Journal of Environmental Management 299 (2021) 113566

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: www.elsevier.com/locate/jenvman

Environmental risk assessment by using disability adjusted life year via constructing of a generalized linear model for morbidity estimation of waterborne pathogens

Tingting Gao [a,b,*], Pengcheng Xu [b], Rong Chen [b], Xiaochang C. Wang [b], Mawuli Dzakpasu [b]

[a] School of Environmental and Chemical Engineering, Xi'an Polytechnic University, No.19 South Jinhua Road, Xi'an, 710048, PR China
[b] International S&T Cooperation Centre for Urban Alternative Water Resources Development, Key Lab of Environmental Engineering, Shaanxi Province, Xi'an University of Architecture and Technology, No.13 Yanta Road, Xi'an, 710055, PR China

ARTICLE INFO

Keywords:
Environmental disease burden
Morbidity analysis
Generalized linear model
Water reuse

ABSTRACT

The Environmental burden of disease (EBD) quantitatively evaluates the health impacts of pathogens by using the disability-adjusted life year (DALY) method. The life loss due to morbidity is a general expression for the EBD outcome and, thus, morbidity analysis is indispensable. Considering the deficiency of previous morbidity analysis methods, the objective of this study was to construct a linear morbidity model by using a generalized linear model (GLM) as a template and introducing exposure dose, pathogen toxicity and human immunity as impact variables. Human experimental data were collected for model fitting, and the results indicated a good fit of the majority of the pathogen data. Consequently, two practical cases of water reuse in Xi'an Siyuan University (Case 1) and Lake Cui, Kunming (Case 2) were selected for model validation. Results for case 1 indicated the major EBD to be attributed to rotaviruses (5.57×10^{-7} DALYs, 95% confidence interval (CI): 4.46×10^{-7}-1.72×10^{-4} DALYs) and sprinkling irrigation (5.12×10^{-7} DALYs, 95% CI: 1.95×10^{-7}-1.47×10^{1} DALYs). Conversely, that for case 2 is mainly attributed to noroviruses (1.42×10^{-7} DALYs, 95% CI: 7.51×10^{-11}-2.67×10^{-4} DALYs) and road flushing (1.62×10^{-7} DALYs, 95% CI: 1.16×10^{-7}-2.67×10^{-4} DALYs). However, comparison with the suggested threshold of 10^{-6} DALYs indicated the EBDs for both cases are acceptable and, thus, water reuse is confirmed to be safe. The methodology for morbidity modelling proposed in this research can effectively compensate for missing data in DALY calculation and, thereby, help to optimize the process for EBD evaluation.

1. Introduction

The Environmental burden of disease (EBD) analysis comprise a method for the qualitative and quantitative evaluation of the health impact caused by environmental pollution (WHO, 2002). Since pathogen exposure via different environmental media may also stimulate environmental pollution issues, and may lead to various health impacts, the evaluation of the disease burden caused by environmental pathogens is of vital importance. A previous method adopted for the quantitative microbial risk assessment (QMRA) was to estimate the infection rate (P_{inf}) caused by certain pathogens, and the health impact for this method is interpreted and quantified as infection risk (Haas et al., 1999). However, an infection can only be regarded as the first step of health outcome after exposure to pathogens, and infection may further develop into specific disease types, thus an objective analysis of the overall perspective of health development is required. The traditional QMRA analysis cannot provide such information to further describe the disease development after infection occurs nor to indicate the severity or duration of specific health impacts. However, due to lack of data, existing studies for the prediction of the probability and severity of illness given infection mainly based on epidemiological investigations or rough estimations of the prevalence of disease in developed countries which may lead to inaccurate or over-estimations of the health impacts for other countries or regions (Gyan et al., 2017). Besides, the characteristics of pathogen toxicity and human immunity are not considered in the previous health risk assessment method. Therefore, an extended health evaluation methodology needs to be put forward to assess the health impact caused by environmental pathogens in a comprehensive manner.

Disability-adjusted life year (DALY) is a method proposed by the

* Corresponding author. School of Environmental and Chemical Engineering, Xi'an Polytechnic University, No.19 South Jinhua Road, Xi'an, 710048, PR China.
E-mail address: 358964614@qq.com (T. Gao).

https://doi.org/10.1016/j.jenvman.2021.113566
Received 2 April 2021; Received in revised form 17 August 2021; Accepted 18 August 2021
Available online 20 August 2021

World Health Organization (WHO) for evaluating the overall disease burden attributable to global or regional disease outbreaks. It is expressed as the sum of years of healthy life lost due to disability and premature death. It is designed with the consideration of a variety of social factors, including disability weight, time preference, age weight and life expectancy. The DALY method can be used to quantitatively and comprehensively describe the health hazards of disease outbreaks due to certain environmental pollution events. The results can be used for the comparison of disease impacts occurring in different regions and among different populations, ages, or genders (Ishaq et al., 2020). Consequently, major risk factors, pollution transmission pathways, and sensitive populations for certain pollution events can be identified, and dynamic monitoring of human health status can be achieved accordingly (Ishaq et al., 2020). Considering the deficiency of the traditional health risk assessment method, and the absence of a theoretical basis and practical models for EBD analysis, the objective of this research is to build a framework for the EBD assessment specific to environmental pathogens by using the DALY method. This framework could provide a scientific basis and guidance for further health policy-making and disease prevention for certain environmental pollution issues (Salgot and Esther, 2006).

The major difficulty of an EBD study using DALY is building a relationship to describe the development from pathogen exposure to the occurrence of a certain health impact, and morbidity analysis would be the major link. Through morbidity analysis, the exposure dose of a certain pathogen can be converted to the probability of illness, namely morbidity, which is regarded as a key input data for DALY calculation. The interpretation of health impact can be extended to a practical disease outcome of morbidity instead of infection as that obtained from the previous QMRA method (Gao et al., 2015). Therefore, morbidity analysis is indispensable for the EBD assessment. Traditional methods for morbidity analyses mainly include: (1) a rough statistics of morbidity based on epidemiological investigations or animal experimental results obtained mainly from developed countries, (2) an assumption that infection rate approximately equals to morbidity, (3) a calculation of morbidity based on the product of infection rate (P_{inf}) and the risk of disease given infection ($P_{ill/inf}$) which also be determined from the method (1) (Havelaar et al., 2000; Dietz et al., 2000; Wei et al., 2012; Gyan et al., 2017). However, data obtained from methods (1) and (3) is of poor applicability and may lead to overestimation of the morbidity for undeveloped countries or regions, and data obtained from method (2) may affect the accuracy of the risk assessment results (Gyan et al., 2017). Also, as discussed in literatures, pathogens may cause various health effects depending on the properties of both host and the pathogenic organism (Teunis et al., 1999). The pathogenic actions of certain microorganisms mainly depend on its virulence, which means the ability to invade a host and cause different severity of health damages (Teunis et al., 1999). The human immune system can prevent the invasion of pathogens and interfere with their pathogenic actions, thus the health impacts also differ between hosts (Teunis et al., 1999). The health damage and its severity ultimately depend on the balance between the colonization potential of pathogens and the strength of host defences (Teunis et al., 1999). However, those factors of pathogenic characteristics of different pathogens and the immune status of various exposed human populations are not considered in the previous studies of morbidity analysis. Therefore, the accuracy and objectivity of the EBD analysis results may be affected. On this basis, a comprehensive method for the morbidity analysis should be developed and a reliable calculation model needs to be established for applying the EBD to environmental pathogens. Such a method will not only provide reliable and accurate disease data but also optimize the methodology for EBD studies.

A generalized linear model (GLM) was selected as a template to build a morbidity calculation model, which was first proposed by Nelder and Wedderburn in 1972 (David et al., 1999). The GLM is designed by the direct generalization of a traditional linear model and is used to describe the quantitative nonlinear logical relationship between independent and dependent variables (David et al., 1999). In practice, a variety of statistical models can be defined as GLM, such as logistic regression models, Probit regression models, Poisson regression models and negative binomial regression models (David et al., 1999). According to the basic definition of GLM, the average value of the dependent variable depends on the linear predicted value of an independent variable through a nonlinear link function. Thus, the GLM can be used to fit attribute variables or variables with values at specific intervals, such as the probability of incidents occurring (David et al., 1999). For instance, Namata et al. (2008) applied a GLM to analyse the infection rate of pathogens. Besides, according to the GLM definitions, the probability distribution of the independent variable in the GLM is generalized to the entire exponential distribution with interspersed parameters. Thus, the GLM is suitable for fitting continuous/discrete variables, or symmetrical variables as well, such as morbidity (David et al., 1999). For example, Bollaerts et al. (2008) used GLM to simulate the illness incidence under certain pathogen exposure doses. Therefore, it is practicable to take GLM as a template to construct a morbidity calculation model under given conditions.

The main objective of this research is to build a mathematical relationship for the quantitative estimation of morbidity by using GLM as a template, considering the limitations of traditional morbidity analysis methods. Based on the study of disease progression and pathogenesis of environmental pathogens, three major pathogenic factors including exposure dose, pathogen toxicity and human immunity are considered for morbidity modelling (Pulcini et al., 2014). Human experimental data based on previous studies are collected for model fitting (Mathewson et al., 1986; Hornick et al., 1970; Black et al., 1988; Levine et al., 1973, 1973, 1988; Ward et al., 1986; Teunis et al., 2008; Chappell et al., 1999). Two practical cases, namely, water reuse in Xi'an Siyuan University and Lake Cui, Kunming, are selected for model verification. The results can be used for disease prevention and risk management.

