A mechanistic individual-based two-host interaction model for the transmission of a parasitic disease

Heinzmann, D ; Barbour, A D ; Torgerson, P R

DOI: https://doi.org/10.1142/S1793524511001313

International Journal of Biomathematics

A mechanistic individual-based two-host interaction model for the transmission of a parasitic disease

Dominik Heinzmann

Institute of Mathematics, University of Zurich, Winterthurerstrasse 190
Zurich, 8057, Switzerland
dominik.heinzmann@math.uzh.ch

A.D. Barbour

Institute of Mathematics, University of Zurich, Winterthurerstrasse 190
Zurich, 8057, Switzerland
a.d.barbour@math.uzh.ch

Paul R. Torgerson

School of Veterinary Medicine, Ross University, Basseterre Box 334
St. Kitts, West Indies
ptorgerson@rossvet.edu.kn

Abstract. A mechanistic individual-based model for the infection dynamics of the parasite *Echinococcus granulosus* in a two-host transmission system is proposed. The model describes the individual densities of the parasites in the two host populations. The architecture consists of two sub-processes for the acquisition and severity of infection in the host populations and a superimposed infection contact scheme between the hosts. The parasite dynamics within the host population are modeled using a compound mixed Poisson process for the sheep and a shot-noise process for the dogs. All model parameters are estimated based on available data. The fitted model is then used for simulations of the transmission dynamics between the two hosts to investigate environmental factors and evaluate intervention programs.

Keywords: Shot noise process; individual-based; basic reproduction number; generation time

1. Introduction

In this paper, a individual-based model is presented to describe the infection dynamics of the parasite *Echinococcus granulosus* [6] between its definitive hosts, dogs, and its intermediate hosts, sheep. The parasite causes cystic echinococcosis which is a zoonotic parasitic disease, endemic in many parts of the world [7, 24]. Only a few countries are free of the parasite [6]. The disease has a worldwide distribution and its prevalence in humans and animals changes between different regions. The highest prevalence of the parasite in humans is found in rural areas where people are in close contact with livestock, as for example a prevalence of about 9% in the Central Peruvian Andes [15], 3.8% in Qinghai (China) [29] and 2% in Southern Kazakhstan [25]. The transmission dynamics of the parasite modelled here can be described as follows. Adult worms mature in the intestine of the dog and the infective eggs are released in the feces. Conditional on ingestion of such infective biomass by sheep, hydatid cysts can form in organs such as the liver (where on average $60 - 70\%$ of the total cyst burden of a sheep can be found), lungs and brain. The cyst develops over years in the sheep. Dogs gain access to infective materials by consumption of infected viscera of sheep. In rural areas, sheep are slaughtered on the farms and the offal is accessible to dogs. Dogs can also become infected by scavenging dead sheep in fields.

[17] constructed a model of the *Echinococcus granulosus* life-cycle. They used integrodifferential equations for the mean number of worms in dogs and cysts in sheep. The host populations were divided into classes according to the number of infections they carry in order to focus on acquired immunity in hosts. Animals were assumed to lose infection independently of the parasite burden. The model mechanistically described the prevalence and the development of the mean burden of

parasites in the host, but not the parasite densities. A negative binomial distribution was used to describe the densities in hosts, but without linking it to an underlying infection process or to the animals' ages.

Here, a mechanistic individual-based model describing the development of the worm loads in dogs and cyst burdens in sheep is presented. The architecture is based on the two intra-population models for the host populations introduced in [13, 14] to model the infection dynamics in the two host populations separately. In [13], a compound mixed Poisson process is used to model the acquisition of hydatid cysts in sheep, whereas in [14], a shot noise process is used to model the infection pattern in dogs. Shot noise processes are a classical extension of compound Poisson processes in that they allow for a (deterministic) decrease of the accumulated jumps [5] between consecutive jumps. Such a decrease is useful to model deaths of parasites in a host. The model in this paper links these two models by superimposing a biologically reasonable infection contact pattern between the hosts. This yields an individual-based simulation model for the whole life-cycle of the parasite, with biologically interpretable parameters estimated based on observed data. The fitted model is then used to evaluate the influence of environmental factors and control intervention programs on the transmission cycle. The simulation experiments indicate that the model provides a valuable tool for investigating the dynamics initiated by natural or man-made changes to the life-cycle of *Echinococcus granulosus*.

2. Intra-population models

In this section, the intra-population models for the sheep and dog populations proposed in [13, 14] are briefly recalled. These models are an important part of the architecture of the final model for the whole life-cycle of the parasite. The sub-models are based on a clumped infection mechanism. It is assumed in both models

that the transmission cycle of *Echinococcus granulosus* is endemically stable, so that the ingested parasite clumps are identically distributed over time. Endemicity of the transmission cycle is suggested by various studies [10, 27, 21, 22, 26, 30]. Furthermore, a low incidence rate in hosts is also assumed in both models so that the acquisition process can be described by a (mixed) Poisson process. A low infection rate of sheep is suggested by [3], [10] and [26], and a low infection rate of dogs is suggested by [9], [17], [22] and [24]. Finally, the clump sizes are treated as independent in the models, making compound processes for sheep and shot noise processes for dogs reasonable models.

2.1. *Infection dynamics in sheep population*

In [13], it is shown that a compound mixed Poisson process with a zero-truncated negative binomial distribution for the number of established cysts per ingested clump of infective material provides an adequate fit for the age-dependent cyst distribution in sheep. The model indicates that the rate of ingestion of clumps is heterogeneous within the sheep population, and that the clump sizes are aggregated.

Let the random variable Y_t be the total number of cysts established in an individual up to time t and let V_k $(k = 1, 2, \ldots)$ be independent random variables describing the numbers of cysts acquired at a single infection, with common distribution $\mathcal{Q}$. $\mathcal{Q}$ is assumed to be the zero-truncated version of a negative binomial distribution with parameters θ and ζ, whose mean and variance are given by

$$\mathbb{E}(V_k) = \frac{\theta\zeta}{1-(1/(\zeta+1))^{\theta}}$$

and

$$\mathrm{Var}(V_k) = \left[\frac{\theta\zeta(1+\zeta+\theta\zeta)}{1-(1/(\zeta+1))^{\theta}} - \left(\frac{\theta\zeta}{1-(1/(\zeta+1))^{\theta}}\right)^2\right].$$

