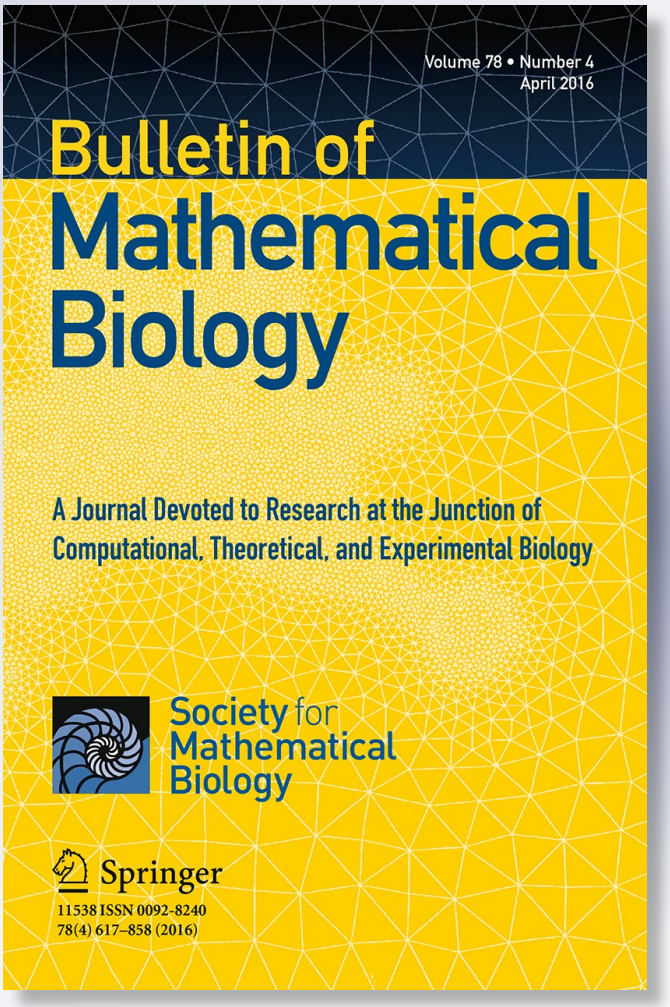
The Dynamics of **Ascaris lumbricoides** *Infections*

A. C. Fowler & T. Déirdre Hollingsworth

Bulletin of Mathematical Biology
A Journal Devoted to Research at the Junction of Computational, Theoretical and Experimental Biology Official Journal of The Society for Mathematical Biology

ISSN 0092-8240
Volume 78
Number 4

Bull Math Biol (2016) 78:815-833
DOI 10.1007/s11538-016-0164-2



Springer

Bull Math Biol (2016) 78:815–833
DOI 10.1007/s11538-016-0164-2

Society for Mathematical Biology

ORIGINAL ARTICLE

The Dynamics of *Ascaris lumbricoides* Infections

A. C. Fowler[1,2] · **T. Déirdre Hollingsworth**[3,4]

Received: 18 November 2015 / Accepted: 29 March 2016 / Published online: 11 April 2016

Abstract The Anderson–May model of human parasite infections and specifically that for the intestinal worm *Ascaris lumbricoides* is reconsidered, with a view to deriving the observed characteristic negative binomial distribution which is frequently found in human communities. The means to obtaining this result lies in reformulating the continuous Anderson–May model as a stochastic process involving two essential populations, the density of mature worms in the gut, and the density of mature eggs in the environment. The resulting partial differential equation for the generating function of the joint probability distribution of eggs and worms can be partially solved in the appropriate limit where the worm lifetime is much greater than that of the mature eggs in the environment. Allowing for a mean field nonlinearity, and for egg immigration from neighbouring communities, a negative binomial worm distribution can be predicted, whose parameters are determined by those in the continuous Anderson–May model; this result assumes no variability in predisposition to the infection.

Keywords Infectious diseases · *Ascaris lumbricoides* · Mathematical model · Negative binomial distribution

✉ A. C. Fowler
andrew.fowler@ul.ie

1 MACSI, University of Limerick, Limerick, Ireland

2 OCIAM, University of Oxford, Oxford, UK

3 Mathematics Institute, University of Warwick, Coventry, UK

4 Liverpool School of Tropical Medicine, Liverpool, UK

1 Introduction

Ascaris lumbricoides, or roundworm, is a ubiquitous infection of low-income populations with poor sanitation in tropical countries (Scott 2008). It has been present in human populations for thousands of years (Cox 2002). It was originally widely prevalent throughout the world (Tyson 1683; Stoll 1947; Crompton 2001), but was largely eradicated from developed countries in the twentieth century [(the Japanese experience is described by WHO (1996)], and developed countries (North America, Europe, Russia) now only register a handful of cases (Pullan et al. 2014; Crompton 2001). Recent estimates put the number of people infected at 820 million, with considerably more at risk (Pullan et al. 2014). Infection is caused when eggs excreted in faeces are ingested. Maturing to a larval stage, they migrate through the blood to the lungs, before being coughed up and reingested to the small intestine, where the adult worm matures. Infection with *A. lumbricoides* rarely causes death, but can lead to chronic disability, leading to poor physical and cognitive development and school achievement (Bethony et al. 2006).

