Censored regression modelling to predict virus inactivation in wastewaters

Dr. Julii Brainard[1], *Dr. Katherine Pond*[2], *Professor Paul R. Hunter**[1]

[1] Norwich Medical School, University of East Anglia Norwich NR4 7TJ, United Kingdom

[2] Department of Civil and Environmental Engineering, Robens Centre for Public and Environmental Health, University of Surrey, Guildford GU2 7XH

* Professor Paul Hunter, Norwich Medical School, University of East Anglia Norwich NR4 7TJ, United Kingdom. Tel. 01603-456161. Fax. 01603-593752. Paul.Hunter@uea.ac.uk.

KEYWORDS wastewaters, viruses, inactivation, faeces, model, censored regression

ABSTRACT

Among the many uncertainties presented by poorly studied pathogens is possible transmission via human faecal material or wastewaters. Such worries were a documented concern during the 2013 Ebola outbreak in West Africa. Using published experimental data on virus inactivation rates in wastewater and similar matrices, we extracted data to construct a model predicting the T90 (1 x $\log_{10}$ inactivation measured in seconds) of a virus. Extracted data were: RNA or DNA genome, enveloped or not, primary transmission pathway, temperature, pH, light levels and matrix. From the primary details, we further determined matrix level of contamination, genus and taxonomic family. Prior to model construction, three records were separated for verification. A censored normal regression model provided the best fit model, which predicted T90 from DNA or RNA structure, enveloped status, whether primary transmission pathway was faecal-oral, temperature and whether contamination was low, medium or high. Model residuals and predicted values were evaluated against observed values. Mean values of model predictions were compared to independent data, and considering 95% confidence ranges (which could be quite large). A relatively simple model can predict virus inactivation rates from virus and matrix attributes, providing valuable input when formulating risk management strategies for little studied pathogens.

INTRODUCTION

The emergence of new or re-emergence of previously known viral infections is often followed by concerns about the risks of environmental transmission. This potential exposure path was identified in the epidemic of SARS infection [1], for avian influenza [2] and more recently the Ebola epidemic in West Africa [3]. When concerns are raised about the risk of environmental transmission, attention naturally turns to questions of survival and persistence of the implicated virus in the environment. One of the areas of particular interest in in the survival of virus in wastewater and latrine sludge.

In the recent Ebola epidemic in 2014-15, the World Health Organisation (WHO) issued guidance about handling latrine waste contaminated by Ebola virus (EBOV). However, it was acknowledged in August 2014 that relevant scientific data were sparse, and initial guidelines [3] stated that EBOV-contaminated latrines should be kept secure for a minimum of four weeks after last use, with any subsequent desludging to involve wearing full personal protective equipment. However, other authors expressed concerns that posited transmission risks from latrine materials were poorly evidenced in emerging research [4]. In part, this disagreement reflected the lack of data on the survival of EBOV in wastewater [5].

The initial cautionary guidelines on keeping latrine sludge for four weeks proved difficult to maintain and subsequent hazard and critical control analysis [6] as well as hazard assessment and experimental data [5, 7, 8] allowed a reappraisal of the guidance. WHO guidelines about how long to keep an Ebola-contaminated latrine secure and when desludging could commence were correspondingly revised in 2015 [9] to recommend storage for a minimum of seven days after last receipt of infectious material.

Clearly, better knowledge of the environmental survival of viral pathogens early in any future epidemic would aid guidance formulation. However, as with the Ebola epidemic getting this data directly through experimental or observational studies may not be easy. Part of the problem with EBOV was the need to ensure strict safety standards for any experimental work which delayed the start of any such research [5]. This led us to investigate whether or not it was possible to predict viral survival in environmental matrices given a relatively limited amount of data on a particular virus.

The aim of this paper was to collect relevant data useful to explore and quantify [10] the relationship between possible predictive variables and viral persistence in faecally-contaminated matrices. We did this by constructing a model to best predict virus deactivation in surface waters, wastewaters and other matrices which are potentially contaminated with organic matter, especially faecal material.

METHODS

Primary data on virus inactivation in eligible media were collected by extracting data from published experiments and observational findings. Potentially suitable articles were found by searching two bibliographic databases (Pubmed and Scopus) using the below phrase (af = all text field;, tw = in title, abstract or keywords; exp = expanded alternatives). We only selected data from articles in peer-reviewed literature. There were no date or language restrictions.

exp viruses/ (or for scopus *virus).tw.

AND
(stool or feces or faeces or wastewater or manure).af.
AND
(inactivation or survival or removal or persistence or viability).af.

A single reviewer (JB or KP) screened each title and abstract for articles that indicated they contained time-series data about virus inactivation in eligible media. Articles were excluded if they *only* had data for sterile water-based media or tissue culture. Note that data for sterile media *were* included for extraction when reported in an article that also reported data about virus inactivation in contaminated media. This data selection strategy was done purposefully so as to collect some data on sterile media for our modelling, but to not try to exhaustively search and record all such data for sterile media. Full text of each article that could not be excluded from title and abstract was screened to confirm or reject eligibility. In addition, some other articles were known to the authors to have suitable data, and we also checked references in two previously compiled literature inventories for information about virus persistence in faecally-contaminated material, fresh-water, wastewaters or wet tissue culture [11, 12].