2. Methods

The framework for applying GLM to quantitatively estimating morbidity in an EBD study for environmental pathogens is shown in Fig. 1. In this figure, the establishment of a generalized linear morbidity model primarily consists of three steps including template selection, variable definition, and parameter determination, which are discussed in detail in section 2.1. The construction of the morbidity model cannot only provide a reliable disease data but also create a link between pathogen exposure analysis and DALY calculation and, therefore, help to optimize the framework for EBD assessment. The EBD framework constructed with the introduction of the morbidity model is mainly comprised of three parts, namely exposure analysis, morbidity analysis and disease burden calculation, which is discussed in section 2.2. Two practical cases are further selected for model application as discussed in section 3, and as a result, major risk factors during water reuse in Xi'an Siyuan University and Lake Cui are determined accordingly.

2.1. Construction of a linear morbidity calculation model

2.1.1. Fundamental methods for constructing a linear morbidity calculation model

According to the assumption proposed by McCullagh and Nelder (1999), if there is a nonlinear logic relationship between a dimensional random variable Y and a multidimensional random variable X, and the variable Y follows an exponential distribution, then a GLM, as shown in Eq. (1), can be used to describe the nonlinear relationship between X and Y (McCullagh and Nelder, 1999):

$$g(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon \quad (1)$$

where β is the model coefficient, ε is a random error term that generally follows a normal distribution, $g(Y)$ is the link function for the GLM

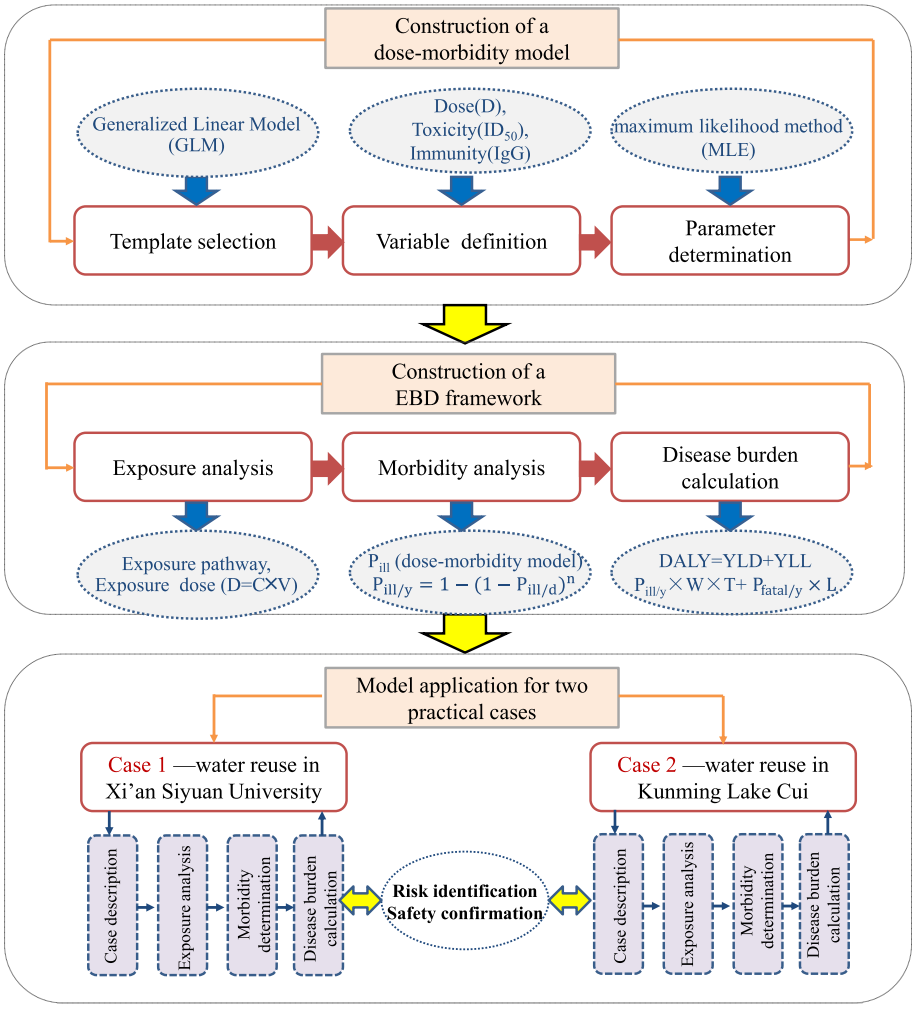


Fig. 1. Method configuration for applying a generalized linear model to quantitative estimation of morbidity in a disease burden study of environmental pathogens.

which has different expressions according to the probability distribution of variable *Y*.

By using Eq. (1) as a template to construct a morbidity calculation model, the dependent variable (*Y*) in Eq. (1) is determined as morbidity caused by a specific pathogen, which is expressed as *P*. The independent variables (*X*) are determined based on the analysis of the pathogenic mechanism and processes (Pulcini et al., 2014). As a result, three major variables that may affect the probability of illness are considered, according to the literature, namely, the pathogen exposure dose (*D*), pathogen toxicity (*T*) and human immunity (*S*) (Pulcini et al., 2014). Of these three variables, assuming exposure dose and toxicity are in co-determination of the pathogenic intensity of certain pathogens and human immunity affects the capability of the human body to eliminate pathogens, then the balance between these three factors can ultimately affect the development of certain health hazards (Pulcini et al., 2014). Therefore, *D*, *T* and *S* are determined as major independent variables (*X*) for the linear morbidity model (Eq. (1)) and are assumed to act independently of each other. Moreover, for a specific pathogen, we assume that the value of the variable *P* increases with those of the variables *D* and *T*, and decreases with that of the variable *S*. Thus, a general expression of the linear morbidity model can be derived as shown in Eq. (2):

$$g(p) = \beta_0 + \beta_1 \mathrm{D} + \beta_2 T + \beta_3 / S + \varepsilon \quad (2)$$

where p represents morbidity; β_0, β_1, β_2, and β_3 are defined as model coefficients; and ε is the deviations between $g(p)$ and its linear predicted values, which is normally distributed. The characteristics of each variable are summarized in Table 1.

Additionally, as shown in Table 1, the units of measure for the variables *D*, *T*, and *S* are inconsistent. Thus, dimensionless processing of these variables in Eq. (2) is indispensable for further model calculation. For unit conversion, we assume an average human body weight of 70 kg and serum content of the blood of a normal human is 37.5 mL/kg (Zu et al., 1994). Accordingly, the nondimensionalization of the above

Table 1
Implication and characterization of each variable in the linear morbidity model.

Variable	Characterization	Unit	Reference
D[a]	Characterized by internal dose or absorbed dose	mg	Chen et al. (2006)
T[b]	Characterized by median infective dose (ID_{50})	mg/kg	Zambriski et al. (2013)
S[c]	Characterized by serum antibody levels (*IgG*)	μg/mL	Teunis et al. (2002)

Note.
[a] The amount of pathogen that enters a human body and interact with human cells, usually can be measured in the unit of 'mg' (Chen et al., 2006).
[b] The minimum amount of pathogen that can cause infection among half the populations during a given period via a particular route usually can be measured in the unit of 'mg' per 'kg' of body weight (Zambriski et al., 2013).
[c] The amount of immune serum globulin is expressed by the IgG level because of its high concentration in the serum, which can be measured in the unit of 'ug' per 'ml' of the serum (Teunis et al., 2002).

variable units can be achieved through the following calculations: (1) $D(mg)\times\frac{1kg}{70kg\times10^6mg}=\mathrm{D}\times\frac{1}{7\times10^7}$, (2) $T\left(\frac{1mg}{kg}\right)\times\frac{1kg}{10^6mg}=T\times\frac{1}{10^6}$ and (3) $S\left(\frac{1\mu g}{1ml}\right)\times\frac{37.5ml}{10^9\mu g}=S\times\frac{37.5}{10^9}$. Therefore, three dimensionless coefficients are defined for the nondimensionalization of the variables *D*, *T*, and *S* as $a=\frac{1}{7\times10^7}$, $b=\frac{1}{10^6}$ and $c=\frac{37.5}{10^9}$, respectively.