Let N_t be a negative binomial random variable with mean $\psi\xi t$ and variance $\psi\xi t(1+\xi t)$, where ψ is a shape and ξ a scale parameter, and denote the distribution of Y_t by $\mathcal{P}_t$. Then the model presumes that

$$Y_t = \sum_{k=1}^{N_t} V_k \quad with \quad \mathbb{P}(N_t = m) = \frac{\Gamma(\psi+m)}{\Gamma(\psi)m!}\left(\frac{1}{t\xi+1}\right)^{\psi}\left(\frac{t\xi}{t\xi+1}\right)^{m},$$

so that

$$\mathcal{P}_t = \sum_{m=0}^{\infty} \mathbb{P}(N_t = m)\mathcal{Q}^{\star m}\,, \tag{2.1}$$

where $\mathcal{Q}^{\star m}$ is the mth convolution of $\mathcal{Q}$ and N_t is independent of the V_k's. In [13], a Poisson process as alternative for the counting process N_t is also considered, but it is shown that it is inferior to the above specification of N_t. The heterogeneity of acquisition of clumped infections proposed by model (2.1) may result from behavioral differences of sheep on pasture, or from differences in the immune system of sheep.

The model was fitted to a sheep sample from Kazakhstan [26], containing 2505 individual records of the variables age and hydatid cyst burden in sheep. The resulting parameter estimates are represented in Table 1. The estimates imply that a sheep ingests an infectious clump roughly every $1/\hat{\psi}\hat{\xi} \approx 3$ years, each clump leading on average to about 4 established cysts.

2.2. *Infection dynamics in dog population*

[14] showed that shot noise processes provide a good fit to the age-dependent prevalences and the distribution of *Echinococcus granulosus* in dogs.

Let the random variable X_t denote the total number of parasites carried by a dog at time t. $(X_t)_{t\geq 0}$ is modeled as a continuous-time stochastic process. Infection events occur at the times $0 < \tau_1 < \tau_2 \cdots$ of a Poisson process with rate β. At each

Table 1. Maximum likelihood estimates of the parameters of the sub-process (2.1) for the sheep population and of the PT respectively CS sub-process for the dog population from the corresponding Kazakhstan samples, together with 95% confidence intervals computed by the bootstrap percentile method. Note that " − " for the dog models indicates that the corresponding parameter is not specified in that model.

Sheep	Sub-process	
$\hat{\psi}$	$0.941\,(0.629, 1.260)$	
$\hat{\xi}$	$0.343\,(0.225, 0.741)$	
$\hat{\theta}$	$0.351\,(0.139, 0.617)$	
$\hat{\zeta}$	$5.859\,(3.215, 9.763)$	
Dogs	PT-subprocess	CS-subprocess
$\hat{\beta}$	$0.445\,(0.317, 0.918)$	$0.340\,(0.213, 0.881)$
$\hat{\mu}$	$6.001\,(4.305, 7.054)$	$4.302\,(3.723, 4.928)$
$\hat{\sigma}$	$2.955\,(2.437, 3.306)$	$2.616\,(2.182, 2.882)$
$\hat{\lambda}$	$8.833\,(6.319, 13.176)$	−
$\hat{t}_d$	−	$0.744\,(0.580, 1.108)$

τ_i, the value of X_t increases by a random amount, assumed to have a log-normal distribution $\mathrm{LN}(\mu, \sigma^2)$, so that the mean increase is $\exp(\mu + \sigma^2/2)$. Thereafter, the amount declines according to a predetermined scheme. Thus

$$X_t = \sum_{k=1}^{M_t} U_k h(t - \tau_k)\,, \quad t \geq 0\,, \tag{2.2}$$

where M_t is a Poisson random variable with mean βt, $U_k \sim \mathrm{LN}(\mu, \sigma^2)$ $(k = 1, 2, \ldots)$ are independent, and $h(t)$, $t \geq 0$, denotes the proportion of parasites still surviving t time units after infection. We take $h(t) = 0$ for $t < 0$.

There is experimental evidence that dogs eventually completely lose their infection [1, 6, 10]; this should also be reflected in the model. The natural choice $h(t) = \exp(-\lambda t)$, for some $\lambda > 0$, is not of this form. For this version, we prefer a more faithful model, in which U_k is replaced by $\tilde{U}_k \sim \mathrm{Po}(U_k)$, and each of the $\tilde{U}_k$ parasites has an independent, exponentially distributed lifetime with mean $1/\lambda$. Then an infection load of m parasites disappears completely after a random time of

mean $(\gamma+\log m)/\lambda$, where γ is the Euler-Mascheroni constant. For $U_k \sim \mathrm{LN}(\mu, \sigma^2)$, this gives a mean survival time of a single infection load of about $(\gamma + \mu)/\lambda$. The resulting distribution of X_t is mixed Poisson. This model is referred to as the PT (Poisson transform) model.

An alternative is offered by the CS (constant survival) model, where

$$h(t) = \begin{cases} 1 & \text{if } t \leq t_d \\ 0 & \text{if } t > t_d \,, \end{cases}$$

with t_d the (fixed) duration of infection in a dog. Here, there is no decay of the burden between infection and complete loss at t_d time units later.

The PT and CS models were fitted to a dog sample from Kazakhstan [22] containing 606 individual records of the variables age and parasite counts in dogs. Note that the study leading to the dog sample was conducted at the same year (2000) and in the same region as the study for the sheep sample above. The resulting estimates from the dog sample are given in Table (1). The results suggest that the infection rate is $0.35 - 0.45$ per dog per year, indicating that a dog is exposed to infection on average once every 2.5 years, with a mean load of several thousands. The fitted models also suggest that a single clumped infection survives for about 9 months.

3. Inter-population models

This section introduces the sub-models describing the infection contact pattern between sheep and dogs.

3.1. *Infection of dogs by sheep*

An infection of a dog takes place if an infective sheep dies. A sheep is infective if it contains cysts harboring protoscoleces. A protoscolex is a larval stage of the parasite

formed during asexual budding processes inside the cysts inside the intermediate hosts ([6]). Hence a model for the fertility (presence of protoscoleces) in cysts is required to describe the infection of dogs by sheep.

Let $k(t)$ be the probability that a cyst at age t has formed protoscoleces and thus is fertile. Assume that once a cyst is fertile, it remains so. Since $k(0) = 0$ and the shape of $k(t)$ should allow a flexible fit, a reasonable choice for $k(t)$ is the Bass model [2]:

$$k(t) = k^* \frac{1 - e^{-(a+b)t}}{1 + \frac{b}{a} e^{-(a+b)t}}, \tag{3.1}$$

where k^* is the asymptotic probability of fertility and a and b are adjustable coefficients. Equation (3.1) allows the modeling of S-shaped curves.