Data concerned with virus removal by physical means (eg, filtering), or matrices that were purposefully disinfected by a chemical agent, were ineligible. Inactivation data for matrices exposed to temperatures > 55 degrees C were excluded (because we wanted to exclude infeasible outdoor air temperatures, and did not want to capture data relating to efficacy of sterilization methods). Data were extracted by a single investigator (JB or KP) and verified by another researcher (KP or JB). Data about virus inactivation expressed as T90 (1 x $\log_{10}$ decline) in any faecally-contaminated matrix, water-based media or (wet) cell culture were extracted. Dried media, or media to which disinfection agents had been added, were both excluded. From all eligible articles, the following variables were extracted into standardised forms:

Bibliographic details, virus, temperature of experiment, matrix virus was kept in, inactivation time (T90, in seconds), lighting conditions (that matrix was exposed to during experimental run) and pH. Where a large number of very similar experiments were undertaken (see for example, Magri *et al*. 2015 [13]), which had very similar media, temperature and other conditions, with corresponding similar T90 results, then a grouped average T90 was recorded with median/mean values extracted for predictors (such as temperature, contamination level, pH, etc). This grouping was done to try to prevent a large set of data from a relatively small number of articles (and their specific experimental methods) dominating the model outputs.

Using the data available from primary extraction, and by consulting a large range of sources (Supporting Information, List S1) we also recorded various characteristics of the virus: genetic material (RNA or DNA), enveloped virus or not (a binary 1/0 variable), and primary transmission pathway(s) (airborne, body contact/fluids, faecal-oral, insect vector, respiratory, rodents or multiple). The variable 'faecal-oral' was generated for each record and defined to equal 1 for primary transmission pathway = faecal-oral, and 0 otherwise. Some reports gave experiment temperature as 'room temperature', which was recoded to 20° C. The matrix was also categorized as having a high, medium or low level of faecal contamination according to the logic: media with no faecal or urine content were categorized as low, while wastewaters and media with unclear faecal content or ≥ 10% faecal material were categorized as high. All other matrices were categorized as medium level of contamination, except when diluted to ≤ 1%, causing the contamination category to move down one level (ie, faecally-contaminated

wastewater diluted to 1/1000 moved from high to medium). Light conditions that matrices were exposed to were recorded (eg., dark, solar UV, etc). Matrix pH during the monitoring period from start until final time point or T90 was reached, was also extracted. A variable, pHdiff7, was generated which was the absolute difference in pH from 7.0 (ie. pHdiff7 = abs(pH – 7)).

Inactivation times (T90s, time in seconds to decline 90% or $\log_{10}$ 1) were often stated precisely (usually in tables), but sometimes only available to read on graphic figures or in supplemental data. Incomplete and imprecise data were common, often due to a finite monitoring period. Hence T90s were sometimes recoded as follows: <5% apparent decline during the full observation period meant the record was excluded (insufficient information). Decline of 50%-89% of peak value at the last time-monitoring point, the last time point was recorded as T90, and as a censored value (relevant to regression modelling, see below). If last observed viral load was 26%-49% of peak viral load, T90 was recoded as 1.5 x last time point (right censored). Where observed viral load at last monitoring point was less than peak but below 26% of peak viral load value, T90 was coded as 2 x last observation time (also right censored). If viral load had fallen below limit of detection at first time period, the first time point was taken as T90 and the data noted as left censored.

After eligibility screening but prior to model construction, three records in three studies [8, 14, 15] were separated to provide independent data to test the final model against. These three articles were chosen because they were relatively recently published (2013-15), included both DNA and RNA viruses, provided a diversity of primary transmission pathways (faecal-oral, body fluids and respiratory), three different levels of faecal contamination (low, medium, high) and three different genera.

The extracted data were input to a regression model within Stata (v.14.0, cnreg command [16]) to predict virus inactivation (logarithm with base 10, of T90 expressed in seconds) as a function of available attributes of either or both virus and matrix. Many transformations of predictor and response variables were tried (square roots, logarithmic, exponential, etc). The primary aim of the model was to best predict T90 from the available data. The preferred model utilized easily obtainable virus and media attributes while minimizing overall uncertainty, as indicated by the robust standard error of the residuals [17, 18]. The robust standard error was determined using a clustered sandwich estimator for the standard deviation of model residuals (vce option in Stata [19], clustering by genus). Other desirable model features were statistically significant p-values ($\leq$0.05) for variable coefficients, and credible relationships between predictors and dependent variable. For the preferred model, fit and reliability were explored by comparing residuals to fitted values, comparing fitted with predicted values and by comparing model predictions with independent data not used in model construction.