On this basis, a multiplication of the dimensionless coefficients (*a*, *b*, *c*) with variables (*D*, *T*, *S*) is included in Eq. (2), and a dimensionless expression of the linear morbidity model can be derived as shown in Eq. (3):

$$g(p)=\beta_0+\beta_1(aD)+\beta_2(bT)+\frac{\beta_3}{cS}+\varepsilon \tag{3}$$

Additionally, for a given disease outbreak, assuming morbidity (*P*) follows a binomial distribution, a logit function is used for expressing the link function in Eq. (3). It is proved to fit well to the majority of data obtained from a disease outbreak or experimental cases, according to the definition of GLM (McCullagh and Nelder, 1999; Scarpello and Ritelli, 2008). The expression of the logit function is shown in Eq. (4) (McCullagh and Nelder, 1999). The ultimate logistic expression for the linear morbidity model can be derived as Eq. (5):

$$Logit\ functiong:\ (p)=\ln\left(\frac{p}{1-p}\right) \tag{4}$$

$$ln(p/(1-p))=\beta_0+\beta_1(aD)+\beta_2(bT)+\beta_3/(cS)+\varepsilon \tag{5}$$

2.1.2. Method for parameter determination of the linear morbidity model-

A maximum likelihood method (MLE) is applied for parameter determination of the GLMs (Yue and Chen, 2004). This method firstly assumes the number of the sample set is N, and n samples are randomly selected. The observed values of these random samples are defined as y_1, y_2, …, y_n, of which $y_i=1$ represents the occurrence of illness and $y_i=0$ represents non-illness status ($i=1,\ldots,\mathrm{n}$). On this basis, the probability of illness under a given condition of x_i is determined as $p_i=p\ (y_i=1|x_i)$, while the probability of non-illness status is expressed as $1–p_i$ (Yue and Chen, 2004). Therefore, for given observed outcomes of $y_1, y_2, \ldots, y_n$, the probability of each observed value $y_i=1$ or 0, $P(y_i)$ can be calculated by using Eq. (6):

$$P(y_i)=p_i^{y_i}(1-p_i)^{1-y_i} \tag{6}$$

where p_i can be calculated by using Eq. (7), which is transformed from Eq. (5), where $D'=aD$, $T'=bT$, $S'=cS$. The variables and parameters in Eq. (7) are the same as those introduced above.

$$p_i=\frac{e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}{1+e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}} \tag{7}$$

Secondly, assuming each observed value (y_i) is independent, a likelihood function is constructed according to the basic theory of MLE (Verbeke and Molenberghs, 2000), as shown in Eq. (8):

$$f(\beta)=\prod_{i=1}^{n}p_i^{y_i}(1-p_i)^{1-y_i},\quad \beta=(\beta_0,\ \beta_1,\ \beta_2,\ \beta_3) \tag{8}$$

Additionally, for calculation convenience, a log-likelihood function is derived by calculating the logarithm of both sides of Eq. (8), as shown in Eq. (9). Because the variable *ln(β)* in Eq. (9) is a monotonic function of the variable *f(β)*, the value of *f(β)* reaches its maximum when the value of *ln(β)* increases to the maximum at a given value of *β* (Verbeke and Molenberghs, 2000):

$$\ln(\beta)=\sum_{i=1}^{n}y_i(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')-ln\left(1+e^{\beta_0+\beta_1D'+\beta_2T'+\beta_3/S'}\right) \tag{9}$$

Finally, according to the basic theory of MLE, the determination of parameter values should ensure that the value of $\ln(\beta)$ reaches its maximum (Verbeke and Molenberghs, 2000). Therefore, partial derivatives specific to each parameter (β_0, β_1, β_2, β_3) are calculated on both sides of Eq. (9), and the value of each partial derivative is determined as 0. Therefore, a series of logarithmic likelihood equations can be derived as Eq. (10.1)–Eq. (10.4):

$$\frac{\partial ln(\beta)}{\partial\beta_0}=\sum_{i=1}^{n}\left(y_i-\frac{e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}{1+e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}\right)=0 \tag{10.1}$$

$$\frac{\partial\ln(\beta)}{\partial\beta_1}=\sum_{i=1}^{n}\left(y_i-\frac{e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}{1+e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}\right)D'=0 \tag{10.2}$$

$$\frac{\partial\ln(\beta)}{\partial\beta_2}=\sum_{i=1}^{n}\left(y_i-\frac{e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}{1+e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}\right)T'=0 \tag{10.3}$$

$$\frac{\partial\ln(\beta)}{\partial\beta_3}=\sum_{i=1}^{n}\left(y_i-\frac{e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}{1+e^{(\beta_0+\beta_1D'+\beta_2T'+\beta_3/S')}}\right)\frac{1}{S'}=0 \tag{10.4}$$

Solutions obtained from Eq. (10.1)–Eq. (10.4) are determined as estimates of the parameters β_0, β_1, β_2, and β_3. However, because Eq. (10.1)–Eq. (10.4) are nonlinear, direct solutions are difficult to obtain, and a suitable software is indispensable for practical calculations. In this study, the software *Origin 8.0* was applied for model fitting, and human experimental (HE) data for typical pathogens available from previous studies were collected for parameter determination (Teunis et al., 1996). Consequently, the linear morbidity calculation model (Eq. (5)) was fitted to the data of *Escherichia coli*, *Salmonella*, *Campylobacter*, *Shigella*, *Vibrio cholerae*, Rotavirus, Norovirus, and *Cryptosporidium*, respectively, and the results are shown in Fig. 2 (Mathewson et al., 1986; Hornick et al., 1970; Black et al., 1988; Levine et al., 1973, 1973, 1988; Ward et al., 1986; Teunis et al., 2008; Chappell et al., 1999). In this figure, the scattered points represent the morbidity obtained from the human experimental data, and the continuous curves represent the morbidity calculated by using the linear dose-morbidity model. Most of the experimental points fall on or near to the model curves, and can predict the tendency that morbidity varies with exposure doses accurately, thus showing a good fit of the model with the available experimental data for the majority of the pathogens. Therefore, the dose-morbidity model can be used for the proper estimation of the probability of illness directly given certain doses of typical environmental pathogens. Consequently, the EBD caused by typical pathogens can be evaluated accordingly by taking morbidity as input data. However, the human experiments conducted so far are all in small sample size, and the data points obtained in this research are limited, therefore the degree of freedom are relatively small in this study. The results for parameter fitting specific to *Escherichia coli*, *Salmonella*, *Campylobacter*, *Shigella*, *Vibrio cholerae*, Rotavirus, Norovirus, and *Cryptosporidium* are summarized in Table 2.

2.2. Construction of EBD framework incorporating generalized linear morbidity model

2.2.1. Exposure analysis

Water sampling and Real-time quantitative Polymerase Chain Reaction (PCR) were for the quantification of exposure concentrations by using a SYBR Primix Dimer Eraser™ kit (Takara, Japan) in an iCycler iQ5 Real Time PCR Detection System (Bio-Rad, Hercules, California, USA). The method and procedure for PCR detection are suggested based on a previous study (Zhou et al., 2015).

Exposure pathways for certain pathogens are determined based on field investigations, and an average exposure dose for certain pathogens specific to each pathway is calculated by using Eq. (11) (Gao et al., 2015):

$$D=C\times V \tag{11}$$

where D represents the average exposure dose of a certain pathogen

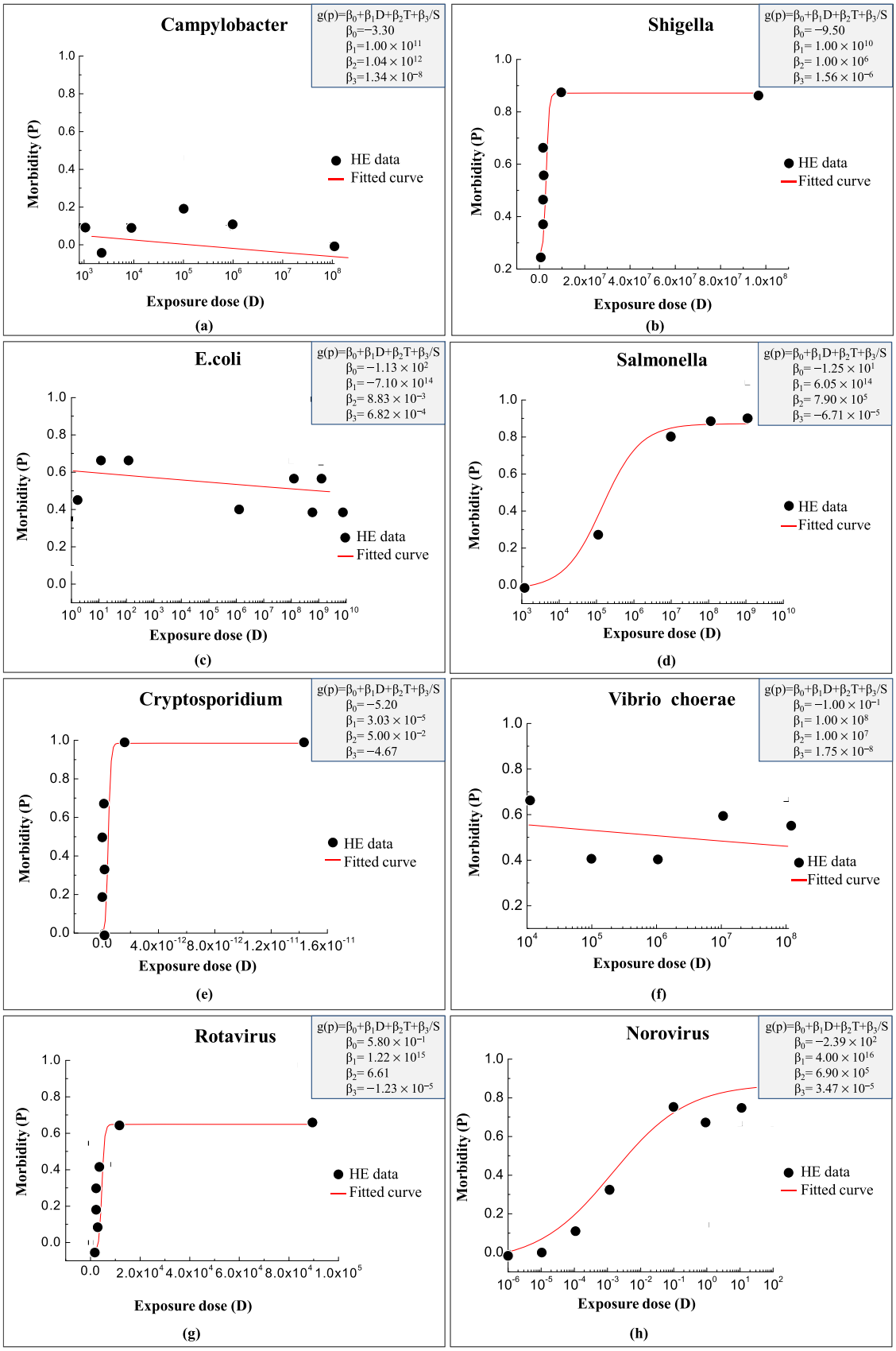


Fig. 2. Model fitting results for pathogens: (a) *Campylobacter*, (b) *Shigella*, (c) *E. coli*, (d) *Salmonella*, (e) *Cryptosporidium*, (f) *V. cholerae*, (g) Rotavirus and (h) Norovirus (Mathewson et al., 1986; Hornick et al., 1970; Black et al., 1988; Levine et al., 1973, 1973, 1988; Ward et al., 1986; Teunis et al., 2008; Chappell et al., 1999).