The data used to fit the model (see below) contain records of the number of fertile and non-fertile cysts for a sheep at age t, but not the ages of the cysts themselves. Thus (3.1) is a latent process and needs to be coupled to the underlying mixed Poisson infection process of (2.1) in order to compare it to the data. Each animal has a fixed infection rate, and the acquisition process of cysts is Poisson. Let $q(t)$ denote the probability that a cyst is fertile in a sheep at age t. Thus the following equation for $q(t)$ is appropriate to model the fertility in cysts:

$$\begin{aligned} q(t) &= \int_0^t k(s) \frac{1}{t} ds = \frac{k^*}{t} \int_0^t \frac{1 - e^{-(a+b)s}}{1 + \frac{b}{a} e^{-(a+b)s}} ds \\ &= \frac{k^*}{t} \left[t + \frac{1}{b} \left\{ \log(1 + \frac{b}{a} e^{-(a+b)t}) - \log(1 + \frac{b}{a}) \right\} \right]. \end{aligned} \tag{3.2}$$

The fertility model is fitted to an unpublished data set from Kyrgyzstan, collected in 2007. A total of 1081 sheep slaughtered in an abattoir were examined for *Echinococcus granulosus* cysts. The cysts were investigated for the presence of protoscoleces. Table 2 summarizes the observed counts and the resulting propor-

Table 2. Number of cysts n_1, number of fertile cysts n_2 and the resulting observed proportion of fertile cysts $q(t)$ (with 95% binomial confidence interval) for different ages. Ages were recorded as integers in the sample with 6 years the maximum.

Age	n_1	n_2	$q(t)$
1	529	5	0.009 (0.003, 0.022)
2	810	14	0.017 (0.009, 0.029)
3	662	34	0.051 (0.036, 0.071)
4	783	44	0.056 (0.041, 0.075)
5	617	38	0.062 (0.044, 0.084)
6	490	43	0.088 (0.064, 0.116)
overall	3891	178	0.046 (0.039, 0.053)

tions of cysts having protoscoleces for different age classes of the sheep. The overall prevalence of infection is 0.046, indicating a low fertility of cysts in general.

The maximum likelihood estimates of the parameters in (3.2) are $\hat{k}^* = 0.103$ (95% bootstrap confidence interval: $(0.079, 0.121)$), $\hat{a} = 0.124$ $(0.101, 0.147)$ and $\hat{b} = 1.394$ $(1.204, 2.426)$. Figure 1 displays the resulting latent model (3.1) and the mixture model (3.2). The latent model suggests that a cyst has a chance of about 6% to be fertile at cyst age 2 and a chance of 10% at age 4. Between $1-3$ years, cyst fertility increases approximately linearly with age. The mixture model (3.2) suggests that the asymptotic proportion of fertile cysts in older sheep is approximately 10%, indicating that one cyst out of ten is on average fertile. Figure 1(b) shows that the mixture model is in reasonable agreement with the observed proportions of cyst fertility in different age classes.

The final infection of dogs by sheep is thus as follows. A sheep dies (or is slaughtered) at age t, having a number of cysts n_c, specified by its infection history. The cadaver, ingested by a single (randomly selected) dog, is considered to be infective if at least one cyst is fertile. Thus the probability of infectiousness of the sheep is $1 - [1 - q(t)]^{n_c}$, where $q(t)$ is as in (3.2). If the sheep cadaver is infective, the dog

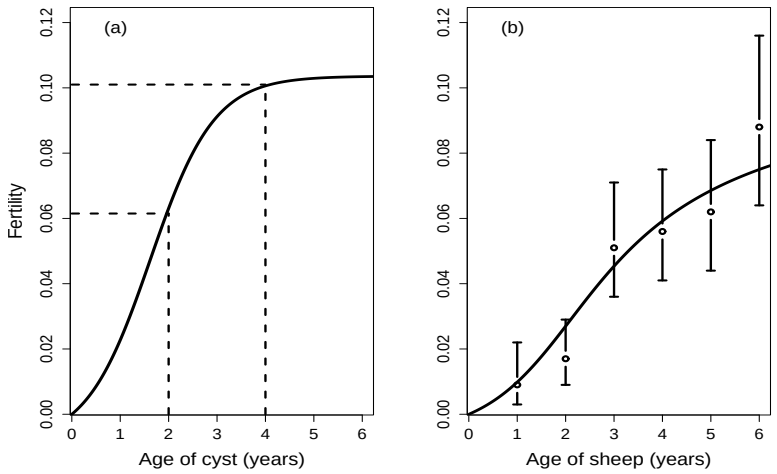

Fig. 1. (a) Estimated age-dependent fertility equation (3.1) for *Echinococcus granulosus* cysts. (b) The resulting fertility of cysts in function of the age of the sheep based on the mixture model (3.2), together with the observed prevalences of infection given in Table 2.

eating it gets infected. In the regions of Kazakhstan, where the samples of this paper where collected, slaughtering is often done at home, and in most cases households have only a single dog (personal observations from the authors). Hence the above assumption of a single dog getting infected per infectious sheep is reasonable. Even if more than one dog eat from that sheep cadaver, the low cyst burden per sheep and the small fertility probability of a cyst imply that in most cases, only a single cyst is fertile, thus leading to an infection of one of the dogs.

3.2. *Infection of sheep by dogs*

Sheep catch an infection if they make contact with excreta of infective dogs, i.e. excreta containing parasite eggs. There are no experimental studies for *Echinococcus granulosus* available which appropriately investigate the parasite burden in dogs and the resulting egg output in excreta. However, experiments with *Echinococcus multilocularis*, a comparable parasite, suggest that the total number of eggs in feces

of infected dogs are not correlated with the number of worms in the dog. Hence it is assumed that all dogs are equally infectious, independent of their parasite burden, and the following model is used to describe the infection of sheep by dogs.

Let η be a random variable for the rate of potentially infectious contacts of the sheep with excreta of dogs. Let δ denote the prevalence of infection in dogs at time t. Since only a fraction δ of the contacts of sheep are with excreta of infected dogs, the infection rate of sheep is $\eta\delta$. The sub-process for the sheep population (2.1) indicates that the infection rate $\eta\delta$ has a gamma distribution with parameters ψ and ξ. Hence the contact rate η has a gamma distribution with the same shape parameter ψ, but different scale parameter $\tilde{\xi} = \xi/\delta$. Hence each sheep obtains at birth an individual contact rate η (drawn from the corresponding gamma distribution) which stays constant over its life, so that the resulting infection rate at time t is $\eta\delta$, specifying its potential to get infected. Thus given the prevalence in dogs δ at time t, the probability of infection through excreta from infective dogs is heterogeneous in the sheep population.