RESULTS AND DISCUSSION

The data search and study selection is described in Figure 1. Article searches were undertaken on 18 May 2016. From 2088 partly duplicated articles found in the primary search, there were 619 studies that could not be excluded after screening title and abstract. Of those, 583 articles were excluded after full text review. A further 19 articles were known by the authors to have relevant data, or were found by reading other literature inventories. A final total of 55 articles (containing 467 data points representing 52 unique viruses) were

found that had observation data points suitable for input to or testing of our regression model(s). Three papers [8, 14, 15] each containing one record suitable for model testing were removed from the model construction data set (as described in Methods). The final number of observations used in model construction was 464, from data in 52 papers about 51 unique viruses.

In the models discussed below, the square root of the Logarithmic (base 10) transformation of the T90 (expressed in seconds) value was the dependent variable. T90 expressed in seconds allowed for observations taken within one minute of virus inoculation into a matrix. The logarithmic and square root transformations led to minimal robust standard error for the residuals. Censored linear regression was appropriate due to the observation limits of the dependent variable [16, 20]; the dependent variable (T90) was sometimes only recorded as below or above a specified detection limit (see Methods).

Light and pH. Only about 25% of records had data on matrix exposure to light; we judged this insufficient for our study purposes and light levels were thus disregarded in the modelling. There were also several problems with using pH as a predictor of T90. About 30% (141 of 464) of records did not provide information on matrix pH during the T90 observation period. Observed values of pH were relatively limited compared to possible real world conditions (extracted pH values = 2.1-2.6, and 6-9.3). Moreover, many of the experiments reported pH that changed during the experimental run; this variability is likely to be replicated in field conditions and yet could be difficult to reliably predict prospectively. We therefore decided that our preferred model should not include pH as a predictor. A possible best fit model that incorporates the variable pHdiff7 as a predictor, with additional discussion about possible caveats is described in Supporting Information (Section S2).

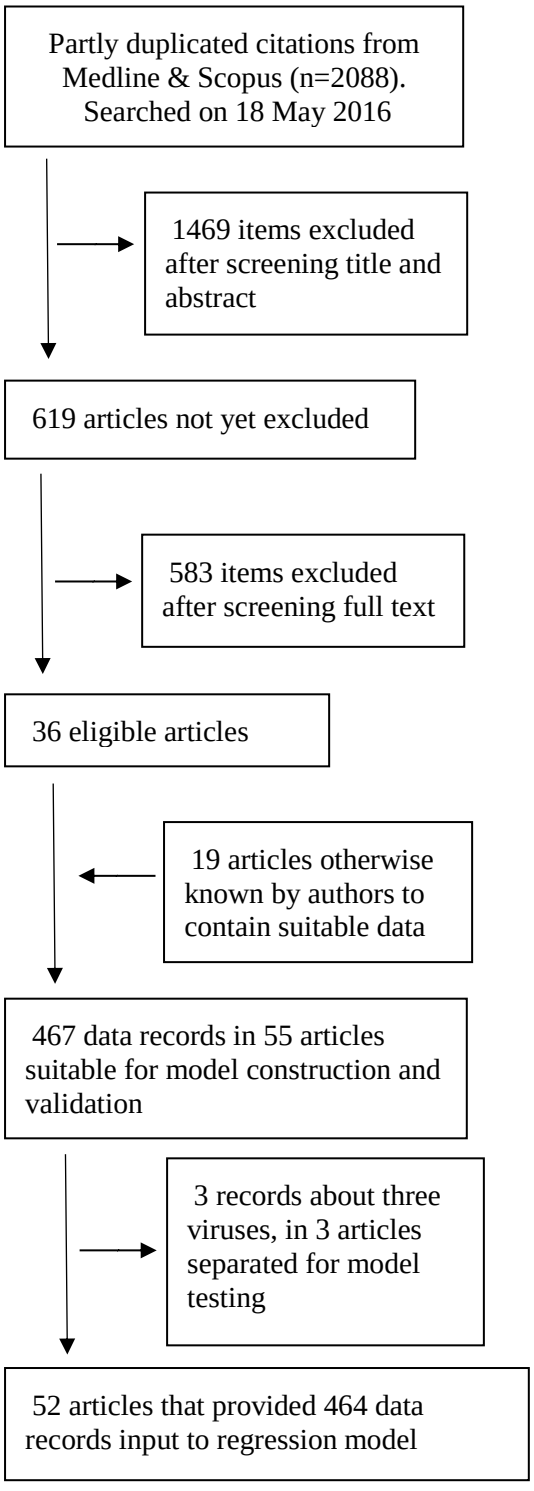


Figure 1. Study Selection Flow Chart.

