hosts belonging to a single species (1), genus (2), family (3), order (4), class (5), or multiple classes (6). Just as with host species richness, the ability to infect humans was not included in estimates of taxonomic diversity.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.79	0.90	-2.71	-1.37	-0.75	-0.19	0.87	2261	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.34	-1.33	-1.32	-1.30	6000	1.00
	Evolutionary Isolation	1.70	0.13	1.45	1.61	1.70	1.78	1.93	5833	1.00
	Host Taxonomic Diversity	-0.01	0.20	-0.39	-0.15	-0.02	0.12	0.38	2687	1.00
	Vectored	-0.26	0.66	-1.59	-0.69	-0.26	0.18	1.02	5847	1.00
Parasite	Reproduction	0.04	0.67	-1.28	-0.40	0.05	0.48	1.36	5347	1.00
	Environmental Resting Stage	0.30	0.98	-1.56	-0.34	0.27	0.90	2.39	6000	1.00
Country	Latitude	-0.20	0.43	-1.05	-0.48	-0.21	0.08	0.65	5261	1.00
	GDP per capita	-0.55	0.37	-1.27	-0.80	-0.55	-0.30	0.18	6000	1.00

Table SI 5: Summary of model with host taxonomic diversity for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

2.2.3 Parasites with avian reservoirs

As an extension of our main model, we include whether or not a parasite uses an avian reservoir (Eastern equine encephalitis, Western equine encephalitis, Venezuelan equine encephalitis, Fowlpox, Newcastle Disease, West Nile Virus, *Pasturella multocida*), as we hypothesize that this might correlate with whether domesticated mammals represent dead-end hosts from which the parasite is not transmitted further, such as is the case for West Nile Virus and other encephalitic viruses that spillover from birds to horses [31]. The use of avian species as reservoir hosts were taken from OIE publications (disease summaries from the OIE Terrestial Manual [11] and OIE technical disease cards), and from Lèfevre et al. [21], and coded as a binary predictor.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.59	0.72	-2.04	-1.07	-0.59	-0.11	0.80	1568	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.33	-1.33	-1.32	-1.30	5980	1.00
	Evolutionary Isolation	1.69	0.12	1.46	1.61	1.69	1.78	1.93	5935	1.00
	Host Species Richness	0.90	0.73	-0.45	0.41	0.87	1.36	2.42	3574	1.00
	Vectored	0.28	0.65	-1.60	-0.70	-0.26	0.16	1.01	4318	1.00
Parasite	Reproduction	0.14	0.65	-1.14	-0.30	0.13	0.58	1.42	5209	1.00
	Environmental Resting Stage	0.23	0.97	-1.61	-0.39	0.20	0.82	2.29	5934	1.00
	Avian Reservoir	0.18	0.83	-1.47	-0.36	0.17	0.68	1.83	5474	1.00
Country	Latitude	-0.20	0.42	-1.05	-0.48	-0.20	0.09	0.62	3896	1.00
	GDP per capita	-0.55	0.37	-1.29	-0.80	-0.55	-0.31	0.15	5942	1.00

Table SI 6: Summary of model including indicator for avian reservoir for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

2.2.4 Model with parasite taxonomic family

In an effort to explore the potential influence of phylogenetic non-inependence among parasite taxa, we provide an extension of our model with parasite type replaced by parasite taxonomic information at the level of family.

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.68	0.67	-2.01	-1.11	-0.67	-0.24	0.66	1980	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.34	-1.33	-1.32	-1.30	5760	1.00
	Evolutionary Isolation	1.69	0.12	1.45	1.61	1.69	1.78	1.94	5767	1.00
	Host Species Richness	0.88	0.71	-0.43	0.40	0.86	1.35	2.35	2924	1.00
Parasite	Vectored	-0.25	0.65	-1.55	-0.68	-0.25	0.19	1.01	3324	1.00
	Reproduction	0.07	0.66	-1.23	-0.36	0.07	0.49	1.37	3724	1.00
	Environmental Resting Stage	0.18	0.95	-1.65	-0.44	0.15	0.77	2.12	5085	1.00
Country	Latitude	-0.21	0.43	-1.04	-0.50	-0.21	0.08	0.64	4510	1.00
	GDP per capita	-0.56	0.37	-1.29	-0.80	-0.56	-0.31	0.17	5952	1.00

Table SI 7: Summary of model with parasite taxonomic family for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

2.2.5 Model with citation count per parasite

In an effort to explore the potential influence of study effort on parasite-induced mortality, we provide an extension of our model with the number of Web of Knowledge citations per parasite included as another parasite-level predictor. Web of Knowledge searches were conducted on January 29, 2019 from the University of Georgia. Searching within "All databases", we recorded the number of citations returned with either the disease common name or parasite Latin binomial in either the "Tile" or "Topic" (ex. *"TI=("African horse sickness" OR "Orbivirus african horse sickness virus") OR TS=("African horse sickness" OR "Orbivirus african horse sickness virus")"*). Where two diseases were caused by a single parasite species (ex. Fowl cholera and Haemorrhagic septicaemia caused by *Pasteurella multocida*) we searched only for the parasite Latin binomial. To account for potential increased study effort for zoonotic diseases, we conducted the same searches with the added search term *"NOT TI=(human* OR "Homo sapiens") NOT TS=(human* OR "Homo sapiens")"*, which excludes the terms "human*" and "Homo sapiens" from either the topic or title. Citation counts excluding humans were always lesser or equal to those including humans, however the two counts are highly correlated (Pearson's $r = 0.99$). We conduct our model using citation counts excluding humans, but provide both predictors and the exact Web of Knowledge search terms in the supplemental data ("citations.csv").