4. Final model for the whole transmission cycle

4.1. *Simulation model*

The final individual-based simulation model consists of simply combining the intra-population models introduced in Section 2 and the inter-population models presented in Section 3. All model parameters are given and fixed in Table 3. The fixed values for the parameters are discussed in Subsection 4.3.

Starting with $n^{(1)}$ dogs and $n^{(2)}$ sheep, the host populations are initialized by attributing an actual age with corresponding load and a remaining life duration to all animals. For each dog, a lifespan is generated by using an exponential distribution with mean r years and the corresponding load is computed based on the PT or CS

Table 3. Summary of all parameters with fixed values needed to simulate the mechanistic two-host model. Note that " − " for the dog models indicates that the corresponding parameter is not specified in that model. The infection rate β is only used for the initialization of the dog population.

Dogs	PT	CS	Explanation
β	0.445	0.340	Parameter of the PT/CS sub-processes for dogs
μ	6.001	4.302	" "
σ	2.955	2.616	" "
λ	8.833	−	Parameter of the PT sub-process for dogs
t_d	−	0.744	Parameter of the CS sub-process for dogs
Sheep			
ψ	0.941		Parameter for the gamma mixture of contact rates
$\tilde{\xi}$	1.491		" "
θ	0.351		Parameter of the clump size distribution
ζ	5.859		" "
Fertile			
k^*	0.103		Parameter in the fertility model (3.2)
a	0.124		" "
b	1.394		" "
Model			
$n^{(1)}$	3030		Population size dogs
ρ	10.7		$\rho = n^{(2)}/n^{(1)}$, so that population size sheep $n^{(2)} = 32421$
r	3		Sample mean age of dogs

processes. To obtain an appropriate starting distribution for the ages of the sheep population, we note that the age at death distribution of sheep is length-biased [18]. The lifetime L is thus sampled from the empirical age at death distribution, weighted in proportion to lifetime. The age of the sheep is determined by a realization of a $U[0, L]$-distribution, and a load for the sheep is then computed based on the process (2.1). This provides a satisfactory initialization so that equilibrium can be reached reasonably quickly. The transmission is assumed to take place in a homogeneous, homogeneously mixing closed community.

The simulation of the full model given the above initializations is carried out

using the tau-leaping method (Gillespie 2001), an approximation to the Gillespie algorithm [12] to speed up simulation. The choice of τ is discussed in the Subsection 4.3. Given a subinterval of length τ, the expected number of death and infection events assuming constant rates over the interval is determined and executed and the state variables of the system such as individual ages and loads are updated.

After each time step τ, the number of death events in both host populations is computed and the dead animals are removed and replaced by uninfected newborns with age 0 and load 0. The ages of the remaining animals are updated. The loads of dogs still alive are adjusted accordingly to the decay process specified with the PT respectively the CS model. Note that for the PT model, a dog having at time t j parasites will lose on average $j(1-\exp(-\lambda\tau))$ during the τ. Thus for the simulation with the PT model, we assume that parasites die independently with probability $1-\exp(-\lambda\tau)$ during τ. The number k of new infections in the dog population is then determined based on presence and fertility of cysts in the dead sheep as described in Subsection 3.1, and k dogs are randomly selected from the $n^{(1)}$ dogs in the population. A realization of $\mathrm{LN}(\mu,\sigma^2)$ is attributed to each of those k dogs. Note that if the selected dog is already infected, then the additional worm load is additive to the existing load at the time of infection.

Sheep dying during τ are replaced by newborns, for which an individual contact rate η is drawn from a gamma distribution with shape parameter ψ and scale parameter $\tilde{\xi}$ as seen in the previous Section. The prevalence of infection in dogs δ is evaluated, and the number of infectious events per sheep experienced during τ is given as a realization l of $\mathrm{Po}(\eta\delta\tau)$. The corresponding sheep is then infected l times with a number of cysts specified by a realization of a zero-truncated version of a negative binomial random variable with shape parameter θ and scale parameter ζ.

4.2. *Basic reproduction number*

Assume that we have a single infected animal in an otherwise fully susceptible dog population of size $n^{(1)}$, and that all $n^{(2)}$ sheep are uninfected and susceptible. The infected dog infects sheep with cysts at a mean rate of $\varphi := \psi\tilde{\xi}\rho$, where $\rho = n^{(2)}/n^{(1)}$. A sheep is infectious, and then transmits the disease to exactly one dog, only if it has fertile cysts. Let ω denote the mean proportion of sheep that are infectious conditional on having cysts. Given model (3.2) for the fertility of a cyst in a sheep of age t and model (2.1) describing the cyst load of a sheep at age t, the mean fertility of a sheep at age t having cysts becomes $\omega_t = \sum_{j=1}^{\infty} \mathbb{P}(Y_t = j)[1-(1-q(t))^j]$, where $\mathbb{P}(Y_t = j)$ is specified by $\mathcal{P}_t$ in (2.1). If $t_1, \ldots, t_n$ are the ages of the sheep in the sample, then ω can be defined as $\omega := (1/n)\sum_{l=1}^{n} \omega_{t_l}$.

Hence on average $\varphi\omega$ infections will arise in the dog population, per unit duration of the dog's infectious period. Let α_1 be the mean loss rate of infection in dogs, which is the sum of the natural death rate of dogs and a rate of loss of infection through death of parasites. For the PT model, we have seen that the mean survival time of a single infection is given as $\mathbb{E}(T) = (\gamma+\mu)/\lambda$ with γ the Euler-Mascheroni constant, and for the CS model, the duration is fixed with t_d. Given that the dogs have an exponential life time with mean 3 years, we have $\alpha_1 = 1/3 + \lambda/(\gamma+\mu)$ for the PT model and $\alpha_1 = 1/3 + 1/t_d$ for the CS model. Then a single infected dog indirectly infects on average a total of $\varphi\omega/\alpha_1$ dogs if all sheep are susceptible.

Theorem 4.1. *The basic reproduction number*

$$R_0 := \frac{\varphi\omega}{\alpha_1},$$

with parameters as above, is a threshold such that, as $t \to \infty$, $R_0 < 1$ implies that the disease dies out and $R_0 > 1$ implies that the disease may persist.