In recent years, there has been an enormous investment in providing free treatments to children in affected areas (see http://unitingtocombatntds.org). These drugs effectively clear infection, but do not affect the environment and so reinfection occurs rapidly. Therefore, there are a number of questions arising concerning how to design these treatment programmes, including how rapid is "bounceback" following mass treatment, and therefore how frequently should treatment be given to push infection levels down (Anderson et al. 2012; Jia et al. 2012; Truscott et al. 2014). While previous analysis has shown that this is likely to be dominated by the life expectancy of the worm (Anderson and May 1991), we do not yet have approximations which include the background transmission rate, which is likely to play an important role in bounceback.

It has long been noticed that a feature of macroparasites, such as *Ascaris*, is that they are very over-dispersed in the population, with a small proportion of the population harbouring the highest number, or intensity, of worms and this is commonly represented as a negative binomial distribution (Anderson and May 1978; May and Anderson 1978). Since severity of symptoms is related to the intensity of infection (Bethony et al. 2006) and is expected to be correlated with infectivity, it is essential to understand the drivers of this distribution of worms and the impact of treatment upon it. The negative binomial distribution has been shown to be generated by varying susceptibility across hosts (Bartlett 1960; May and Anderson 1978), but is commonly assumed as a given property of the population in modelling (Anderson and May 1985; 1991).

Our purpose in this paper is to examine the way in which the dynamics of the infection can provide a cause for the observation of a negative binomial distribution. As mentioned, this can be due to a distribution of host susceptibility, but this explanation simply pushes the observation back to the question of why susceptibility should be gamma distributed (Bartlett 1960). Rather, we are interested in whether a model of the disease can in itself produce the observed negative binomial distribution. The answer to this is yes, but it relies in the model on assuming a nonzero immigration rate of mature eggs.

The way in which *Ascaris* infections are promoted and distributed in poor communities has been described by, for example, Otto et al. (1931), Cort and Stoll (1931), Forrester et al. (1988), Anderson et al. (1993), and Williams et al. (1974), who variously describe the nature of infection in a number of places, including Tennessee and Virginia, USA, China, Korea, Japan, Mexico, and Guatemala. Many of these studies focus on the transmissibility of the disease, establishing the domestic environment as a primary hot bed of infection, but also emphasising the wider avenues of transmission (Cairncross et al. 1996; Otto et al. 1931; Cort and Stoll 1931). In essence, the vehicle of transmission lies in faecal deposition containing viable eggs on the open ground, and its subsequent migration via the vector of social contact, and also in some cases the use of human excrement as fertiliser, leading to installation of viable eggs on exported vegetables (Cort and Stoll 1931). Essentially immigration of eggs occurs through the normal modes of human contact, and special efforts would be required to suppress it.

Some information on the spatial spread of *Ascaris* can be inferred from genetic studies. For example, Anderson et al. (1995) found that there were strong genetic similarities of *Ascaris* between families in the same village, but less commonality with different villages, while Betson et al. (2012) found evidence of high gene flow of *Ascaris* between two villages in Uganda. Evidently, both social and economic intercourse allow for an effective diffusion of *Ascaris* eggs, and in a particular community, this provides a mechanism for effective immigration. We comment further on this in the conclusions.

Our method proceeds by reformulating the continuous Anderson and May (1991) model in a stochastic form (which is really the underlying description whence a continuous model is derived), which describes the stochastic evolution of two coupled populations (free-living egg stage and adult worm). This leads to a partial differential equation for the corresponding generating function, and we find that an approximate solution is possible, based on the disparity of the timescales in each equation. While similar ideas have been used before (e. g., Hadeler and Dietz 1983; Kretschmar and Adler 1993), our novelty lies in providing an explicit approximate solution of the three characteristic equations which describe the generating function p.d.e. It is as a consequence of this solution that we are able to predict the occurrence of a negative binomial distribution.

Stochastic models in epidemiology, and particularly for helminth infections, have a long history. For example, a simple comparison between deterministic and stochastic models was studied for simple epidemics by Jacquez and O'Neill (1991). Generally, comparable analytic results for the stochastic model as for the deterministic model are not readily available, and this has been a common observation (e. g., Isham 1995). Stochastic models of parasite distributions were probably first studied by Kostitzin (1934, p. 20 ff.), who opined that his infinite system of nonlinear differential equations 'presented nearly insurmountable difficulties'. Tallis and Leyton (1966) consider a general form of stochastic model, while Tallis and Leyton (1969) consider the more particular case of helminth infections and in certain cases obtain negative binomial distributions, similarly to Bartlett (1960), depending on the assumed probability distribution of infection.

Hadeler and Dietz (1983) consider a quasi-linear model for the distribution of parasites in an infected host population, where the nonlinearity of the model is affected

through a dependence of the (larval) uptake rate on the mean parasite burden, based on the idea that the larval population is in quasi-equilibrium. This is somewhat similar to the approach that we take here. A similar model was considered by Kretschmar and Adler (1993).