Table 2
Parameter fitting results for a linear morbidity model according to available human experimental data on typical environmental pathogens.

Pathogen	Parameter estimates				D[a]	Df[b] (n[c])	References
	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$			
E. coli	-1.13×10^{2}	-7.10×10^{14}	8.83×10^{-3}	6.82×10^{-4}	0.149	1 (5)	Mathewson et al. (1986)
Salmonella	-1.25×10^{1}	6.05×10^{14}	7.90×10^{5}	-6.71×10^{-5}	5.42	3 (5)	Hornick et al. (1970)
Campylobacter	−3.30	1.00×10^{11}	1.04×10^{12}	1.34×10^{-8}	2.422	3 (6)	Black et al. (1988)
Shigella	−9.50	1.00×10^{10}	1.00×10^{6}	1.56×10^{-6}	1.08	2 (5)	Levine et al. (1973)
V. cholerae	-1.00×10^{-1}	1.00×10^{8}	1.00×10^{7}	1.75×10^{-8}	1.75	5 (5)	Levine et al. (1988)
Rotavirus	5.80×10^{-1}	1.22×10^{15}	6.61	-1.23×10^{-5}	6.18	6 (8)	Ward et al. (1986)
Norovirus	-2.39×10^{2}	4.00×10^{16}	6.90×10^{5}	3.47×10^{-5}	5.126	4 (8)	Teunis et al. (2008)
Cryptosporidium	−5.20	3.03×10^{-5}	5.00×10^{-2}	−4.67	0.360	7 (8)	Chappell et al. (1999)

Note.
[a] deviance.
[b] numbers of degrees of freedom.
[c] the total number of samples.

(copies/d); C represents the average exposure concentration of a specific pathogen detected by the PCR method (copies/mL); V represents the volume of reclaimed water ingested via a certain pathway (mL/d), which is determined from the literature (Gao et al., 2015).

2.2.2. Morbidity analysis

In previous studies, morbidity is usually estimated by using Eq. (12) (Gao et al., 2015). In this equation, the morbidity (P_{ill}) is calculated by the product of infection rate (P_{inf}) and the probability of illness given infection ($P_{ill/inf}$) (Gao et al., 2015). For this method, the infection rate is usually estimated by using an exponential or Beta-Poisson dose-response model with determined dose-response parameters of 'r', 'α' and 'N_{50}', shown as Eq. (13) and Eq. (14), respectively (Haas et al., 1999). The probability of illness given infection ($P_{ill/inf}$) is calculated by using an 'Illness-infection' model with determined dose-response parameters of 'γ' and 'σ', shown as Eq. (15) (Teunis et al., 1999):

$$P_{ill} = P_{inf} \times P_{ill/inf} \tag{12}$$

$$P_{inf} = 1 - e^{-rD} \tag{13}$$

$$P_{inf} = 1 - \left[1 + \frac{D}{N_{50}}\left(2^{\frac{1}{\alpha}} - 1\right)\right]^{-\alpha} \tag{14}$$

$$P_{ill/inf} = 1 - (1 + \gamma D)^{-\sigma} \tag{15}$$

In this research, the morbidity caused by specific pathogens are calculated by using Eq. (5) under given dose and toxicity of pathogens and immune status of human bodies. In this equation, the variables of ID_{50} and IgG represent the toxicity of pathogens and the immunity of human bodies respectively, and are determined based on previous studies. The dose-morbidity coefficients (β_0, β_1, β_2, β_3) specific to each pathogen are fitted by experimental data collected from previous studies as shown in Fig. 2 (Mathewson et al., 1986; Hornick et al., 1970; Black et al., 1988; Levine et al., 1973, 1973, 1988; Ward et al., 1986; Teunis et al., 2008; Chappell et al., 1999).

A daily average morbidity ($P_{ill/d}$) caused by specific pathogens is calculated by using Eq. (5), and the conventional method shown as Eq. (12) ~ Eq. (15) can also be adopted for results comparison. An annual average morbidity ($P_{ill/y}$) used for the annual burden of disease calculation is computed by using Eq. (16) (Gao et al., 2015):

$$P_{ill/y} = 1 - \left(1 - P_{ill/d}\right)^{n} \tag{16}$$

where n represents the annual exposure frequency specific to each exposure pathway and is determined based on field investigations.

To account for the uncertainty of morbidity estimations, a Monte Carlo simulation is further performed (Alexander, 2003). It is a numerical simulation calculation process. For this method, a random sampling of the variables according to their probability distributions is conducted by computers and the calculation is repeated hundreds or thousands of times (Alexander, 2003). In this research, variables D is calculated based on variable C (Eq. (11)), and variable C is lognormal distributed based on a previous study (Dean, 1981), thus variable C is random sampled and iteratively calculated for 10,000 iterations in this study (Alexander, 2003). Variable ID_{50} and IgG are determined as point estimates according to literatures (Teunis et al., 1996, 2008; Vardinon et al., 1999; Szu et al., 2013; Menon et al., 2013; Clemente et al., 2015). As a result, the value of morbidity is outputted in a certain interval rather than a single value to account for its uncertainties (Alexander, 2003). The Monte Carlo simulation is the most practical and effective method to solve the problems of randomness and uncertainty in risk assessment (Alexander, 2003).

2.2.3. Disease burden calculation

The disease burden is evaluated in terms of DALYs by using Eq. (17) ~(19), where YLD represents the healthy years of life lost due to disability. It is computed by a multiplication of the annual morbidity ($P_{ill/y}$), disease weight (W) and disease duration (T). YLL represents years of healthy life lost due to premature death. It is calculated as a product of annual mortality ($P_{fatal/y}$) and life expectancy (L) (Murray and Lopez, 1996). The disease parameters (W, T, L) are determined from the literature. The sum of YLD and YLL can be used to describe the total burden of disease (DALYs) caused by environmental pathogen exposures (Murray and Lopez, 1996). Additionally, a Monte Carlo simulation was also performed to account for the uncertainty of DALY evaluations (Alexander, 2003).

$$DALY = YLD + YLL \tag{17}$$

$$YLD = P_{ill/y} \times W \times T \tag{18}$$

$$YLL = P_{fatal/y} \times L \tag{19}$$

3. Case study

3.1. Case description

Xi'an Siyuan University (Case 1) is located in the south-eastern suburbs of Xi'an in China. Water supply in this university depends mainly on groundwater pumping at a combined maximum capacity of 3000 m^3/d. However, the actual water demand for this campus is estimated to be approximately 6000 m^3/d, which far exceeds the water supply capability. Therefore, to alleviate the pressure on water supply and to reduce the amount of sewage discharge, wastewater in this campus is collected and treated in the Siyuan wastewater treatment plant (WWTP). The main treatment process of this WWTP is a combination of Anoxic-Anaerobic-Aerobic (A^2O) with a Membrane Bio-Reactor (MBR). After treatment, wastewater is discharged into the

Siyuan Lake for water replenishment and reused for sprinkling irrigation and toilet flushing.

Lake Cui (Case 2) is located in the Wuhua district of Kunming, Yunnan in China. It is a small and shallow lake with a surface area of 15 $\times 10^4$ m^2. The major source of water replenishment for this lake is the discharge of reclaimed water produced by the Kunming 4th WWTP. The main treatment process of this WWTP is A^2O-MBR, which produces 6000 m^3/d of reclaimed water. The replenished reclaimed water in Lake Cui is reused for watercourse cleaning, boating, and road flushing.

However, a variety of pathogenic microorganisms can be detected in the reclaimed water and may raise potential health risks as a consequence. The probability of illness and EBD caused by typical pathogens during water reuse in Xi'an Siyuan University and Lake Cui, Kunming should be quantitatively evaluated to ensure water reuse safety.

3.2. Exposure analysis

According to epidemiological investigations, *E. coli*, *Salmonella,* and Rotavirus are identified as major risk factors in case 1, and *E. coli* and Norovirus are identified as typical pathogens for case 2 (Gerba et al., 1985; Tao et al., 2003; Bitton, 2005). Water sampling and Real-time PCR are carried out for the quantification of *E. coli*, *Salmonella,* Rotavirus, and Norovirus concentrations. The primers/probes used for PCR detection is obtained from the literature (Zhou et al., 2015).

The concentration (C) detected for typical pathogens in the reclaimed water is lognormal distributed (Dean, 1981). The average C detected for *E. coli* in case 1 is 9.36 $\times 10^2$ copies/100 mL, with a 95% confidence interval (95%CI) of 8.41 $\times 10^1$–3.50 $\times 10^3$ copies/100 mL, and a standard deviation (SD) of 7.21 $\times 10^2$. For *Salmonella*, it is 9.90 $\times 10^2$ copies/100 mL (95%CI: 2.08 $\times 10^1$–9.45 $\times 10^3$, SD: 2.15 $\times 10^3$), while for Rotavirus, it is 1.60 $\times 10^1$ copies/100 mL (95%CI: 1.32–2.32 $\times 10^2$, SD: 4.70 $\times 10^1$). The average C of *E. coli* and Norovirus detected in case 2 are 2.89 $\times 10^2$ (95%CI: 6.18–9.10 $\times 10^2$, SD: 2.87 $\times 10^2$) and 1.45 $\times 10^2$ copies/100 mL (95%CI: 1.23 $\times 10^2$–1.67 $\times 10^2$, SD: 6.20 $\times 10^1$), respectively.