The threshold R_0 can be computed as follows. Fix ψ, $\tilde{\xi}$ and ρ as in Table 3. As seen in Section 2, the mean survival time of a single infection in dogs is comparable in the PT and CS models, with a common value of about 0.75 years. Thus α_1 is fixed to $1/3+1/0.75 \approx 1.65$. And finally, ω is fixed as follows. Let the model parameters for (3.2) and (2.1) be fixed by their estimates as before. For each age t in the Kazakhstan sample, we can approximate ω_t by $\omega_t(m) := \sum_{j=1}^{m} \mathbb{P}(Y_t = j)[1-(1-q(t))^j]$, where m is chosen such that $\omega_t(m) - \omega_t(m-1) < 10^{-8}$. Given the ages of the sheep $t_1, \ldots, t_n$, we have $\omega = (1/n)\sum_{l=1}^{n} \omega_{t_l}(m_l)$, where m_l is the value of m computed for age t_l. Here, $\omega = 0.198$, which is close to the corresponding empirical value of 0.205, where instead of the theoretical distribution for the number of cysts the observed distribution is used. Thus we obtain $R_0 = 1.806$.

As a marginal note, subdividing the ages of the sheep sample into age classes $(0,1]$, $(1,2]$, $(2,3]$, $(3,4]$, $(4,5]$ and 5+, where the class 5+ summarizes all sheep older than 5 years, and computing ω, the mean infectiousness conditional on having cysts, for all of the classes separately results in 0.051, 0.155, 0.268, 0.354, 0.414 and 0.455. This indicates that ω restricted to young sheep ≤ 1 year is about 5%, whereas it is about 45% in sheep older than 5 years.

4.3. *Simulation specifications*

The fixed values for the parameters in Table 3 are determined as follows. For the intra-population model for the dog population, the corresponding parameters are fixed by their estimates given in Table 1. Note that the parameter β is only used to determine a starting distribution for the simulations. After initialization, the infection rate is determined by the number of infective dying sheep. For the intra-population model of the sheep population, the parameters ψ, θ and ζ are fixed by their estimates given in Table 1, and $\tilde{\xi}$ is fixed as $\hat{\xi}/0.230 = 0.343/0.230 = 1.491$ by

noting that 0.230 is the prevalence of infection in dogs of the Kazakhstan sample. The parameters for the age-dependent cyst fertility model for the infection of dogs by sheep, are fixed by their estimates from the Kyrgyzstan sample. Finally, the remaining model parameters are fixed as follows. Assuming that for the dog sample of Kazakhstan, approximately every fifth dog was sampled (rough estimate of participants of that study), we set the constant population size of dogs $n^{(1)} = 3030$, which corresponds to 5 times the sample size of the dog sample from Kazakhstan. The population ratio ρ can be approximated by 10.7, based on field data in Kazakhstan, where $\rho = 10.368$ $(95\%CI : 10.074, 10.706)$ (sample from 1 village; unpublished data), and in Kyrgyzstan, where $\rho = 11.418$ $(95\%CI : 10.593, 12.383)$ (samples from 3 villages; unpublished data), and where during a purgation study in dogs, owners were asked how many sheep and dogs they own. Hence the constant population size of sheep can be set to $n^{(2)} = \rho n^{(1)} = 32421$. And finally, r is fixed by the estimated mean age of 3 years from the dog sample from Kazakhstan.

A choice of 0.01 for τ in the simulation algorithm is reasonable since during τ, the net change of δ, the prevalence of infection in dogs, is about $\tau|\beta - \alpha_1|$, with β the infection rate of dogs approximated by 0.4 per dog per year as mean value of the infection rates β of the PT and CS model given in Table 1, and α_1 approximated by 1.65 as before. Thus the net change is about 1.2%, ensuring that the infection pressure towards sheep does not change greatly during τ. In addition, the mean loads of dogs will decrease during τ on average by a factor of $\exp(-0.01\lambda) = 0.915$, where λ is fixed by its estimate in Table 1. This is sufficiently accurate for our application since all dogs harboring parasites are considered to be equally infectious.

For the simulation, a suitable burn-in period for the system to reach its endemic equilibrium state is determined such that the relative change of the mean of any of the quantities of interest below averaged over consecutive intervals of size 500 time

steps τ (which corresponds to 5 years simulation time) is less than 0.005. For the present application, the burn-in period is fixed at 20000 steps.

5. Application

In this section, the simulation model for the complete transmission cycle of the parasite is used to investigate (a) steady-state behavior, (b) environmental influences, and (c) intervention programs.

5.1. *Steady-state transmission dynamics*

The quantities investigated in this simulation study for the host populations are prevalences of infection δ in dogs and s in sheep, the per capita contact rates $\kappa^{(1)}$ of dogs with sheep viscera and the per capita contact rate $\kappa^{(2)}$ of sheep with excreta from dogs. The contact rate $\kappa^{(1)}$ of dogs is computed at each time step τ in the simulation model as the total number of sheep dying divided by $n^{(1)}\tau$. Similarly, $\kappa^{(2)}$ is computed at each time step τ as the total number of contacts of sheep with excreta of dogs divided by $n^{(2)}\tau$. Hence the above quantities can be determined at each time step τ in the model. We sample their values at every 2000th time unit τ after the burn-in period until we have 1000 values. The average values of the quantities of interest are computed together with the 2.5% and 97.5% quantiles to obtain an idea about their variation.

The results can then be compared to their corresponding estimates derived from the sheep and dog samples from Kazakhstan. In what follows, we refer to these estimates as the true values. The true values of the prevalences of infection δ and s are computed from the corresponding samples, resulting in 0.230 for dogs and 0.363 for sheep. The true value of $\kappa^{(2)}$ is given by $\psi\tilde{\xi}$, where $\tilde{\xi} = \xi/0.230$, and ψ and ξ are fixed by their estimates given in Table 3. Finally, the true value of the

Table 4. Prevalences of infection δ in dogs and s in sheep, and the contact rates $\kappa^{(1)}$ of dogs with viscera from sheep, and $\kappa^{(2)}$ of sheep with excreta from dogs, obtained in the simulation model, together with the observed values, as described in Subsection 5.1. Note that there are two true values of $\kappa^{(1)}$ for either the PT (value=6.191) or the CS model (value=4.731).