Stochastic models of helminth infections have been reviewed by Cornell (2010). See also the articles by Walker *et al*. and Hollingsworth *et al*. in the book edited by Holland (2013). Barbour and Kafetzaki (1991) address the overdispersion of observed parasite distributions with a susceptible infective model with various assumed infection probability distributions. Isham (1995) introduced a stochastic model for parasite burden as a function of host age, and later Herbert and Isham (2000) extended this to three stages: eggs, larvae, and adults. Walker et al. (2010) provide a stochastic model which focusses on the distinction between 'trickle' and 'clumped' infection rates, a theme to which we will return. Bottomley et al. (2005) develop a stochastic model for two species of competing helminths, which leads to two coupled equations for the respective probability densities. The analysis is limited to a linearised system, and computation of the mean and variance of the distributions. Adler and Kretzschmar (1992) consider a stochastic parasite model which ignores the free-living stages of the parasite. Gaba et al. (2006) use a computational stochastic model in a sheep/nematode system. While some analytic progress can be made in some of these models (e.g. derivation of equations for the mean and variance of the resulting distribution), it seems that no direct approximate solution for the time evolution of the probability distribution of coupled worm/egg populations has been provided. We are able to do this here by combining the stochastic description with an asymptotic analysis of the phase plane structure of the characteristic equations describing the generating function, based on a separation of the timescales of the different populations. It is clear that this idea will have wider applicability in other systems.

The structure of the paper is as follows. In Sect. 2, we review the Anderson–May model and show how it reduces to a set of two ordinary differential-delay equations. This model is then analysed in Sect. 3, and the nonlinear effects of mating and fecundity yield the familiar results of bistability and population saturation. Section 4 develops a stochastic version of the two-component population model of Sect. 3 and shows, by means of an approximate solution, how a negative binomial distribution can be predicted. A discussion of the results follows in Sect. 5, and the conclusions follow in Sect. 6. Improvements in modelling structure and analyses form one of the seven challenges facing the study of neglected tropical diseases (Hollingsworth et al. 2015), and the main purpose of our paper is to contribute to this development.

2 Mathematical Model

We begin by reviewing and elaborating the dynamics of an *Ascaris* infection in a single human, who is part of a community of $\bar{N}$ identical individuals. The basic Anderson–May model for directly transmitted helminth infections can be represented by the diagram in Fig. 1. The five boxes represent the five basic variables of the model: E, the immature eggs in the environment, L, the mature infective eggs in the environment, H, the ingested eggs in their larval migratory phase through the body, where they may

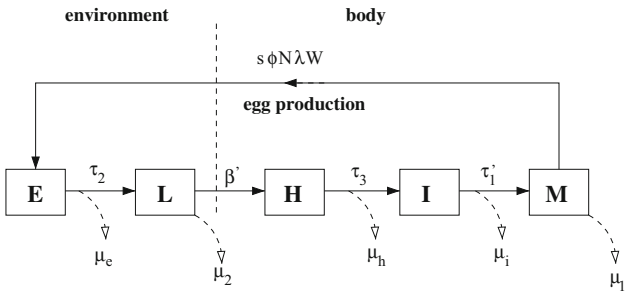


Fig. 1 Schematic representation of the Anderson–May *Ascaris* model. Adult worms (M) in the body produce eggs (E) which are excreted to the environment. These eggs mature over a time τ_2, and the mature eggs (L) are ingested at a rate β'. The ingested eggs enter a larval stage (H) and migrate to the blood and other organs before returning after a time τ_3 to the small intestine (I), where they develop into adult worms over a time τ_1'. Mortality rates at each stage are given by the coefficients μ_k

be systematically attacked by the immune system, I, the ingested infective larvae in the small intestine, and M, the mature worms, these last three being in a single human. The units of L and M are dimensionless, as they are taken as pure numbers. The units of E, H, and I are day^{-1} (per day) because they are distributions with respect to maturation or transit time m: specifically, E, H, and I satisfy the partial differential equations

$$\begin{aligned} \frac{\partial E}{\partial t} + \frac{\partial E}{\partial m} &= -\mu_{\mathrm{e}} E, \\ \frac{\partial H}{\partial t} + \frac{\partial H}{\partial m} &= -\mu_{\mathrm{h}} H, \\ \frac{\partial I}{\partial t} + \frac{\partial I}{\partial m} &= -\mu_i I, \end{aligned} \tag{1}$$

and the respective boundary conditions are

$$\left.\begin{aligned} E &= E_0(t) \equiv s\phi \bar{N} \lambda M(t), \\ H &= \beta' L(t), \\ I &= H(t, \tau_3), \end{aligned}\right\} \quad \text{at} \quad m = 0. \tag{2}$$

Here s is the fraction of female (egg-producing) worms, ϕ is the probability of mating (which will depend on M), $\bar{N}$ is the human community size, λ is the specific egg production rate, and β' is the transmission coefficient; in (1), μ_{e}, μ_{h} and μ_i are the egg mortality rates.