Level	Parameter	mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
	Intercept	-0.58	0.75	-2.07	-1.06	-0.58	-0.08	0.89	1890	1.00
	log (Cases)	-1.33	0.01	-1.35	-1.34	-1.33	-1.32	-1.30	5623	1.00
	Evolutionary Isolation	1.69	0.12	1.46	1.61	1.69	1.77	1.94	5837	1.00
	Host Species Richness	0.97	0.81	-0.48	0.43	0.92	1.47	2.67	5004	1.00
	Vectored	-0.27	0.65	-1.55	-0.70	-0.25	0.17	0.99	4383	1.00
Parasite	Reproduction	0.17	0.68	-1.18	-0.28	0.16	0.60	1.50	5203	1.00
	Environmental Resting Stage	0.22	0.96	-1.60	-0.40	0.20	0.80	2.27	5782	1.00
	Citation Count	-0.11	0.66	-1.44	-0.53	-0.10	0.32	1.19	5679	1.00
Country	Latitude	-0.20	0.42	-1.05	-0.48	-0.20	0.08	0.61	3568	1.00
	GDP per capita	-0.55	0.37	-1.29	-0.81	-0.56	-0.30	0.17	5806	1.00

Table SI 8: Summary of model with citation counts per parasite, for continuous and binary predictors including posterior means, posterior standard deviations, 2.5%, 25%, 50%, 75% and 97.5% quantiles, the effective sample size (n_eff), and the potential scale reduction statistic (Rhat).

References

[1] *World Animal Health in 2005* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-664-4, 2007).

[2] *World Animal Health in 2006* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-667-5, 2008).

[3] *World Animal Health in 2007* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-670-5, 2008).

[4] *World Animal Health in 2008 (Volumes 1 & 2)* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-792-4; 978-92-9044-793-1, 2009).

[5] *World Animal Health in 2009 (Volumes 1 & 2)* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-798-6; 978-92-9044-799-3, 2010).

[6] *World Animal Health in 2010 (Volumes 1 & 2)* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-810-5; 978-92-6044-811-2, 2011).

[7] *World Animal Health in 2011 (Volumes 1 & 2)* (World Organisation for Animal Health (OIE), Paris, France. ISBN: 978-92-9044-867-9; 978-92-9044-868-6, 2012).

[8] Stephens, P. R. *et al.* Global Mammal Parasite Database version 2.0. *Ecology* 42–49 (2017).

[9] Wardeh, M., Risley, C., McIntyre, M. K., Setzkorn, C. & Baylis, M. Database of host-pathogen and related species interactions, and their global distribution. *Scientific Data* **2**, 150049 (2015).

[10] Clutton-Brock, J. *A Natural History of Domesticated Mammals* (Cambridge University Press, Cambridge, United Kingdom, 1999), 2nd edn.

[11] World Organisation for Animal Health (OIE). *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals* (2012), 7th edition edn.

[12] Lefkowitz, E. J. *et al.* Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research* **46**, D708–D717 (2018).

[13] Wilson, D. E. & Reeder, D. M. *Mammal Species of the World: A Taxonomic and Geographic Reference* (Johns Hopkins University Press, Baltimore, Maryland, 2005), 3 edn.

[14] Jones, K., Bielby, J., Cardillo, M. & Fritz, S. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**, 2648 (2009).

[15] Fritz, S. A., Bininda-Emonds, O. R. P. & Purvis, A. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology letters* **12**, 538–549 (2009).

[16] Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).

[17] Alizon, S., Hurford, a., Mideo, N. & Van Baalen, M. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *Journal of evolutionary biology* **22**, 245–59 (2009).

[18] Ewald, P. Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics* **14**, 465–485 (1983).

[19] Rigaud, T., Perrot-Minnot, M.-J. & Brown, M. J. F. Parasite and host assemblages: embracing the reality will improve our knowledge of parasite transmission and virulence. *Proceedings of the Royal Society Biological Sciences* **277**, 3693–702 (2010).

[20] Cressler, C. E., McLeod, D. V., Rozins, C., Van den Hoogen, J. & Day, T. The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology* **143**, 915–930 (2016).

[21] Lefèvre, P.-C., Blancou, J., Chermette, R. & Uilenberg, G. *Infectious and parasitic diseases of livestock* (Lavoisier - Editions Médicales Internationales, Cachan, France, 2010).

[22] Schemske, D. W., Mittelbach, G. G., Cornell, H. V., Sobel, J. M. & Roy, K. Is There a Latitudinal Gradient in the Importance of Biotic Interactions? *Annual Review of Ecology, Evolution, and Systematics* **40**, 245–269 (2009).

[23] Arel-Bundock, V. *WDI: World Development Indicators (World Bank)* (2013). R package version 2.4.

[24] Gelman, A., Jakulin, A., Pittau, M. G. & Su, Y.-S. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* **2**, 1360–1383 (2008).

[25] Ghosh, J., Li, Y. & Mitra, R. On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression. *arXiv preprint* **1507.07170**, 1–59 (2015).

[26] The Stan Development Team. The Stan Core Library, Version 2.17.0 (2017).

[27] Carpenter, B. *et al.* Stan: A probabilistic programming language. *Journal of Statistical Software, Articles* **76**, 1–32 (2017).

[28] R Core Team. R: A Language and Environment for Statistical Computing (2015).

[29] Stan Development Team. RStan: the R interface to Stan (2017). R package version 2.17.2.

[30] Stan Development Team. shinystan: Interactive visual and numerical diagnostics and posterior analysis for bayesian models. (2017). R package version 2.4.0.

[31] Weaver, S. C. & Barrett, A. D. T. Transmission cycles, host range, evolution and emergence of arboviral disease. *Nature reviews. Microbiology* **2**, 789–801 (2004).