	Observed	Simulation/PT	Simulation/CS
δ	0.230	0.232 (0.211, 0.255)	0.252 (0.235, 0.274)
s	0.363	0.346 (0.322, 0.368)	0.366 (0.341, 0.379)
$\kappa^{(1)}$	6.191/4.731	5.986 (3.408, 9.737)	6.099 (3.833, 10.127)
$\kappa^{(2)}$	1.348	1.384 (1.096, 1.739)	1.277 (1.069, 1.684)

contact rate $\kappa^{(1)}$ of dogs is given by $\beta/s\omega$, where β is the infection rate β for the PT respectively the CS model as discussed in Section 2, and ω is the average probability that a sheep harboring cysts is infectious, as discussed in the previous Section and computed as 0.198. Hence using the estimates of β given in Table 3, it follows that $\kappa^{(1)}$ is 6.191 in the PT setting and 4.731 in the CS setting.

Table 4 shows the above quantities of interest of the interaction model with the PT respectively CS sub-model for the dog population, together with the true values introduced above. The values of the interaction model with the PT sub-process are in line with the true values. The computed contact rate $\kappa^{(1)}$ of dogs with viscera from sheep in the simulation model is with 5.986 close to the true value 6.191 in Table 1. In the model with the CS sub-process, the prevalence of infection in the dog population is not well reflected, and the contact rate $\kappa^{(1)}$ of 6.099 is much larger than the corresponding true value 4.731 in Table 1. Thus we will use the interaction model with the PT sub-process for the dog population for investigating the influence of environmental factors and control interventions on the dynamics of the life-cycle of *Echinococcus granulosus*. Figure 2 shows the positive cyst burdens in sheep for different age classes, obtained from a snapshot of the simulation model after the burn-in period, together with the observed positive burdens in the sheep sample

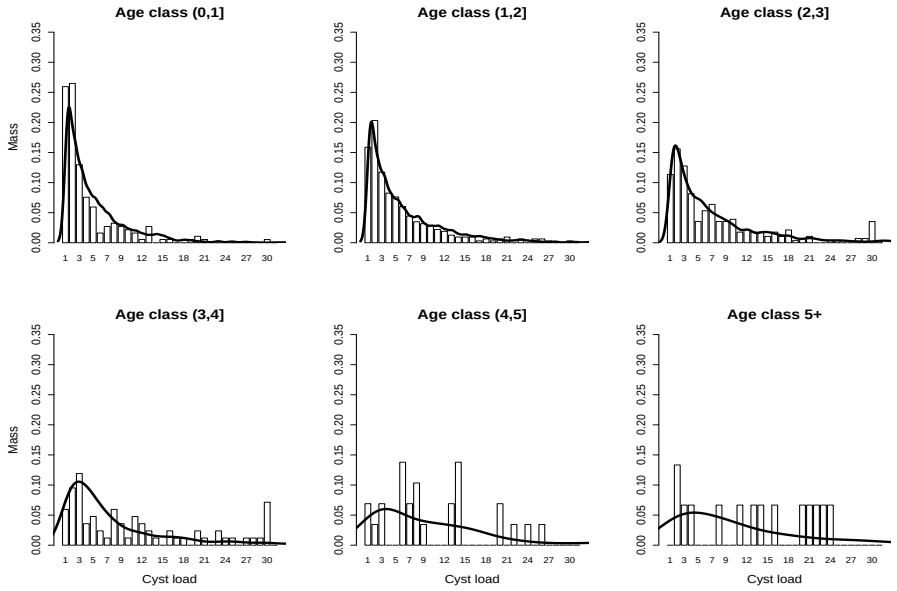

Fig. 2. Snapshot of the distribution of positive burdens of *Echinococcus granulosus* cysts in sheep from the simulation model (solid curve; computed by a kernel density estimation) with a histogram of the observed positive burdens of the sheep sample from Kazakhstan for different age classes. The age classification is taken from [13], with 5+ the age class summarizing all sheep older than 5 years.

from Kazakhstan. For simplification, we represent the positive cyst counts from the simulation output through a kernel density estimator. Analogously, Figure 3 represents a snapshot for the log-transformed positive counts of dogs together with a histogram of the corresponding observed quantities from the dog sample from Kazakhstan. The outputs are in line with the observed counts for sheep and dogs.

5.2. *Environmental influence: Seasonality*

Eggs released by dogs are subject to environmental effects. Climate and temperature are density-independent constraints limiting the survival of eggs [20]. Infectious eggs have been found in water and damp sand for 20 days at $30°C$, 32 days at $10-21°C$. and 225 days at $6°C$ [20]. The eggs survive in general for only short periods if they are exposed to direct sunlight and dry conditions. [19] have shown that the survival

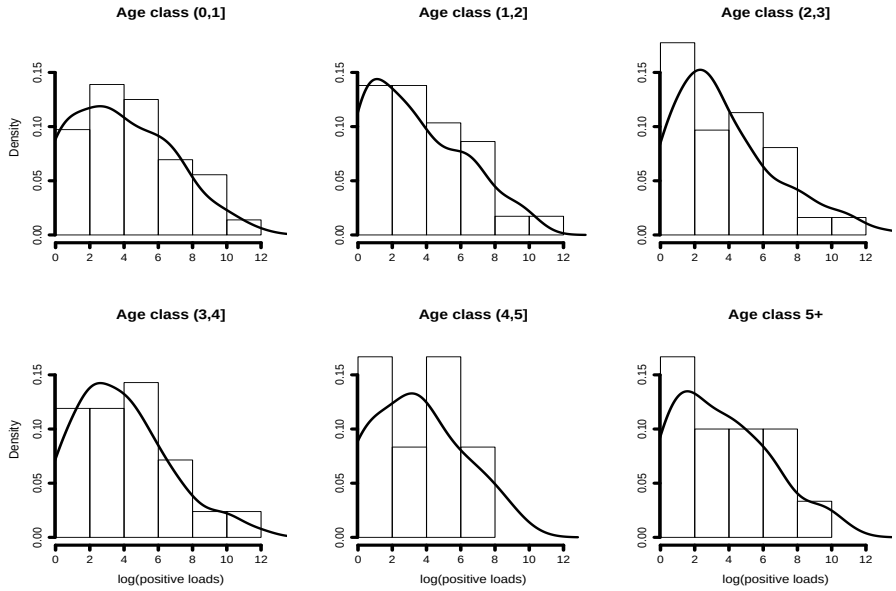

Fig. 3. Snapshot of the distribution of the log-transformed positive parasite loads in dogs from the simulation model (solid curve; computed by a kernel density estimation) with a histogram of the observed loads of the dog sample from Kazakhstan for different age classes, analogously to Figure 2 for sheep.

of infectious *Echinococcus granulosus* eggs can reach up to 41 months in nature under varying temperatures ranging from $-3°C$ to $38°C$. [28] suggest that the survival for *Echinococcus multilocularis* eggs, which are comparable to *Echinococcus granulosus* eggs, is of the order of 100 days. In Kazakhstan, the maximum average monthly temperature is accounted in July at $25°C$ and the lowest in January at $-6°C$. From mid June to late August, the temperatures are approximately $20°C$. There is an average monthly rainfall of $35mm$, ranging from $26mm$ around July and $42mm$ around December.