Solution of the equations is straightforward using the method of characteristics (details are provided in the Appendix), and we find

$$E = E_0(t-m)\mathrm{e}^{-\mu_{\mathrm{e}} m}, \quad H = \beta' L(t-m)\mathrm{e}^{-\mu_{\mathrm{h}} m}, \quad I = H(t-m, \tau_3)\mathrm{e}^{-\mu_i m}. \tag{3}$$

The equations for L and M take the form

$$\begin{aligned} \dot{L} &= E|_{m=\tau_2} - \mu_2 L - \beta' \bar{N} L, \\ \dot{M} &= I|_{m=\tau_1'} - \mu_1 M, \end{aligned} \tag{4}$$

where μ_2 is the mature egg mortality rate and μ_1 is the mature worm mortality rate; thus,

$$\dot{L} = rM(t-\tau_2) - \mu_2 L - \beta' \bar{N} L,$$
$$\dot{M} = \nu_0 L(t-\tau_1) - \mu_1 M, \tag{5}$$

where

$$\tau_1 = \tau_1' + \tau_3, \quad r = sd_2\phi\bar{N}\lambda, \quad \nu_0 = \beta d_1, \quad \beta = \beta' d_3, \tag{6}$$

and the survival coefficients d_1, d_2, and d_3 are defined by

$$d_1 = \mathrm{e}^{-\mu_i \tau_1'}, \quad d_2 = \mathrm{e}^{-\mu_\mathrm{e}\tau_2}, \quad d_3 = \mathrm{e}^{-\mu_\mathrm{h}\tau_3}. \tag{7}$$

This is the basic Anderson–May model for L and M. We will provide elaborations of this model later. There are some differences in detail between (5) and Eqs. (16.7) and (16.8) of Anderson and May (1991). We ignore a loss term in the equation for M due to human mortality; this is in any case small. The main difference in the present version of the model is that Anderson and May take $\beta = \beta'$, which is equivalent to ignoring mortality in the migratory H phase of the egg population. In fact, on page 470 of their book, β is indeed defined analogously to its definition here, and Anderson and May also suppose (page 472) that $\mu_2 \gg \beta'\bar{N}$, so in practice there is little difference.

3 Simplification and Analysis

Estimates of the parameters of the model in (5) are given in Table 1, based on values provided by Anderson and May (1991). We suppose that $\mu_\mathrm{e} = \mu_2$, whence our estimate for $d_2 = 0.7$, slightly higher than the in vivo value of $d_1 \approx 0.6$, where there is a hostile environment, and longer maturation time. There is no estimate for migratory mortality, and our value of 0.5 is nominal. In addition, Anderson and May provide no estimate for β, but we can infer the value of β from the estimates of the basic reproduction rate $R_0 \gtrsim 1$, as given in their table 16.3. The detail of this calculation is given following (11).

We begin by supposing that the natural mortality of infectious eggs in the environment is much greater than that of the uptake by humans, that is, $\beta' N \ll \mu_2$ (this assumption is cosmetic, in the sense that otherwise we simply replace μ_2 by $\mu_2 + \beta' N$ below, which is equivalent to taking a smaller value of μ_2^{-1} in Table 1). We nondimensionalise the model (5) by scaling the variables as

$$M \sim M_0, \quad L \sim \frac{rM_0}{\mu_2}, \quad t \sim \frac{1}{\mu_1}, \tag{8}$$

and then, the dimensionless form of the equations is

$$\varepsilon \dot{L} = M_{\varepsilon_2} - L,$$
$$\dot{M} = R_0 L_{\varepsilon_1} - M, \tag{9}$$

Table 1 Parameter values

Symbol	Meaning	Typical value
d_1	Ingested egg survival coefficient	0.5–0.7
d_2	Egg survival coefficient in wild	0.7
d_3	Larval survival coefficient in migration	0.5
$\bar{N}$	Human community size	10^2
R_0	Basic reproduction rate	1–5
s	Proportion of female worms	0.5
M_0	Mean worm burden	10–20
β	Transmission coefficient	$\sim 10^{-10}$ day^{-1}
β'	Egg uptake rate	$\sim 2 \times 10^{-10}$ day^{-1}
λ, λ_0	Egg production rate	2×10^5 day^{-1}
μ_1^{-1}	Worm life expectancy	1–2 years
μ_2^{-1}	Mature egg life expectancy	28–84 days
ϕ	Mating probability	$\in (0, 1)$
τ_1	Internal egg maturation time	50–80 days
τ_2	External egg maturation time	10–30 days

where

$$F_\tau \equiv F(t - \tau), \tag{10}$$

and the dimensionless parameters are defined by

$$\varepsilon = \frac{\mu_1}{\mu_2}, \quad \varepsilon_2 = \mu_1 \tau_2, \quad \varepsilon_1 = \mu_1 \tau_1, \quad R_0 = \frac{\nu_0 r}{\mu_1 \mu_2} = \frac{\beta d_1 s d_2 \phi N \lambda}{\mu_1 \mu_2}. \tag{11}$$