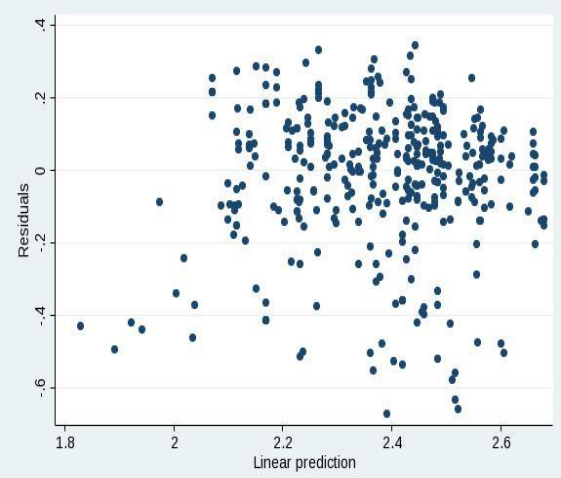
Model 1. Model 1, in Table 1, is our preferred model that does not use pH. An analysis of how well response and predictor meet required model data assumptions is available (Supporting Information). Coefficients, standard error, 95% confidence intervals, t- and p-values are shown below. There are five inputs: faecal-oral as a primary transmission pathway (or not), enveloped structure (or not), DNA rather than RNA structure, temperature and level of matrix contamination with faecal material. Linear temperature was a better predictor than logarithmic transformed temperature values or linear difference from room temperature (20° C). The robust standard error for residuals generated from Model 1 was 0.0190121.

Table 1. Model 1 coefficients and attributes, Censored regression to predict sqrt(T90secs).

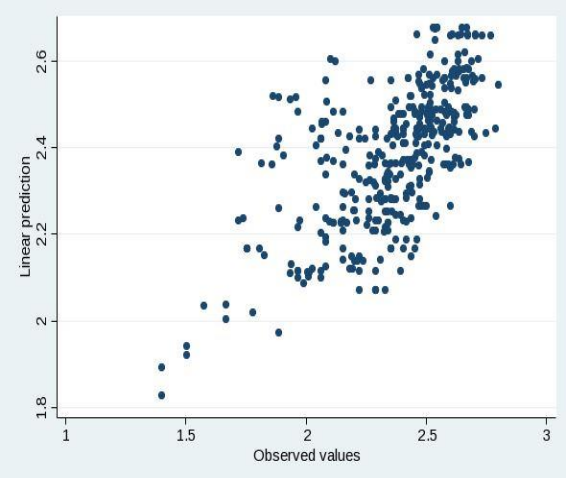
		95% CI for coeff. values		
	Coefficient	**Lower bound**	**Upper bound**	**p-value**
Model constant	2.56883	2.49456	2.64310	< 0.001
Faecal oral transmission pathway (y)	0.12877	0.07305	0.18448	< 0.001
Enveloped virus (y)	-0.09392	-0.15091	-0.03925	0.001
DNA virus (y)	0.01523	-0.02873	0.05918	0.496
Temperature in C°	-0.00971	-0.01136	-0.00805	< 0.001
Low contamination	0	na	na	Na
Medium contamination	0.00428	-0.04468	0.05323	0.864
High contamination	-0.11271	-0.15790	-0.06752	< 0.001

Notes: sqrt(T90secs) = square root[$\log_{10}$(T90 in seconds)]. Enveloped virus (y) = 1 when enveloped, else 0. Faecal oral (y) = 1 when faecal oral is primary transmission pathway, else 0. DNA virus (y) = 1 for DNA virus, else 0. Model default is when level of contamination = low, else model adjusts for when contamination is medium or high as indicated.

Figure 2 shows (a) residuals plotted on fitted values for all uncensored data; (b) fitted plotted on all uncensored observed values. Depicting and analyzing only uncensored residuals is appropriate because of the expected high errors for censored data. Mean value of residuals = -0.1898, standard deviation = 0.1920. An alternative model fit to the same data minus the most influential observations is available in Supporting Information (Section S4); the 95% confidence intervals for coefficients in this alternative model overlap generously with our preferred model so we do not explore this alternative further.



2a. Residuals plotted on predicted values.



2b. Predicted plotted on observed values.

Figure 2. Model 1 residuals, fitted and observed T90 values (log_{10} transformation). (a) Residuals (= observed – predicted values) plotted against prediResicted values, (b) fitted values plotted against observed values.

Virus inactivation as a function of temperature only in Model 1. Model 1 predicts T90 as a function of three binary variables, a three level categorical variable (contamination) and one interval input (temperature). There are eight combinations for the three binary variables (DNA or not, enveloped or not and faecal-oral or not, which are listed in Table 2. The opportunity arises to forecast T90s as a function of these finite combinations and the three levels of matrix contamination (low, medium or high), to relate the predicted T90s otherwise to only temperature, as shown in Figures 3a-3c. Figure 3a shows estimated virus survival (T90s expressed in hours) for a matrix with low contamination, Figure 3b shows corresponding data for a matrix with medium contamination, and estimated virus survival times in a highly contaminated matrix are shown in Figure 3c.