According to field investigations, the exposure pathways for case 1 are sprinkling irrigation, fountains, and toilet flushing, and the contact route is identified as inhalation. For case 2, the general exposure pathways are watercourse cleaning, road flushing, and boating, and the contact routes are inhalation and skin contact. The average concentrations (C) for pathogens are used for exposure dose estimates specific to case 1 and 2 by using Eq. (11), with the unit of 'copies/d', where C is determined according to the PCR detection and V is determined based on previous literature as shown in Table 3 (USEPA, 1997; Chen et al., 2006; Xie et al., 2009; He et al., 2006).

Results for the average value of exposure doses calculated for cases 1 and 2 are shown in Tables 4 and 5, respectively. The unit of exposure dose is further converted to 'mg/d' under the assumption that the number of pathogens in a "1 mg" sample is supposed to be 10^9 copies (Teunis et al., 1996). Results show that, in Case 1, *Salmonella* and sprinkling irrigation lead to the largest exposure doses compared with other pathogens and pathways, with estimations of average values of 1.90 $\times 10^{-6}$ mg/d and 3.40 $\times 10^{-6}$ mg/d, respectively. In Case 2, the largest exposure doses are identified as *E. coli* and road flushing with estimations of average values of 1.95 $\times 10^{-7}$ mg/d and 2.60 $\times 10^{-7}$ mg/d respectively.

Table 3
Ingestion volume (V) and exposure frequency (n) for reclaimed water specific to each pathway in Cases 1 and 2.

Case	Exposure route	V (mL/d)	n (d/y)	References
Case 1	Sprinkling irrigation	175	275	USEPA, 1997; Chen et al. (2006); Xie et al. (2009); He et al. (2006)
	Fountains	16.8	244	
	Toilet flushing	0.05	365	
Case 2	Watercourse cleaning	3.7	40	
	Boating	3.7	14.9	
	Road flushing	60	164.8	

Table 4
Results for exposure dose calculation for case 1, in mg/d.

Case1	*E. coli*	*Salmonella*	Rotavirus	Total
Sprinkling irrigation	1.64×10^{-6}	1.73×10^{-6}	2.85×10^{-8}	3.40×10^{-6}
Fountains	1.58×10^{-7}	1.67×10^{-7}	2.74×10^{-9}	3.27×10^{-7}
Toilet flushing	9.36×10^{-11}	9.90×10^{-11}	1.63×10^{-12}	1.94×10^{-10}
Total	1.80×10^{-6}	1.90×10^{-6}	3.13×10^{-8}	

Table 5
Results for exposure dose calculation for case 2, in mg/d.

Case 2	*E. coli*	Norovirus	Total
Watercourse cleaning	1.07×10^{-8}	5.37×10^{-9}	1.61×10^{-8}
Boating	1.07×10^{-8}	5.37×10^{-9}	1.61×10^{-8}
Road flushing	1.73×10^{-7}	8.97×10^{-8}	2.60×10^{-7}
Total	1.95×10^{-7}	9.77×10^{-8}	

3.3. Morbidity determination

The morbidity caused by specific pathogens is calculated by using Eq. (5), and the variables ID_{50} and IgG are determined as point values by referring to literature (Teunis et al., 1996, 2008; Vardinon et al., 1999; Szu et al., 2013; Menon et al., 2013; Clemente et al., 2015). In Case 1, the value of variable ID_{50} determined for *E. coli*, *Salmonella* and Rotavirus are 6 $\times 10^9$ mg/kg, 36.3 mg/kg and 9.17 $\times 10^4$ mg/kg respectively (Teunis et al., 1996). Also, 470 μg/mL, 32.79 μg/mL and 14 μg/mL are suggested for the value of variable IgG, respectively (Vardinon et al., 1999; Szu et al., 2013; Clemente et al., 2015). In Case 2, the value of variables ID_{50} and IgG for *E. coli* are determined to be the same as in Case 1 (Teunis et al., 1996; Vardinon et al., 1999), and those values for Norovirus are suggested to be 1.34 $\times 10^2$ and 12.2 μg/mL, respectively (Teunis et al., 2008; Menon et al., 2013). The dose-morbidity parameters ($\beta_0, \beta_1, \beta_2, \beta_3$) specific to *E. coli*, *Salmonella*, Rotavirus, and Norovirus are fitted based on previous experimental data as shown in Fig. 2.

The daily average morbidity ($P_{ill/d}$) and annual average morbidity ($P_{ill/y}$) caused by typical pathogens in Cases 1 and 2 are calculated by using Eq. (5) and Eq. (16) respectively. The annual exposure frequency specific to each pathway is presented in Table 3 (USEPA, 1997; Chen et al., 2006; Xie et al., 2009; He et al., 2006). A Monte Carlo simulation is performed to account for uncertainties, and the model was run for 10, 000 iterations (Alexander, 2003). The results are shown in Fig. 3. In Case 1, Rotavirus and toilet flushing lead to the highest annual morbidity with average estimations of 7.73 $\times 10^{-8}$ (95% CI: 6.47 $\times 10^{-8}$-2.39 $\times 10^{-5}$) and 9.07 $\times 10^{-8}$ (95% CI: 9.10 $\times 10^{-8}$-8.99 $\times 10^{-8}$) respectively, and contribute 49.16% and 57.64%, respectively, to the total morbidity. In Case 2, Noroviruses (75.00%) and road flushing (84.14%) are identified as the largest contributors to the annual morbidity estimates, with average estimations of 4.25 $\times 10^{-8}$ (95% CI: 2.25 $\times 10^{-11}$-8.03 $\times 10^{-5}$) and 7.16 $\times 10^{-8}$ (95% CI: 4.35 $\times 10^{-8}$-1.20 $\times 10^{-4}$), respectively. However, comparing with the WHO recommended value of 10^{-5} for health risk probability, the morbidity caused by typical pathogens in Cases 1 and 2 are both far below the threshold (WHO, 2011).

3.4. Disease burden calculation

According to epidemiological investigations, major health outcomes

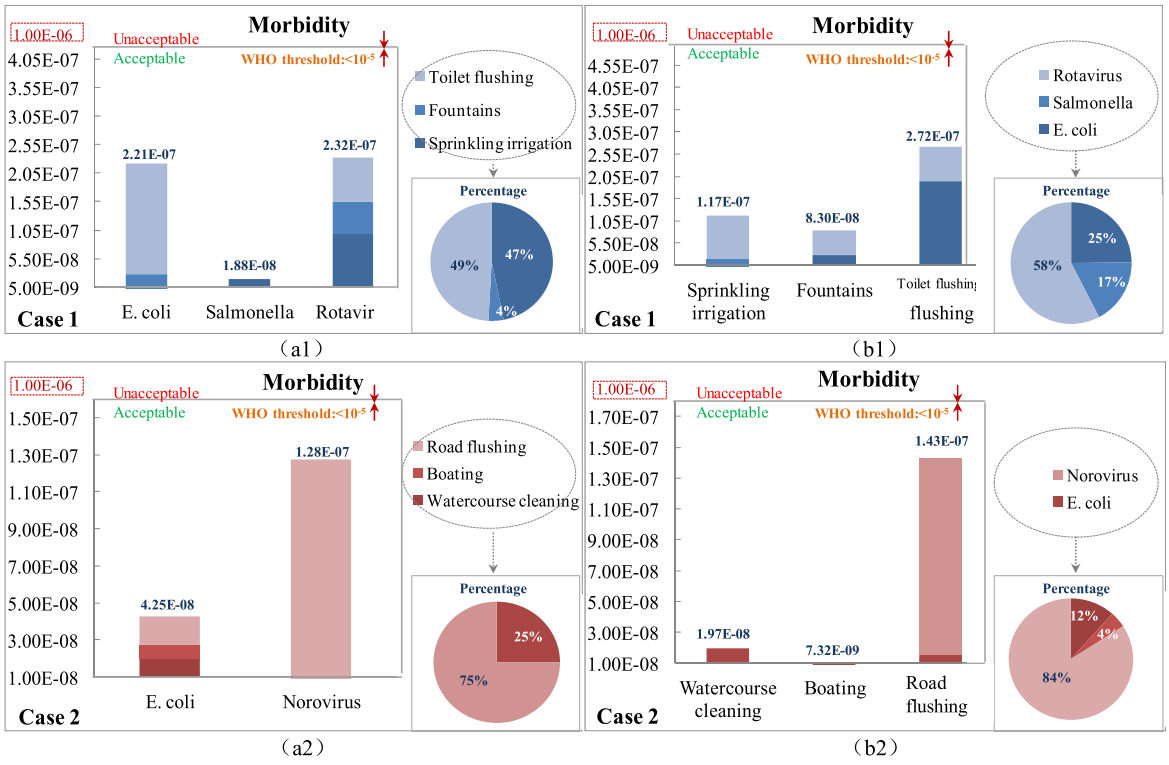


Fig. 3. Results for annual morbidity risk ($P_{ill/y}$) for each pathogen (a1, a2)/exposure pathway (b1, b2) in Cases 1 and 2.