This seasonality is implemented in the interaction model as follows. Given the individual contact rates η for each sheep, its infection rate is defined as $\eta\delta$, where δ is the prevalence of infection in dogs. Under the above conditions, the survival time of eggs in the 2.5 months from mid June to late August is with about 32 days

approximately one third as large as the survival time in the remainder months, which is about 100 days [6]. Hence we choose the infection rate for a sheep to be equal to $\eta\delta r$ for mid June to late August and equal to $\eta\delta r/3$ for the remainder of the year, where r is chosen such that $(9.5[\eta\delta r]+2.5[\eta\delta r/3])/12 = \eta\delta$, indicating that the mean infection rate over the year of that sheep is still $\eta\delta$. This yields $r = 1.161$. Note that we did not consider variation of temperatures in these periods which may be as important than average temperatures [16].

The simulation with the implemented seasonality effect yields prevalences of infection of $\delta = 0.233$ $(0.211, 0.256)$ in dogs and $s = 0.349$ $(0.326, 0.371)$ in sheep. The contact rates are $\kappa^{(1)} = 6.102$ $(3.632, 9.825)$ and $\kappa^{(2)} = 1.401$ $(1.124, 1.773)$. These are all close to the simulation values as above. This indicates that seasonality does not greatly influence the transmission dynamics, the sheep population acting as buffer in the transmission.

5.3. *Control interventions in dog population*

To develop and evaluate public health control interventions against *Echinococcus granulosus*, it is imperative to understand the reaction of the transmission system on the interventions. Mass dog treatment programs are widely applied in practice to control or eradicate *Echinococcus granulosus* [8, 11, 20, 23].

Based on our model, two scenarios are tested, both are based on a treatment of a certain proportion of dogs with an anti-parasitic drug every 6 weeks. The 6-week interval is based on the prepatent period of infection with *Echinococcus granulosus* and is the suggested treatment frequency for such control interventions [4]. It is assumed that the drug eliminates the disease. In Kazakhstan, dogs stay mostly around households (discussion with study participants). Thus in scenario 1, we assume that 75% of dogs from the whole population are randomly selected and treated at each

intervention. In scenario 2, we increase the percentage to 95%, reflecting a larger control effort. The distribution of the time to extinction is then approximated by simulation. The treatment is started after the initial burn-in period. For scenario 1, the mean time to extinction of the disease is 13.6 years (95%CI:11.2, 16.3), whereas for the scenario 2, it is 11.7 years (10.2, 14.5).

The large values of the mean extinction time are due to the fact that the mean generation time for the cycle of infection in dogs is several years long, because the parasite has to wait in a sheep until it dies. The control interventions considered have the effect of reducing the mean duration $1/\alpha_1$ of infection, where α_1 is the sum of the death rate of dogs $1/r = 1/3$ and the inverse of the mean duration of an infection. For both scenarios, given that a dog is infected at the beginning of treatment k, he will lose its infection at the $(k+i)$th treatment $(i = 0, 1, \ldots)$ according to a geometric distribution with probability $p = 0.75$ respectively $p = 0.95$. Since the mean of the geometric distribution is $(1-p)/p$, the dog will stay infected for a mean time of $(1-p)/p \cdot 6$weeks, resulting in 2 weeks for scenario 1 and 0.3 for scenario 2. The infection time point of the dog is uniformly distributed between the $(k-1)$th and kth treatment, thus on average 3 weeks prior to its first treatment. Hence the mean duration of an infection becomes 5 weeks for scenario 1 and 3.3 weeks for scenario 2. We have seen in Subsection 4.2 that $R_0 = 2.980/\alpha_1$. Hence $R_0 \approx 0.3$ for scenario 1 and $R_0 \approx 0.2$ in scenario 2. Let the generation time be the time of infection of a dog until it infects indirectly another dog. Hence, starting from around $0.230n^{(1)} \approx 700$ infected dogs, scenario 1 implies that $5-6$ generations are needed to eliminate infection since $700(R_1)^6 < 1$, and in scenario 2, $4-5$ generations are needed. This suggests that, with the lifetime distribution observed in the sheep population, the mean generation time of an infection is around 2.5 years.

If a prepatent period of infection were included in the model, a rather smaller

number of generations would be needed to eliminate infection.

6. Discussion

In this paper, a mechanistic individual-based model is proposed to describe the whole life-cycle of the parasite *Echinococcus granulosus* in its two-host system between dogs and sheep. The findings significantly contribute to the understanding of the infection dynamics of the parasite.

It is shown that *Echinococcus granulosus* cysts at age 2 have an average probability of 6% to be fertile and the asymptotic probability of fertility of a cyst is 10%, indicating that 1 out of 10 older cysts are fertile. Furthermore, the mean infectiousness of a sheep conditional on harboring cysts is about 20%. In sheep at age 1 or younger, the mean infectiousness is about 5%, whereas in sheep older than 5 years, it is about 45%. The mean infectiousness in older sheep is higher since they have on average more and older cysts than younger sheep. The model suggests low infection rates in both hosts, indicating that on average a sheep gets infected once every 3 year, and a dog once every 2.5 years. However the mean loads per infection differs remarkably, with a mean of about 5 cysts in sheep and a mean parasite load of several thousands in dogs.

Simulation studies indicate that by using traditional intervention programs, the time to extinction of the disease can be as long as $11 - 14$ years, questioning commonly used short term control policies for *Echinococcus granulosus* [23].

Acknowledgments

The authors would like to thank the two referees for their valuable comments and suggestions. The work was supported by the Schweizerischer Nationalfonds (SNF), project no. 107726.

References

[1] M. Aminzhanov. Duration of the life of Echinococcus granulosus in the organism of dogs. *Veterinariia*, 12:70–72, 1975.

[2] F. Bass. A new product growth model for consumer durables. *Management Science*, 15:215–227, 1969.

[3] P. A. Cabrera, G. Haran, U. Benavidez, S. Valledor, G. Perera, S. Lloyd, M. A. Gemmell, M. Baraibar, A. Morana, J. Maissonave, and M. Carballo. Transmission dynamics of Echinococcus granulosus, Taenia hydatigena and Taenia ovis in sheep in Uruguay. *Int J Parasitol*, 25:807–813, 1995.