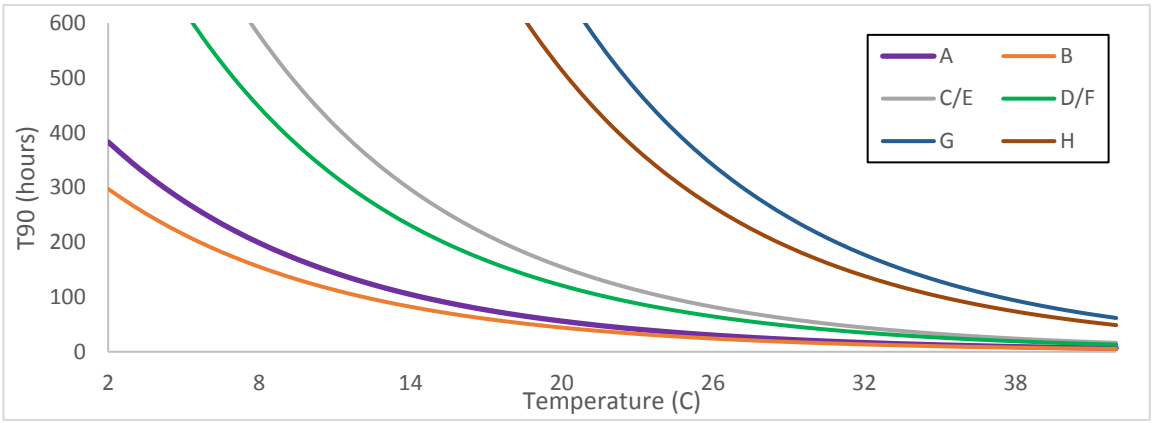
Table 2. Finite combinations of virus attributes applicable to Model 1 and Figure 3.

Group	Primary transmission pathway = Faecal-Oral	Enveloped virus	DNA = nucleic acid	Examples	% of input records within each group
A	0	1	1	Herpes simplex	3.9%
B	0	1	0	SARS coronavirus	23.7%
C	0	0	1	H. Adenovirus 2	5.6%
D	0	0	0	Human rhinovirus	3.4%
E	1	1	1	(none found)	0%
F	1	1	0	Swine fever	6.7%
G	1	0	1	Phi-X174 phage	19.4%
H	1	0	0	Poliovirus	37.3%

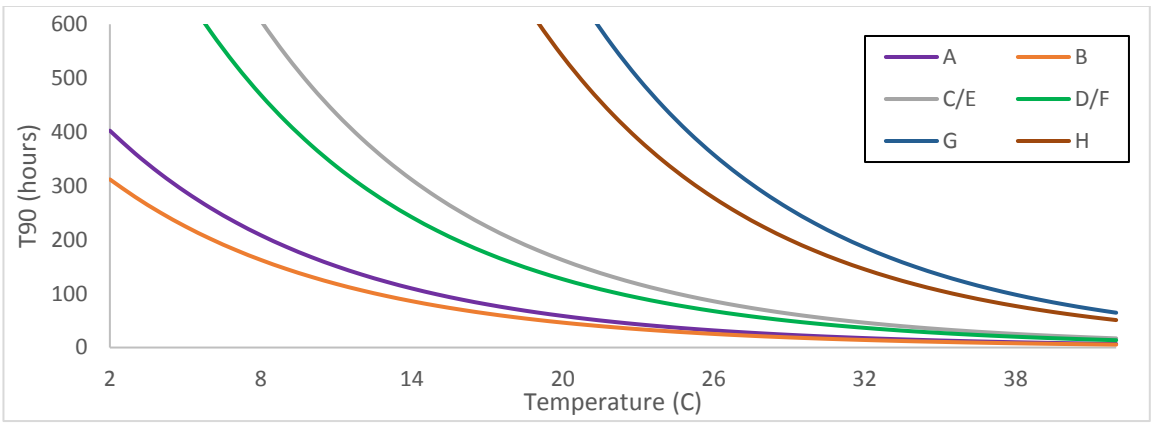
Note: 1=yes, 0 = no.

At the scale shown on Figures 3a-3c, Groups C and E are indistinguishable from each other, and likewise for Groups D and F. There are otherwise three visually apparent macro-groupings, (1) A and B, (2) C-F, (3) G and H. Groups A and B yield similar predicted T90s, the lowest predicted. Groups A and B are different from the other groups in being enveloped and not primarily faecal-oral-transmitted viruses. In contrast, groups G and H are the most long-lived groups: these are not-enveloped viruses with faecal-oral as their primary transmission pathway. The other four virus groups (C-F) form a third visually distinct cluster on Figures 3a-3c, comprising attribute combinations not in A,B, G and H. T90s for the Groups in matrices that have low or medium contamination are extremely similar: it is hard to tell Figures 3a and 3b apart. This result may be expected because of the lack of significance for the p-value on the medium level of contamination (category 2) in Model 1. However, estimated T90s are noticeably much more reduced when the matrix is highly contaminated (Figure 3c). The highly contaminated environment is relatively much more hostile to viral persistence, even for those viruses which are highly adapted to be transmitted through the faecal-oral route.