We mentioned earlier that Anderson and May (1991) do not provide an estimate for the transmission coefficient β. They do, however, provide estimates for the basic reproduction rate of the infection, which is the parameter R_0 defined above. Three different estimates from Iran, Burma, and Bangladesh (table 16.3, p. 478) lie in the range $R_0 = 1 - 5$. From (11), we have

$$\beta = \frac{R_0}{\mu_1^{-1} \mu_2^{-1} d_1 s d_2 \phi N \lambda}, \tag{12}$$

and using the values of the other parameters in Table 1, together with $\phi = \frac{1}{2}$, we obtain a range for β of $0.07 - 2.8 \times 10^{-10}$ day^{-1}. If we accept the value of R_0 of $O(1)$, which is inferred from the recovery timescale of the infection (e.g., Anderson and May 1991, figure 17.4), then this calculated value of β seems very small, and it raises the issue of whether this model for egg uptake is realistic; we come back to this issue later in the discussion (Sect. 5). First, we complete the analysis of the model on the basis that (5) is essentially correct.

From Table 1, we have $\varepsilon, \varepsilon_1, \varepsilon_2 \ll 1$, and thus, the delays can be ignored, the infected egg population L rapidly approaches equilibrium, and the worm population satisfies the approximate equation

$$\dot{M} = (R_0 - 1)\, M, \tag{13}$$

where, as mentioned above, Anderson and May (1991) estimate R_0 in different communities as having typical values $R_0 \approx 1$–5.

The worm population scale M_0 is undetermined, because the model (9) or (13) is linear. In particular, if $R_0 > 1$ then unbounded growth occurs; in reality, the worm population is limited by nonlinearities, as discussed below.

3.1 Nonlinearity and Saturation

Anderson and May (1991) address the issue of the linearity of (5) by proposing two nonlinear dependences of the parameters in the definition of R_0. The mating probability must be a function of M, since we must have $\phi = 0$ for $M = 1$, $\phi = 0.5$ for $M = 2$, and so on. Generally, ϕ is an increasing function of M, asymptoting to one for large M. A simple estimate for ϕ follows from assuming mating occurs if there is at least one male and one female worm. In this case,

$$\phi = 1 - \frac{1}{2^{M-1}}. \tag{14}$$

Similarly, egg fecundity λ is a decreasing function of M. Measurements indicate that while λ decreases (Anderson and May 1991, figure 15.14), the net production $M\lambda$ increases with M, as we might expect (Hall and Holland 2000). The simplest choice for a decreasing fecundity which satisfies these constraints is the algebraic decay function

$$\lambda = \frac{\lambda_0 M_0}{M + M_0}, \tag{15}$$

where M_0 then provides the natural scale for the worm population. This is not dissimilar to other algebraic data fits (Anderson and Medley 1985). With these modifications, the dimensionless Anderson–May model (13) takes the form

$$\dot{M} = [R_0 \psi(M) - 1]\, M, \tag{16}$$

where

$$\psi(M) = \frac{\phi(M)}{1 + M}, \tag{17}$$

and a reasonable representation for ϕ is, from (14),

$$\phi = 1 - \mathrm{e}^{-\alpha M}, \tag{18}$$

where $\alpha \approx M_0 \ln 2$, and has a typical value in the range $7 - 14$.

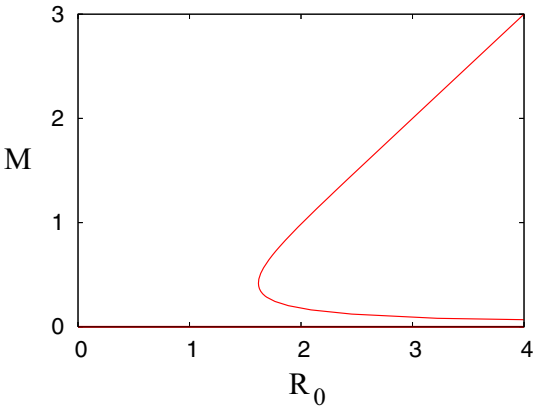


Fig. 2 The bistable equilibria of (16), assuming (17) and (18), with $\alpha = 5$

The nonlinear model (16) provides a classical hysteretic transition from the stable state $M = 0$ to the stable upper branch of Fig. 2. Since in practice α is large, the threshold for transition is very low, and we can approximately take $\phi = 1$ in (17). In this case (16) is simply

$$\dot{M} = \left(\frac{R_0}{1+M} - 1 \right) M, \tag{19}$$

and the stable steady state is just $M \approx R_0 - 1$. The general solution of (19) is just

$$t = \frac{1}{R_0 - 1} \ln \left[\frac{M}{(R_0 - 1 - M)^{R_0}} \right]. \tag{20}$$

In particular, the dimensionless time, following an intervention which reduces the worm burden to a fraction f_I of the steady state, for it to recover to a fraction f_R of the steady state, is just

$$t_R = \frac{1}{(R_0 - 1)} \ln \left(\frac{f_\mathrm{R}}{f_\mathrm{I}} \right) + \frac{R_0}{(R_0 - 1)} \ln \left(\frac{1 - f_\mathrm{I}}{1 - f_\mathrm{R}} \right). \tag{21}$$

This gives a simple expression for the bounceback time.