caused by *E. coli*, *Salmonella*, Rotavirus, and Norovirus are determined as diarrhoea (watery/bloody), typhoid, and acute gastroenteritis, respectively (Havelaar and Melse, 2003). On this basis, disease parameters, including disability weight (*W*), disease duration (*T*) and proportion of different health outcomes (p) are determined according to previous studies as summarized in Table 6 (Murray and Lopez, 1996; Havelaar and Melse, 2003; Howard and Pedley, 2004; Gu et al., 2002). The disease burden is calculated by using Eq. (17)~(19). However, since the health outcomes attributed to *E. coli*, *Salmonella*, Rotavirus, and Norovirus in Cases 1 and 2 are generally nonfatal, the calculation of YLL can be neglected, and the evaluation of EBD mainly focuses on YLD:

To account for the uncertainty of DALY estimation, a Monte Carlo simulation was also performed and run for 10,000 iterations (Alexander, 2003). The results for disease burden calculation specific to Cases 1 and 2 are shown in Fig. 4. For Case 1, Rotaviruses and sprinkling irrigation are identified as the major risk factor and exposure pathway during water reuse, leading to an average health loss of 5.57×10^{-7} DALYs (95% CI: 4.66×10^{-7}-1.72×10^{-4} DALYs) and 5.12×10^{-7} DALYs (95% CI: 1.95×10^{-7}-1.47×10^{1} DALYs), respectively, and contributing 49.40% and 45.40%, respectively, to the total burden of disease estimates. In Case 2, Norovirus and road flushing are identified as the key risk factor and transmission pathway during water reuse, leading to an average health loss of 1.42×10^{-7} DALYs (95% CI: 7.51×10^{-11}-2.67×10^{-4} DALYs) and 1.62×10^{-7} DALYs (95% CI: 1.16×10^{-7}-2.67×10^{-4} DALYs), respectively, and contributing 71.47% and 81.90%, respectively, to the total burden of disease evaluation. However, as compared with the WHO threshold of 10^{-6} DALYs (WHO, 2011), the disease burden caused by major pathogens through typical pathways in Cases 1 and 2 are both acceptable. Therefore, water reuse in Xi'an Siyuan University and Lake Cui, Kunming is determined to be safe.

Table 6
Disease parameters including disability weight (W), disease duration (T, days) and proportion of a certain health outcome (p, %) for DALY calculation for *E. coli*, *Salmonella*, Rotavirus and Norovirus.

Pathogen	health outcome	Parameter			References
		W	T	p	
E. coli	Watery diarrhoea	0.067	2.5	53%	Murray and Lopez (1996); Havelaar and Melse (2003); Howard and Pedley (2004); Gu et al. (2002)
	Bloody diarrhoea	0.44	6	47%	
Salmonella	Typhoid	0.60	24.5	100%	
Rotavirus	Acute gastroenteritis	0.60	4	100%	
Norovirus	Acute gastroenteritis	0.740	1.5	100%	

3.5. Comparison with conventional model

Conventional models shown as Eq. (12) ~ Eq. (19) for morbidity and DALY calculation is applied in this study as well, and the results are obtained for comparison. In this research, a Beta-Poisson model is adopted for infection rate (P_{inf}) calculation for *E. coli*, *Salmonella,* and Rotavirus, with determined dose-response parameters of 'α' and 'N_{50}' as shown in Table 7 (Rose et al., 1991; Regli et al., 1991; Teunis et al., 2010). For Norovirus, a fractional Poisson model is used, with determined parameters of 'ρ' and 'μ' as shown in Table 7 (Veraga et al., 2016). The probability of illness given infection ($P_{ill/inf}$) is calculated by using Eq. (15) (Teunis et al., 1999), and the parameters of 'γ' and 'σ' for *E. coli*, *Salmonella*, Rotavirus, and Norovirus are determined as shown in Table 8 (Teunis et al., 1999; Gao et al., 2016; Chen et al., 2016). Disease burden is also evaluated in terms of DALY by using Eq. (17)~(19) according to the morbidity results obtained from conventional methods (Murray and Lopez, 1996).

As a result, the average morbidity (P_{ill}) for Cases 1 and 2 estimated by using the conventional model are 2.78×10^{-7} (95% CI: 7.14×10^{-8}-5.42×10^{-7}) and 4.45×10^{-8} (95% CI: 3.03×10^{-8}-5.80×10^{-8}), respectively. The total average disease burden calculated by using the conventional model specific to Cases 1 and 2 are 6.01×10^{-6} DALYs (95% CI: 1.54×10^{-6}-1.17×10^{-5} DALYs) and 2.97×10^{-7} DALYs (95% CI: 2.02×10^{-7}-3.91×10^{-7} DALYs), respectively. A comparison of the

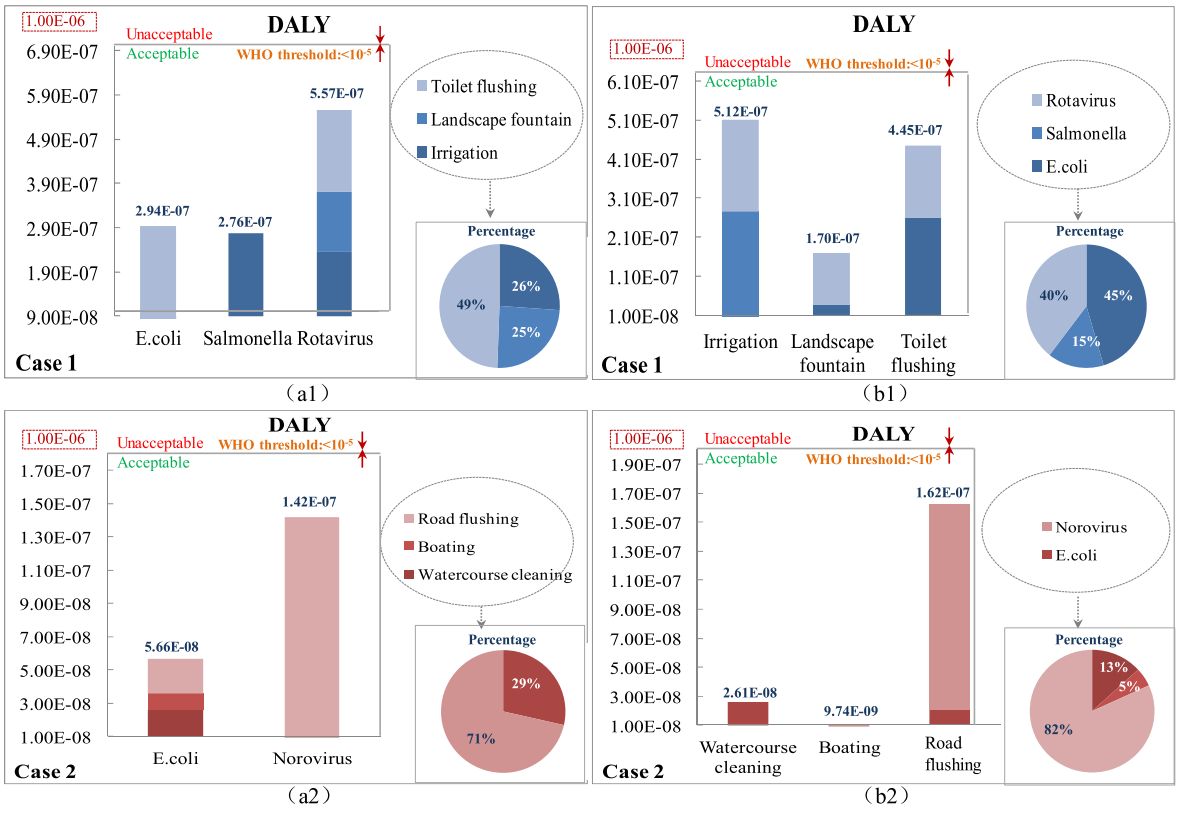


Fig. 4. Results for disease burden calculations (DALYs) for each pathogen (a1, a2)/exposure pathway (b1, b2) in Cases 1 and 2.

Table 7
Dose-infection model (P_{inf}) and corresponding parameters determined for *E. coli*, *Salmonella*, Rotavirus and Norovirus in cases 1 and 2.

Pathogen	Dose-Infection model	Parameter				References
		α	N_{50}	ρ	μ	
E.coli	$P_{inf} = 1 - \left[1 + \frac{D}{N_{50}}\left(\frac{1}{2^{\alpha}} - 1\right)\right]^{-\alpha}$	0.1778	8.6×10^{7}			Rose et al. (1991); Regli et al. (1991); Teunis et al. (1999); Teunis et al. (2010); Gao et al. (2016); Chen et al. (2016)
Salmonella	$P_{inf} = 1 - \left[1 + \frac{D}{N_{50}}\left(\frac{1}{2^{\alpha}} - 1\right)\right]^{-\alpha}$	0.1086	3.6×10^{6}			
Rotavirus	$P_{inf} = 1 - \left[1 + \frac{D}{N_{50}}\left(\frac{1}{2^{\alpha}} - 1\right)\right]^{-\alpha}$	0.2531	6.17			
Norovirus	$P_{inf} = \rho(1 - e^{-D/\mu})$			0.772	1160	

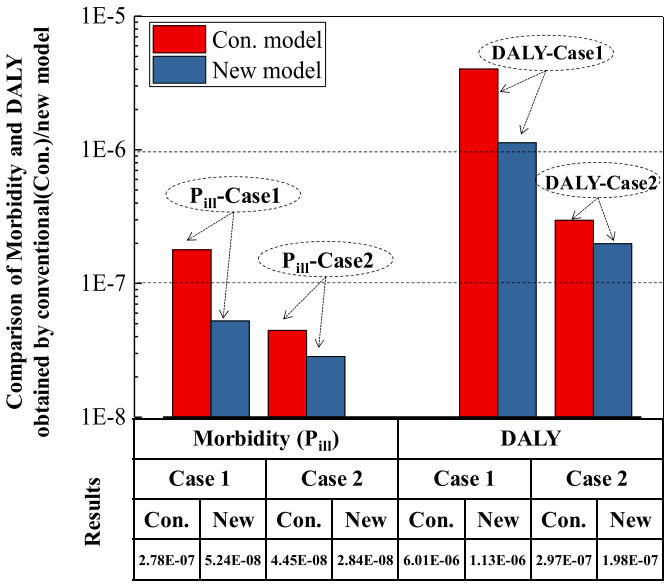
Table 8
Illness-Infection model ($P_{ill/inf}$) and corresponding parameters determined for *E. coli*, *Salmonella*, Rotavirus, and Norovirus in cases 1 and 2.