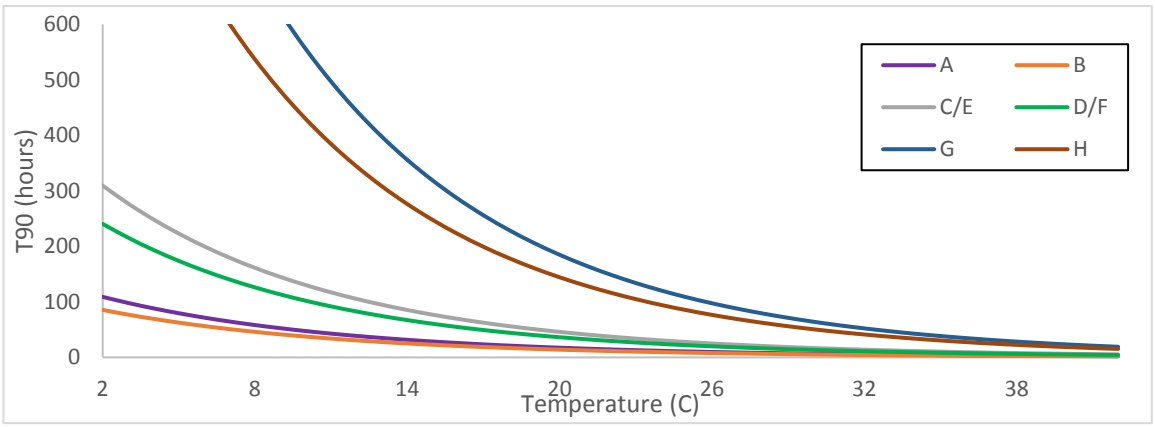
The uncertainty on the mean model estimates is high; some 95% confidence intervals for the model predictions (using medium contaminated matrices as an example) are shown in Supporting Information Section 5.



3a



3b



3c

Note: See Table 2 for Group descriptions, A-H.

Figure 3. T90 plotted for finite virus and matrix attribute combinations. 3a: Matrix has low level of faecal contamination; 3b: medium contamination; 3c: highly contaminated matrix.

Independent test data. Table 3 shows the predicted T90 (in hours) as a function of temperature and the virus/experimental conditions, for data in each of the independent test papers, as predicted by Model 1. Means and 95% confidence intervals were generated for predicted values, based on the 95% confidence intervals for each variable coefficient (as shown in Table 1). All of the confidence intervals are relatively large demonstrating high model uncertainty; still, the mean predictions are sometimes quite encouraging. The midpoint match is good for Fischer et al. 2015 [8], with less than 10% error: Ebola virus, predicted T90 (expressed in hours) = 40.0, observed T90 = 43.2 hours. The mean predicted T90 value is within 30% of the true value for Ahmed et al. 2014 [15]: human adenovirus, predicted T90 = 232 hours, observed T90 = 312 hours. The model output is a poorer fit, at 40% mean underestimate for the data in Adhikari et al. 2013 [14]: P22 phage, predicted T90 = 356 hours and observed T90 = > 500 hours. The observed value of 500 hours from Adhikari et al. is still comfortably within the 95% confidence intervals predicted by Model 1, but the censored nature of this test observation makes it impossible to confirm that the true T90 value is within these boundaries. 500-356 = 144 hours = six days. In absolute terms, six days is not a small error.

Table 3. T90 predictions tested against independent observations.

Reference	Adhikari et al. 2013	Ahmed et al. 2014	Fischer et al. 2015
Virus	P22 phage	Human adenoviruses	Ebola
Faecal oral	Y	N	N
Enveloped	N	N	Y
Nucleic acid	DNA	DNA	RNA
Environmental variables			
Level of contamination	High	Medium	Low
Temperature	14° C	16.7° C	21° C
Model estimates and Observed T90 (hours)			
Lower bound	38	44	7.5
Mean estimate	356	232	40.0
Upper bound	3360	1147	241.6
Observed T90	>500	312	43.2

Notes: Faecal-oral as primary transmission pathway (or not); Enveloped virus (or not). Predicted lower/upper bounds are bounds of 95% confidence interval.

Practical issues and limitations. With a much larger dataset, it could be valuable to develop models for each of the individual scenarios described in Table 2 (Groups A-H). Such customization might well improve model predictions for each combination of virus and environmental traits. It was not practical in this article to develop individual group models, or assess model fit by group, due to diverse sample sizes. For instance, Group E is described in no records in our dataset, whereas Group G conditions apply to 173 (37%) of 464 records). There is also merit in considering whether predictions could be clustered by group: ie, mean predicted inactivation was similar for three distinctive clusters: Groups A-B, C-F and G-H. Guidelines could be developed that treat these clusters of groups as similar in risk management, with regard to expected inactivation rates.