4 The Stochastic Anderson–May Model

The distribution of worm load in humans is highly skewed: most infected carriers have one or two worms, and the number with higher burdens shrinks rapidly, although there is a fat tail to the probability density. Anderson and May (1991) show that a negative binomial distribution fits measured worm burden profiles very well. They give the probability density of having j worms as

$$p_j = \frac{(j+k-1)!}{j!(k-1)!} \alpha^j (1-\alpha)^k, \tag{22}$$

where in Anderson and May's notation,

$$\alpha = \frac{m}{m+k}, \tag{23}$$

and m and k are positive parameters: m is the mean of the distribution. We form the probability generating function

$$G_{\mathrm{W}}(z) = \sum_{0}^{\infty} p_j z^j, \tag{24}$$

from which we find

$$G_{\mathrm{W}}(z) = \left(\frac{1-\alpha}{1-\alpha z}\right)^k, \tag{25}$$

which gives the negative binomial distribution its name.

A natural explanation for the prevalence of negative binomial distributions in the human population can be found through a stochastic process which describes within-host birth, death, and immigration (Bartlett 1960). However, such a model does not apply directly to macroparasite infections, because births do not take place within the host. We now aim to formulate a stochastic model to allow for the Anderson–May environmental dynamics.

We consider a community of $\bar{N}$ people, each of whom has a random number of n worms, and we suppose that the local environment contains m mature eggs, or m cohorts (stools) of mature eggs [in the latter case, we are considering the idea of 'clumped' infection (Isham 1995; Cornell 2010; Walker et al. 2010)]. The joint probability of an individual having n worms and there being m mature eggs is denoted $p_{m,n}$. To derive a stochastic equation for $p_{m,n}$, we suppose the following processes occur: worms die with probability $\mu_1\,\mathrm{d}t$ in a time interval $\mathrm{d}t$, mature eggs are taken in by humans with specific probability $\nu_0\,\mathrm{d}t$, they die with probability $\mu_2\,\mathrm{d}t$, and they are produced by individual worms at a rate r. As will be crucial, we also assume that (mature) eggs are imported from elsewhere at a rate ν_{e}. Most obviously, this is by means of human traffic.

These assumptions lead to the sequence of differential equations

$$\begin{aligned}\dot{p}_{m,n} = &-[n\mu_1 + \nu_0 m + \mu_2 m + rn + \nu_e]p_{m,n} + \mu_1(n+1)p_{m,n+1}\\ &+ \nu_0 m p_{m,n-1} + \mu_2(m+1)p_{m+1,n} + rnp_{m-1,n} + \nu_e p_{m-1,n},\end{aligned} \tag{26}$$

and, defining the joint probability generating function

$$\Pi(z,w) = \sum_{m,n\geq 0} p_{m,n} z^m w^n, \tag{27}$$

(and taking $p_{i,j} = 0$ if i or $j < 0$), we can derive the partial differential equation

$$\begin{aligned}&\frac{\partial \Pi}{\partial t} + [\mu_1(w-1) - r(z-1)w]\frac{\partial \Pi}{\partial w}\\ &+ [\mu_2(z-1) - \nu_0(w-1)z]\frac{\partial \Pi}{\partial z} = \nu_e(z-1)\Pi.\end{aligned} \tag{28}$$

In terms of this distribution, the mean quantities L and M of Sect. 2 are defined by

$$L = \left.\frac{\partial \Pi(z,1)}{\partial z}\right|_{z=1}, \quad M = \left.\frac{\partial \Pi(1,w)}{\partial w}\right|_{w=1}, \tag{29}$$

and if we differentiate (28) with respect to z and w and apply the appropriate limits, we regain (5) in the form

$$\begin{aligned} \dot{L} &= rM - \mu_2 L + \nu_e, \\ \dot{M} &= -\mu_1 M + \nu_0 L, \end{aligned} \tag{30}$$

indicating the values of the egg production rate r and egg intake rate ν_0 are the same as defined earlier in (11). Because of the dependence of ϕ and λ, and thus r, on M, the parameter

$$R = \frac{r\nu_0}{\mu_1 \mu_2} \tag{31}$$

will also be a function of M. For $M = 0$, we have $R = R_0$ as in (11), but the nonlinear dependence of r on M means that the value of $R = R_c$ at the stable steady state is different, and crucially less than one; if we assume (19), for example, we have the simple relation

$$R_c = \frac{1}{R_0}. \tag{32}$$

We come back to discuss the problem when r varies later; for the moment, we just take it as a constant.