[4] P. A. Cabrera, S. Lloyd, G. Haran, L Pineyro, S. Parietti, M. A. Gemmell, O. Correa, A. Morana, and S. Valledor. Control of Echinococcus granulosus in Uruguay: evaluation of different treatment intervals for dogs. *Vet Parasitol*, 103:333–340, 2002.

[5] D. R. Cox and V. Isham. *Point Processes*. New York: Chapman and Hall, 2 edition, 1980.

[6] J. Eckert and P. Deplazes. Biological, epidemiological and clinical aspects of Echinococcosis, a zoonosis of increasing concern. *Clin Microbiol Rev.*, 17:107–135, 2004.

[7] P. Economides and G. Cristofi. *Cestode zoonoses: Echinococcosis and cysticercosis. An emergent and global problem*. NATO Science Series:IOS Press Amsterdam, 3 edition, 2002.

[8] M. A. Gemmell. Hydatid disease in Australia, III. Observations on the incidence and geographical distribution of hydatidiasis in sheep in New South Wales. *Aust Vet J*, 34:269–280, 1958.

[9] M. A. Gemmell. Hydatid diseases in Australia. IV. Observations on the inci-

dence of Echinococcus granulosus on stations and farms in endemic regions of New South Wales. *Aust Vet J*, 35:396–402, 1959.

[10] M. A. Gemmell, J. R. Lawson, and M. G. Roberts. Population dynamics in echinococcosis and cysticercosis: biological parameters of Echinococcus granulosus in dogs and sheep. *Parasitology*, 92:599–620, 1986.

[11] M. A. Gemmell, G. Oudemans, and T. Sakamoto. The effect of bithionol sulphoxide on Echinococcus granulosus and Taenia hydatigena infections in dogs. *Res Vet Sci*, 18:109–110, 1958.

[12] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J Chem Phys*, 81:2340–2361, 1977.

[13] D. Heinzmann, A. D. Barbour, and P. R. Torgerson. Compound processes as models for clumped parasite data. *Math. Biosci.*, 222:27–35, 2009.

[14] D. Heinzmann, A. D. Barbour, and P. R. Torgerson. Shot noise processes for clumped infections with time-dependent decay dynamics. *Submitted*, 2009.

[15] P. L. Moro, J. McDonald, R. H. Gilman, B. Silva, M. Verastegui, V. Malqui, G. Lescano, N. Falcon, G. Montes, and H. Bazalar. Epidemiology of Echinococcus granulosus infection in the central Peruvian Andes. *Bull. W.H.O.*, 75:553–561, 1997.

[16] K. P. Paaijmans, A. F. Read, and M. B. Thomas. Understanding the link between malaria risk and climate. *Proc Natl Acad Sci USA.*, 106:13844–13849, 2009.

[17] M. G. Roberts, J. R. Lawson, and M. A. Gemmell. Population dynamics in echinococcosis and cysticercosis: Mathematical model of the life-cycle of Echninococcus granulosus. *Parasitology*, 92:621–641, 1986.

[18] R. Simon. Length-biased sampling in ethiological studies. *Am. J. Epidem.*, 111:444–452, 1980.

[19] G. K. Sweatman and R. J. Williams. Survival of echinococcus granulosus and taenia hydatigena eggs in two extreme climatic regions of new zealand. *Res. Vet. Sci.*, 4:199–216, 1963.

[20] R. C. A. Thompson and A. J. Lymbery. *The biology of Echinococcus and hydatid disease*. London: George Allen and Unwin, 1986.

[21] B. Todorov and V. Boeva. Human echinococcosis in Bulgaria: a comparative epidemiologiocal analysis. *Bulletin WHO*, 77:110–118, 1999.

[22] P. R. Torgerson, K. K. Burtisurnov, B. S. Shaikenov, A. T. Rysmukhambetova, A. M. Abdybekova, and A. E. Ussenbayev. Modelling the transmission dynamics of Echinococcus granulosus in dogs in rural Kazakhstan. *Parasitology*, 126:417–424, 2003a.

[23] P. R. Torgerson and D. D. Heath. Transmission dynamics and control options for cystic echinococcosis. *Parasitol.*, 127:143–158, 2003d.

[24] P. R. Torgerson, B. Oguljahan, M. E. Muminov, R. R. Karaeva, O. T. Kuttubaev, M. Aminjanov, and B. Shaikenov. Present situation of cystic echinococcosis in Central Asia. *Parasitol Int.*, 55:207–212, 2006.

[25] P. R. Torgerson, K. Rosenheim, I. Tanner, I. Ziadinov, F. Grimm, M. Brunner, S. Shaiken, B. Shaikenov, A. Rysmukhambetova, and P. Deplazes. Echinococcosis, toxocarosis and toxoplasmosis screening in a rural community in eastern kazakhstan. *Trop Med Int Health*, 14:341–348, 2009.

[26] P. R. Torgerson, B. S. Shaikenov, A. T. Rysmukhambetova, A. E. Ussenbayev, A. M. Abdybekova, and K. K. Burtisurnov. Modelling the transmission dynamics of Echinococcus granulosus in sheep and cattle in Kazakhstan. *Vet Parasitol*, 114:143–153, 2003b.

[27] P. R. Torgerson, D. H. Williams, and M. N. Abo-Shehada. Modelling the prevalence of Echinococcus and Taenia species in small ruminants of different

ages in northern Jordan. *Vet Parasitol*, 79:35–51, 1998.

[28] P. Veit, B. Bliger, V. Schad, J. Schaefer, W. Frank, and R. Lucius. Influence of environmental factors in the infectivity of Echinococcus multilocularis eggs. *Parasitology*, 110:79–86, 1995.

[29] H. Wang, J. X. Zhang, P. M. Schantz, A. Ito, P. S. Craig, and X. H. Wu. Epidemiologic survey and analysis on echinococcosis in humans and animals from 1995 to 2005 in Qinghai province. *Chinese Journal of Zoonoses*, 22:11291134, 2006.

[30] I. Ziadinova, A. Mathis, D. Trachsel, A. Rysmukhambetova, T. A. Abdyjaparov, O.T. Kuttubaev, P. Deplazes, and P. R. Torgerson. Canine echinococcosis in Kyrgyzstan: Using prevalence data adjusted for measurement error to develop transmission dynamics models. *Int J Parasitol.*, 38:1179–1190, 2008.