More data are required before we feel confident about including pH in our models. Many previous articles showed links between pH and rates of virus inactivation [21-24], although these sources are not consistent about the optimal pH for virus survival. It is problematic that available pH data are relatively limited in range, while pH data may be hard to reliably obtain or estimate in field conditions. 57% of our records had pH between 6 and 7.99, and another 42% of records had pH between 8 and 9.3. The remaining (1%) four experiments (with specific pH data) reported on pH below 6 (range = 2.1-2.6). Therefore, observed pH data were somewhat incomplete compared to the full possible pH range, and relatively discreet (noticeable gap in distribution of possible pH values).

We tried transformations of other variables in the predictive model, including quadratic and exponential transformations of temperature (away from a relatively microbial friendly condition of 20° C). These variable expressions did not improve model performance.

Geoghegan *et al*. 2016 [10] undertook somewhat similar research to ours, to explore whether biological features of viruses could indicate the likelihood of inter-human transmissibility. They determined that viruses with low host mortality, that establish long-term chronic infections, and that are non-segmented, non-enveloped, and not transmitted by vectors were more likely to be transmissible among humans. However, genome length, genome type, and recombination frequency were not predictive of human transmissibility. Our approach to modelling virus deactivation did not consider as wide a variety of biological traits, but we also did not find biological traits to be the strongest predictors in our modelling: in our Model 1, t-values were greatest for the environmental traits = contamination level and temperature.

There was an inevitable element of subjective judgment in the categorization of contamination level (low, medium or high). Some experimental data were grouped (sets of similar results from many very similar experiments within the same article); there was inevitable subjectivity in the grouping. Some variables were not clearly reported in the eligible articles; we mutually discussed the best representative value to record in such cases (such as for temperature, pH, T90, etc).

Clearly, the models we have described could be improved. Many of the papers did not report all potential predictor variables. In particular pH and whether or not the experiments were run in light or dark conditions was often not reported, even though both these variables are likely to impact on viral survival. To have included both these variables in the model would have meant losing a high proportion of the studies. For the ordinal variable representing degree of contamination of the matrix there was a degree of arbitrariness in the thresholds between the categories. Looking at the primary model it could be argued that the low and moderate contamination categories could be combined meaning that the important cutoff was between

the moderate and high contamination categories. Also, it was usually but not always clear what to assign to the variable 'primary transmission pathway'. If a virus was not normally faecally transmitted, it was not categorized as faecal-oral. However, we acknowledge that transmission pathways for some viruses are not very well understood and most, if not all, viruses *can* be transmitted via a faecal-oral route in at least some circumstances. An example is the epidemic of SARS which was mainly respiratory in transmission, but for which there was evidence of some spread via wastewater [1].

Implications for Public Health. We demonstrated that it is feasible to predict viral survival in different media from key virus and matrix attributes. Clearer reporting in future studies about matrix pH, light level exposure and temperature would probably reduce model uncertainty. While not perfect the model was successful at predicting virus survival to a reasonable degree of accuracy. The model also gives confidence intervals for its predictions. In the absence of more definitive experimental evidence this use of this model would give policy makers estimates of viral survival in different matrices to allow guideline development early in a new epidemic threat. This model should not be seen as an alternative to experimental evidence and does not remove the need to generate such evidence. Clearly, where experimental evidence subsequently conflicts with the predictions of this model then the former should take precedence and guidelines revised in light of this new experimental evidence.

ASSOCIATED CONTENT

Supporting Information Available: S1 List of references consulted to determine virus attributes; S2 Exploratory data analysis; S3 Best fit model that incorporates pH as predictor; S4 Impact of influential observations; S5 95% confidence intervals for viruses in medium contaminated matrices. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Author Contributions

P.R.H. and K.P. designed the study. K.P. and J.B. undertook searches, screened articles and extracted data. P.R.H. and J.B. undertook regression analysis. J.B. undertook other data analysis, wrote the first draft of the article and assembled revisions. All authors substantially commented on draft text and approve of this version of the manuscript.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

Thank you to three quick and anonymous referees who suggested many helpful improvements. This research was funded by the National Institute for Health Research (NIHR) Health Protection Research Unit in Emergency Preparedness and Response in partnership with Public Health England (PHE). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or PHE.