If the initial populations have values M and N, then the initial condition for Π is

$$\Pi = z^M w^N \quad \text{at} \quad t = 0. \tag{33}$$

It is convenient to define

$$Z = z - 1, \quad W = w - 1; \tag{34}$$

then the characteristic equations for (28) are

$$\begin{aligned} \dot{Z} &= \mu_2 Z - \nu_0 W(1+Z), \\ \dot{W} &= \mu_1 W - rZ(1+W), \\ \dot{\Pi} &= \nu_e Z \Pi. \end{aligned} \tag{35}$$

Despite their simplicity, solution of these for Z and W does not appear feasible in closed form. Instead, we take advantage of the fact that $\mu_1 \ll \mu_2$, so that the system $(35)_{1,2}$ is relaxational, with Z being the fast variable. Additionally, note that our concern is with the distribution of the worm numbers. If we denote its generating function as $G_W(w)$, then in terms of z and w,

$$G_W(w) = \Pi(1, w); \tag{36}$$

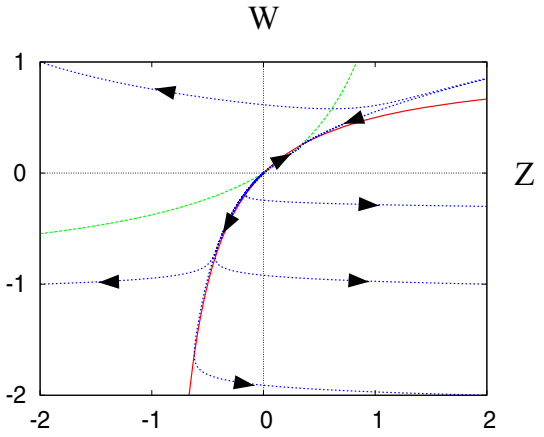


Fig. 3 Phase portrait for $(35)_{1,2}$, using values $\mu_1 = 0.002$ day^{-1}, $\mu_2 = 0.02$ day^{-1}, $R_c = 0.6$, and $\nu_0 = 0.02$ day^{-1}; the value of r is determined from (31). The *solid* (*red online*) curve is the $\dot{Z} = 0$ nullcline, and the *dashed* (*without arrows*, *green online*) *curve* is the $\dot{W} = 0$ nullcline. Apart from the origin, which is an unstable node, there is a second fixed point at $\left(\frac{1}{3}, \frac{1}{2}\right)$, which is a saddle point, and located at the apparent slight discontinuity (which arises through the solution of two separate trajectories)

note that $z = 1$ when $Z = 0$ in Fig. 3.

Figure 3 shows a phase portrait in the (Z, W) phase plane for a typical set of parameters having $\mu_1 \ll \mu_2$, and in which $\nu_0 \sim \mu_2$ (though as we see later the case $\nu_0 \ll \mu_2$ is more likely). The figure illustrates the case where $r < \dfrac{\mu_1\mu_2}{\nu_0}$, for which the fixed point other than the origin lies in the first quadrant; the only difference for the case $r > \dfrac{\mu_1\mu_2}{\nu_0}$ is that it lies in the third, but the phase portrait is otherwise the same.

What concerns us is the large time evolution of Π along the characteristics, since this will give us the limiting distribution, and in view of (36), we are also interested in the initial conditions for (35) which intersect the W axis, on which $Z = 0$, i. e., $z = 1$. It suffices to discuss the trajectories in $W < 0$, which must thus originate from the saddle point at the origin, emerging almost along the lower unstable separatrix. For the case where $\mu_1 \ll \mu_2$, these trajectories remain almost on the Z nullcline until they leave at the last moment and dive almost horizontally to the right. This allows us to obtain an approximate solution for Π.

We consider first the case $\nu_0 \sim \mu_2$. The trajectories in $W < 0$ remain approximately on the Z nullcline

$$W = \frac{\mu_2 Z}{\nu_0(1+Z)}, \tag{37}$$

on which therefore

$$\dot{W} \approx \mu_1 W - \frac{r\nu_0 W(1+W)}{\mu_2 - \nu_0 W}. \tag{38}$$

Similarly, from (35) we can derive the approximate equation for Π on the Z nullcline,

$$\frac{1}{\Pi}\frac{d\Pi}{dW} = \frac{\nu_0\nu_e}{\mu_1(\mu_2 - \nu_0 W) - \nu_0 r(1+W)}, \tag{39}$$

and the appropriate initial condition for long time solutions is

$$\Pi = 1 \quad \text{at} \quad W = 0. \tag{40}$$

Defining

$$\Omega = \frac{\mu_1\mu_2 - r\nu_0}{\nu_0(\mu_1 + r)} = \frac{\mu_2(1-R_c)}{\nu_0 + \mu_2 R_c}, \tag{41}$$

the solution of this is

$$\Pi = \left(\frac{\Omega}{\Omega - W}\right)^k, \tag{42}$$

where

$$k = \frac{\nu_e}{\mu_1 + r}, \tag{43}$$

and rewriting this in terms of $w = 1 + W$ yields

$$\Pi = \left(\frac{1-\alpha}{1-\alpha w}\right)^k, \tag{44}$$

where

$$\alpha = \frac{1}{1+\Omega} = \frac{\nu_0 + \mu_2 R_c}{\nu_0 + \mu_2}. \tag{45}$$

This gives the familiar negative binomial distribution on the Z nullcline, but it remains approximately valid also on the W axis, since the change in W between the two is asymptotically small. Hence, the long-term worm distribution is (approximately) negative binomial.