Pathogen	Illness-Infection model	Parameter		References
		γ	σ	
E.coli	$P_{ill/inf} = 1 - (1 + \gamma D)^{-\sigma}$	2.28×10^{-2}	2.46×10^{-2}	Rose et al. (1991);
Salmonella	$P_{ill/inf} = 1 - (1 + \gamma D)^{-\sigma}$	1.00×10^{-16}	3.40×10^{8}	Regli et al. (1991); Teunis et al. (1999); Teunis et al. (2010);
Rotavirus	$P_{ill/inf} = 1 - (1 + \gamma D)^{-\sigma}$	1.73×10^{-3}	2.46×10^{-3}	Gao et al. (2016);
Norovirus	$P_{ill/inf} = 1 - (1 + \gamma D)^{-\sigma}$	8.73×10^{-4}	9.50×10^{-2}	Chen et al. (2016)

results obtained by the conventional and new models are shown in Fig. 5. From this figure, the morbidity and disease burden calculated by using the conventional model are slightly higher due to the higher estimates of infection rate by using the conventional dose-response model. However, all fall in the same intervals and, thus, similar conclusions for morbidity analysis and EBD evaluation could be drawn. The health hazards caused by typical waterborne pathogens specific to cases 1 and 2 assessed by both methods (conventional or new) are regarded as acceptable. Therefore, water reuse in these two cases is identified as safe. However, since the application of the conventional models is involved with a variety of issues regarding model selection (P_{inf} $P_{ill/inf}$) and parameter determination (α, N_{50}, ρ, μ, γ, σ), results obtained from the conventional method are relatively uncertain. The newly established generalized linear model can be used for direct estimation of disease burdens through the transformation of exposure data to disease data with consideration of certain exposure circumstances and human immunity status. Thus, it can provide a more accurate and objective estimate of the morbidity and can also provide a new perspective for the EBD evaluation.

4. Conclusions

A linear morbidity calculation model was constructed for



Results	Morbidity (P_{ill})				DALY			
	Case 1		Case 2		Case 1		Case 2	
	Con.	New	Con.	New	Con.	New	Con.	New
	2.78E-07	5.24E-08	4.45E-08	2.84E-08	6.01E-06	1.13E-06	2.97E-07	1.98E-07

Fig. 5. Comparison of the results for morbidity and disease burden calculated respectively by conventional (Con.) and new models specific to each case.

transforming pathogen and/or human exposure data to disease data, which is indispensable for the EBD analysis by using a DALY method. Additionally, human testing data for several intestinal pathogens are collected and applied for model fitting and parameter determination. Two practical cases of water reuse are selected for model application. From the results, the following conclusions could be drawn:

1. A generalized linear model (GLM) was studied and firstly used as a template to construct a multivariable linear morbidity model according to the characteristics of the GLM. Three major factors including the exposure dose (D), pathogen toxicity (ID_{50}) and human immunity (IgG), which may lead to the occurrence of certain diseases and affect the tendency of disease development are analysed according to their functional mechanisms and are determined as key independent variables for model construction. For parameter fitting and model validation, human testing data for 8 typical pathogens available from previous studies are collected. The results indicate a good fit of this linear morbidity model with the existing experimental data. Thus, it could be used for predicting the morbidity for typical environmental pathogens under given exposure doses, pathogen toxicity and human immune data directly. The establishment of a linear morbidity model in this research cannot only provide a suitable method for obtaining disease data but also optimize the process of EBD study by establishing a mathematical relationship to transform pathogen and/or human exposure data into disease data.
2. Using the constructed linear morbidity model from this research and the DALY method recommended previously by the WHO, disease burdens for two practical cases of water reuse in Xi'an Siyuan University and Lake Cui, Kunming are evaluated for further risk decision-making. As a result, a major contribution of disease burden is attributed to Rotavirus in case 1 and Norovirus in case 2, which lead to average health losses of 5.57×10^{-7} DALYs (95% CI: 4.66×10^{-7}-1.72×10^{-4} DALYs) and 1.42×10^{-7} DALYs (95% CI: 7.51×10^{-11}-2.67×10^{-4} DALYs), respectively. Besides, sprinkling irrigation and road flushing are determined as major transmission pathways in Cases 1 and 2, which lead to average health losses of 5.12×10^{-7} DALYs (95% CI: 1.95×10^{-7}-1.47×10^{1} DALYs) and 1.62×10^{-7} DALYs (95% CI: 1.16×10^{-7}-2.67×10^{-4} DALYs), respectively. However, comparing with the WHO threshold of 10^{-6} DALYs, the EBD caused by pathogens specific to each case is both far below the restriction level and, thus, is acceptable. Therefore, the safety of water reuse in Xi'an Siyuan University and Lake Cui, Kunming could be guaranteed. By comparison of the results obtained from the conventional method, the newly established generalized linear model cannot only be applied for appropriate estimations of morbidity but also provides new perspectives for the EBD evaluation.

Credit author statement

Tingting Gao: Conceptualization, Methodology, Investigation, Writing, Supervision-Original Draft, Pengcheng Xu: Investigation, Writing, Supervision, Rong Chen: Supervision, Xiaochang C. Wang: Supervision, Mawuli Dzakpasu: Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Natural Science Basic Research Program of Shaanxi Province, China (No. 2020JQ-835), the the Natural Science General Special Scientific Research Program of Shaanxi Provincial Department of Education, China (N0. 21JK0665), the National Natural Science Foundation of China (No. 51308439), and the Program for Innovative Research Team in Shaanxi Province.

References

Alexander, S., 2003. Monte Carlo sampling methods. Handb. Oper. Res. Manag. Sci. 10, 353–425.
Wei, An, Zhang, D.Q., Xiao, S.M., Yu, J.W., Yang, M., 2012. Risk assessment of Giardia in rivers of southern China based on continuous monitoring. J. Environ. Sci. 24, 309–313.
Bitton, G., 2005. Wastewater Microbiology, third ed. John Wiley & Sons, Inc., Hoboken, NJ.
Black, R.E., Levine, M.M., Clements, M.L., Hughes, T.P., Blaser, M.J., 1988. Experimental *Campylobacter jejuni* infection in humans. JID (J. Infect. Dis.) 157 (3), 472–479.
David, K.Blough, Madden, Carolyn W., Hornbrook, Mark C., 1999. Modeling risk using generalized linear models. J. Health Econ. 18 (2), 153–171.
Bollaerts, K., Aerts, M., Faes, C., Grijspeerdt, K., Dewulf, J., Mintiens, K., 2008. Human salmonellosis: estimation of dose-illness from outbreak data. Risk Anal. 28 (2), 427–440.
Chappell, C.L., Okhuysen, P.C., Sterling, C.R., Wang, C., Jakubowski, W., DuPont, H.L., 1999. Infectivity of *Cryptosporidium parvum* in healthy adults with preexisting anti-*C. parvum* serum immunoglobulin G. Am. J. Trop. Med. Hyg. 60 (1), 157–164.
Chen, H.H., Chen, H.W., He, J.T., Liu, F., Shen, Z.L., Han, B., 2006. Health-based risk assessment of contaminated sites: principles and method. Earth Sci. Front. 13 (1), 216–223.
Chen, R., Gao, T.T., Wang, X.C., Zhou, J.H., Xu, L.M., 2016. Health impact assessment of wastewater reuse for replenishing an urban landscape lake by disability adjusted life year. J. Water Reuse Desal. 6, 371–381.
Clemente, M.G., Patton, J.T., Yolken, R., Whitington, P.F., Parashar, U.D., Jiang, B., Raghunathan, T., Schwarz, K.B., 2015. Prevalence of groups A and C rotavirus antibodies in infants with biliary atresia and cholestatic controls. J. Pediatr. 166 (1), 79–84.
Dean, R.B., 1981. Use of log-normal statistics in environmental monitoring. In: Cooper, W.J. (Ed.), Chemistry in Water Reuse. Ann Arbor Science Press, Ann Arbor, MI, USA, 1981.
Dietz, V., Vugia, D., Nelson, R., Wicklund, J., Nadle, J., McCombs, K.G., 2000. Active, multisite, laboratory based surveillance for Cryptosporidium parvum. Am. J. Trop. Med. Hyg. 62 (3), 368–372.
Gao, T.T., Wang, X.C., Chen, R., Ngo, H.H., Guo, W., 2015. Disability adjusted life year (DALY): a useful tool for quantitative assessment of environmental pollution. Sci. Total Environ. 511, 268–287.
Gao, T.T., Chen, R., Wang, X.C., Ngo, H.H., Li, Y.Y., Zhou, J.H., Zhang, L., 2016. Application of disease burden to quantitative assessment of health hazards for a decentralized water reuse system. Sci. Total Environ. 551 (552), 83–91.
Gerba, C.P., Rose, J.B., Singh, S.N., Farrah, S.R., 1985. Waterborne gastroenteritis and viral hepatitis. Crit. Rev. Environ. Sci. Technol. 15 (3), 213–236.
Gu, S., Song, G., Zhou, F., Ji, W., Yang, Q., 2002. The application of DALY for health state evaluation of the resident of Shanghai. J. Prev. Med. Shanghai 55, 143–155.
Gyan, C.S., Hewage, K., Sadiq, R., 2017. Microbial quality of reclaimed water for urban reuses: probabilistic risk-based investigation and recommendations. Sci. Total Environ. 576, 738–751.