REFERENCES

1. McKinney, K. R.; Gong, Y. Y.; Lewis, T. G., Environmental transmission of SARS at Amoy Gardens. *J. Environ. Health* **2006,** *68* (9), 26.
2. Brown, J. D.; Swayne, D. E.; Cooper, R. J.; Burns, R. E.; Stallknecht, D. E., Persistence of H5 and H7 avian influenza viruses in water. *Avian Dis.* **2007,** *51* (s1), 285-289.
3. WHO, Ebola Virus Disease (EVD): Key questions and answers concerning water, sanitation and hygiene. In *WHO/EVD/WSH/14*, Organisation, W. H., Ed. 2014.
4. Lantagne, D. S.; Hunter, P. R., Comment on "Ebola virus persistence in the environment: state of the knowledge and research needs". *Environ. Sci. Technol. Lett.* **2015,** *2* (2), 48-49.
5. Bibby, K.; Fischer, R. J.; Casson, L. W.; Stachler, E.; Haas, C. N.; Munster, V. J., Persistence of Ebola Virus in Sterilized Wastewater. *Environ. Sci. Technol. Lett.* **2015,** *2* (9), 245-249.
6. Edmunds, K.; Elraham, S.; Bell, D. J.; Brainard, J.; Derviservic, S.; Fedha, T. P.; Few, R.; Howard, G.; Lake, I.; Maes, P., Dealing with Ebola-infected waste: a Hazard Analysis of Critical Control Points for reducing the risks to public health. *Bull. World Health Organ.* **2016,** *94* (6), 424-432.
7. Schuit, M.; Miller, D. M.; Reddick-Elick, M. S.; Wlazlowski, C. B.; Filone, C. M.; Herzog, A.; Colf, L. A.; Wahl-Jensen, V.; Hevey, M.; Noah, J. W., Differences in the comparative stability of ebola virus makona-c05 and yambuku-mayinga in blood. *PLoS One* **2016,** *11* (2).
8. Fischer, R.; Judson, S.; Miazgowicz, K.; Bushmaker, T.; Prescott, J.; Munster, V. J., Ebola virus stability on surfaces and in fluids in simulated outbreak environments. *Emerg. Infect. Dis.* **2015,** *21* (7), 1243.
9. World Health Organisation *Rapid Guidance on the Decomissioning of Ebola Care Facilities*; March, 2015; p 45.
10. Geoghegan, J. L.; Senior, A. M.; Giallonardo, F. D.; Holmes, E. C., Virological factors that increase the transmissibility of emerging human viruses. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113* (15), 4170-4175.
11. Wigginton, K.; Ye, Y.; Ellenberg, R., Emerging investigators series: the source and fate of pandemic viruses in the urban water cycle. *Environ. Sci.: Water Res. Technol.* **2015,** *1* (6), 735-746.
12. Sobsey, M. D.; Meschke, J. S., Virus survival in the environment with special attention to survival in sewage droplets and other environmental media of fecal or respiratory origin. *Report for the World Health Organization, Geneva, Switzerland* **2003,** *70*.
13. Magri, M. E.; Fidjeland, J.; Jönsson, H.; Albihn, A.; Vinnerås, B., Inactivation of adenovirus, reovirus and bacteriophages in fecal sludge by pH and ammonia. *Sci. Total Environ.* **2015,** *520*, 213-221.
14. Adhikari, U.; Harrigan, T.; Reinhold, D.; Waldhorn, A. A., Modeling seasonal variation in bacteriophage removal in constructed wetlands using convection–dispersion equation. *Ecol. Eng.* **2013,** *54*, 266-272.
15. Ahmed, W.; Gyawali, P.; Sidhu, J.; Toze, S., Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. *Lett. Appl. Microbiol.* **2014,** *59* (3), 348-354.
16. Cleves, M., *An Introduction to Survival Analysis Using Stata*. Stata Press: College Station, Texas, 2008; p 372.
17. Harrell, F. E.; Lee, K. L.; Califf, R. M.; Pryor, D. B.; Rosati, R. A., Regression modelling strategies for improved prognostic prediction. *Stat. Med.* **1984,** *3* (2), 143-152.
18. Killip, S.; Mahfoud, Z.; Pearce, K., What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann. Fam. Med.* **2004,** *2* (3), 204-208.
19. Baum, C. F.; Schaffer, M. E.; Stillman, S., ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. *Boston College Department of Economics, Statistical Software Components S* **2007,** *425401*, 2007.
20. Wooldridge, J. M., *Introductory Econometrics: A Modern Approach*. 6th ed.; CENGAGE Learning Custom Publishing: 2015.

21. Deng, M. Y.; Cliver, D. O., Persistence of inoculated hepatitis A virus in mixed human and animal wastes. *Appl. Environ. Microbiol.* **1995,** *61* (1), 87-91.
22. Lai, M. Y.; Cheng, P. K.; Lim, W. W., Survival of severe acute respiratory syndrome coronavirus. *Clin. Infect. Dis.* **2005,** *41* (7), e67-e71.
23. Mondal, T.; Rouch, D. A.; Thurbon, N.; Smith, S. R.; Deighton, M. A., Factors affecting decay of Salmonella Birkenhead and coliphage MS2 during mesophilic anaerobic digestion and air drying of sewage sludge. *J. Water Health* **2015,** *13* (2), 459-472.
24. Zhang, C.; Li, W.; Liu, W.; Zou, L.; Yan, C.; Lu, K.; Ren, H., T4-like phage Bp7, a potential antimicrobial agent for controlling drug-resistant Escherichia coli in chickens. *Appl. Environ. Microbiol.* **2013,** *79* (18), 5559-5565.