Haas, C.N., Rose, J.B., Gerba, C.P., 1999. Quantitative Microbial Risk Assessment. John Wiley & Sons, Toronto.
Havelaar, A.H., Melse, J.M., 2003. Quantifying Public Health Risk in the WHO Guidelines for Drinking-Water Quality: a Burden of Disease Approach. RIVM, Bilthoven, Netherlands. RIVM report 734301022/2003.
Havelaar, A.H., de Hollander, A.E.M., Teunis, P.F.M., Evers, E.G., van Kranen, H.J., Versteegh, F.M., 2000. Balancing the risks and benefits of drinking-water disinfection: disability adjusted life years on the scale. Environ. Health Perspect. 108, 315–321.
He, X.H., Ma, S.H., Li, A.D., Pan, X.C., Chen, Q., Wang, J.F., 2006. Exposure assessment of various reclaimed water uses. Environ. Sci. 27 (9), 1912–1915.
Hornick, R., Greisman, S.E., Woodward, T.E., DuPont, H.L., Hawkins, A.T., Snyder, M.J., 1970. Typhoid fever: pathogenesis and immunologic control. N. Engl. J. Med. 283 (14), 739–746.
Howard, G., Pedley, S., 2004. Assessing the risk to public health from water supply using QMRA. In: Godfrey, S., Howard, G. (Eds.), Health, Institutional, Social and Mapping Programmes to Support WSPs. Leicestershire: WEDC. Loughborough University.
Ishaq, S., Sadiq, R., Farooq, S., Chhipi-Shrestha, G., Hewage, K., 2020. Investigating the public health risks of low impact developments at the residential, neighbourhood, and municipal levels. Sci. Total Environ. 744, 140778.
Levine, M.M., DuPont, H.L., Formal, S.B., Hornick, R.B., Takeuchi, A., Gangarosa, E.J., Snyder, M.J., Libonati, J.P., 1973. Pathogenesis of *Shigella dysenteriae* 1 (shiga) dysentery. JID (J. Infect. Dis.) 127 (3), 261–270.
Levine, M.M., Kaper, J.B., Herrington, D., Losonsky, G., Morris, J.G., Clements, M.L., Black, R.E., Tall, B., Hall, R., 1988. Volunteer studies of deletion mutants of *Vibrio cholerae* O1 prepared by recombinant techniques. Infect. Immun. 56 (1), 161–167.
Mathewson, J.J., Johnson, P.C., DuPont, H.L., Satterwhite, T.K., Winsor, D.K., 1986. Pathogenicity of enteroadherent *Escherichia coli* in adult volunteers. JID (J. Infect. Dis.) 154 (3), 524–527.
McCullagh, P., Nelder, J.A., 1999. Generalized Linear Models, second ed. Chapman and Hill, London.
Menon, V.K., George, S., Aladin, F., Nawaz, S., Sarkar, R., Lopman, B., Gray, J.J., Gomara, M.I., Kang, G., 2013. Comparison of age-stratified seroprevalence of antibodies against norovirus GII in India and the United Kingdom. PloS One 8 (2), 239–256.
Murray, C.J.L., Lopez, A.D., 1996. The Global Burden of Disease. World Health Organization, Geneva.
Namata, H., Aerts, M., Faes, C., Teunis, P., 2008. Model averaging in microbial risk assessment using fractional polynomials. Risk Anal. 28 (4), 891–905.
Pulcini, C., Élisabeth, B.N., Oliver, J.D., Stéphan, H., 2014. The impact of infectious disease specialists on antibiotic prescribing in hospitals. Clin. Microbiol. Infect. 20 (10), 963–972.
Regli, S., Rose, J.B., Haas, C.N., Adams, D., Kuntzer, T., Burger, D., 1991. Modeling risk for pathogens in drinking water. J. AWWA (Am. Water Works Assoc.) 83 (11), 76–84.
Rose, J.B., Sun, G.S., Gerba, C.P., Sinclair, N.A., 1991. Microbial quality and persistence of enteric pathogens in graywater from various household sources. Water Res. 25, 37–42.
Salgot, M., Esther, H., 2006. Integrated Concepts for Reuse of Upgraded Wastewater: WP2-Aquarec Guideline for Quality Standards for Water Reuse in Europe. University of Barcelona, Barcelona.
Scarpello, G.M., Ritelli, D., 2008. Closed form solution of a periodically forced logistic model. Annali dell University di Ferrara 54 (1), 85–94.
Szu, S.C., Hunt, S., Xie, G., Robbins, J.B., Schneerson, R., Gupta, R.K., Zhao, Z., Tan, X., 2013. A human IgG anti-Vi reference for *Salmonella typhi* with weight-based antibody units assigned. Vaccine 31 (15), 1970–1974.
Tao, Q., He, P., Xie, Y., 2003. Analysis of prevalence reasons for typhoid and paratyphoid. J. Chin. Publ. Health 19, 1149–1150.
Teunis, P.F.M., van der Heijden, O.G., van der Giessen, J.W.B., Havelaar, A.H., 1996. The Dose–Response Relation in Human Volunteers for Gastro-Intestinal Pathogens. National Institute of Public Health and the Environment (RIVM), The Netherlands.
Teunis, P.F.M., Nagelkerke, N.J.D., Haas, C.N., 1999. Dose response models for infectious gastroenteritis. Risk Anal. 19 (6), 1251–1260.
Teunis, F.M., Chappell, C.L., Okhuysen, P.C., 2002. Cryptosporidium dose–response studies: variation between hosts. Risk Anal. 22, 475–485.
Teunis, P.F.M., Moe, C.L., Liu, P., Miller, S.E., Lindesmith, L., Baric, R.S., Le Pendu, J., Calderon, R.L., 2008. Norwalk virus: how infectious is it? J. Med. Virol. 80 (8), 1468–1476.
Teunis, P.F.M., Kasuga, F., Fazil, A., Ogden, I.D., Rotariu, O., Strachan, N.J.C., 2010. Dose-response modeling of Salmonella using outbreak data. Int. J. Food Microbiol. 144, 243–249.
United States Environmental Protection Agency, 1997. Exposure Factors Handbook. Office of Research and Development, United States Environmental Protection Agency, Washington, DC.
Vardinon, N., Spirer, Z., Goldhar, J., Kacevman, B., Eylan, E., 1999. Human milk anti-*E. coli* antibodies: relationship to maternal parity. Eur. J. Pediatr. 130 (3), 173–180.
Veraga, G.G.R.V., Rose, J.B., Gin, K.Y.H., 2016. Risk assessment of noroviruses and human adenoviruses in recreational surface waters. Water Res. 103, 276–282.
Verbeke, G., Molenberghs, G., 2000. Linear Mixed Models for Longitudinal Data. Springer, New York.
Ward, R.L., Bernstein, D.I., Young, E.C., Sherwood, J.R., Knowlton, D.R., Schiff, G.M., 1986. Human rotavirus studies in volunteers: determination of infectious dose and serological response to infection. JID (J. Infect. Dis.) 154 (5), 871–880.
WHO, 2002. World Health Report 2002—Reducing Risks, Promoting Healthy Life. World Health Organization, Geneva.
WHO, 2011. Guidelines for Drinking Water Quality, fourth ed. World Health Organization, Geneva.
Xie, X., Hu, H.Y., Guo, M.T., Wu, Q.Y., 2009. Assessment method of the pathogenic microbial exposure caused by aerosolization of reclaimed water. Environ. Sci. 30, 70–74.
Yue, L., Chen, X., 2004. Rate of strong consistency of quasi maximum likelihood estimate in generalized linear models. Sci. China, Ser. A: Mathematics 47 (6), 882–893.
Zambriski, J.A., Nydam, D.V., Wilcox, Z.J., Bowman, D.D., Mohammed, H.O., Liotta, J. L., 2013. *Cryptosporidium parvum*: determination of ID50 and the dose-response relationship in experimentally challenged dairy calves. Vet. Parasitol. 197 (1), 104–112.
Zhou, J.H., Wang, X.C., Ji, Z., Xu, L.M., Yu, Z.Z., 2015. Source identification of bacterial and viral pathogens and their survival/fading in the process of wastewater treatment, reclamation, and environmental reuse. World J. Microbiol. Biotechnol. 31, 109–120.
Zu, S.X., Li, J.F., Barrett, L.J., Fayer, R., Shu, S.Y., McAuliffe, J.F., Roche, J.K., Guerrant, R.L., 1994. Seroepidemiologic study of Cryptosporidium infection in children from rural communities of Anhui, China and Fortaleza, Brazil. Am. J. Trop. Med. Hyg. 51 (1), 1–10.