We now discuss the solution of the characteristic equations (35) when $\nu_0 \ll \mu_2$. It is appropriate to rescale the variables as

$$Z \sim \delta = \frac{\nu_0}{\mu_2}, \quad t \sim \frac{1}{\mu_1}, \tag{46}$$

which leads to the (rescaled) equations

$$\begin{aligned} \mu\dot{Z} &= Z - W(1+\delta Z), \\ \dot{W} &= W - RZ(1+W), \\ \dot{\Pi} &= \nu Z\Pi, \end{aligned} \tag{47}$$

where

$$\mu = \frac{\mu_1}{\mu_2}, \quad \nu = \frac{\nu_e\nu_0}{\mu_1\mu_2}. \tag{48}$$

Since $\mu \ll 1$, the Z equation is still fast and the earlier discussion applies, except now the Z nullcline is approximately $Z = W$, since $\delta \ll 1$. Thus on this nullcline,

$$\frac{1}{\Pi}\frac{d\Pi}{dW} \approx \frac{\nu}{1 - R(1+W)}, \tag{49}$$

whose solution (in terms of w) satisfying $\Pi = 1$ on $w = 1$ is

$$\Pi = \left(\frac{1-R}{1-Rw}\right)^{\nu/R}, \tag{50}$$

so that we again obtain the negative binomial distribution (44) with parameters

$$\alpha = R, \quad k = \frac{\nu}{R} = \frac{\nu_e}{r}. \tag{51}$$

Note that at the stable steady state $R = R_c < 1$.

5 Discussion

The negative binomial distribution is commonly written in terms of k and the mean $m = \dfrac{\alpha k}{1-\alpha}$, and thus, we have

$$\begin{aligned} k &= \frac{\nu_e}{\mu_1 + r}, \quad m = \frac{\nu_e(\nu_0 + \mu_2 R_c)}{\mu_2(\mu_1 + r)(1 - R_c)}, \quad \nu_0 \sim \mu_2, \\ k &= \frac{\nu_e}{r}, \quad m = \frac{\nu_e R_c}{r(1 - R_c)}, \quad \nu_0 \ll \mu_2. \end{aligned} \tag{52}$$

Note that if $R_c \sim 1$ as we suppose, then $\nu_0 \ll \mu_2$ only if $\mu_1 \ll r$; so we see that the first result in (52) includes the second as a particular limit and is uniformly valid.

The estimates in Table 1 suggest the egg production rate is $r \sim 10^7$ day^{-1}, whence the assumption that $R_c \sim 1$ suggests $\nu_0 \sim \beta \sim 10^{-10}$ day^{-1}. As we intimated earlier, it may not be the actual egg production which is important, but the infected (clumped) stool production (Isham 1995; Walker et al. 2010). Eggs are not distributed randomly in the environment. They are concealed in faeces which themselves are deposited occasionally and locally. And within these faeces, only the eggs located on the outside should be available for uptake. Thus, although the adult worm produces 10^5 eggs per day, these are packaged in one set of faeces and only a small fraction will be exposed and available. More specifically, if 10^5 eggs of diameter 50 μm are distributed uniformly in a 2 cm diameter stool, then a rough calculation suggests that $\sim 10^3$ will present themselves on the surface. More importantly, it is not really the egg density which is important but the faecal stool density. In terms of infected stool production, a rate $\lambda \sim 1$ day^{-1} is more reasonable. If we take total infected stool production at 10 day^{-1}, then the necessary uptake rate is $\nu_0 \sim 10^{-5}$ day^{-1}, but this

is still much less than μ_2, so it seems the assumption $\nu_0 = \beta d_1 \ll \mu_2$ is safe; in that case we can use the second line of (52) in our discussion.

Observed distributions can then give us some further understanding of the value of the egg uptake rate and the immigration rate. As an example, we consider some Korean data (Anderson and May 1991, figure 15.17), for which $m = 2.2$ and $k = 0.32$; thus $R_c = 0.87$, and more importantly, $\nu_e = kr \lesssim r$; the immigration rate would need to be comparable to the local production rate. This can make sense, if children of neighbouring villages frequently visit.

It is essential in our analysis to include the effects of external immigration. Partly, this is because with the effects of nonlinearity included, immigration is necessary in order to make the zero state unstable. More pertinently, if there is no immigration, then the initial condition is retained in the distribution, although there may be some effect of nonlinearity on this. Indeed, straightforward asymptotic solution of (35) with $\nu_e = 0$ leads to the approximate distribution

$$\Pi = \left[1 + (w-1)\left(\frac{\Omega}{\Omega + 1 - w}\right)^{\nu'} \exp\left(-\frac{r't}{\mu_2}\right)\right]^N, \tag{53}$$

where Ω is as given in (41), and

$$r' = \mu_1\mu_2 - r\nu_0, \quad \nu' = 1 - \frac{\nu_0\Omega}{\mu_2}. \tag{54}$$

The behaviour of the distribution at large time is opaque; however, because with no immigration, the equilibrium worm density is obtained when $r' = 0$ and thus also $\Omega = 0$. Naïve insertion of these limits implies extinction, which cannot be the case, and a more subtle investigation is necessary, but we do not pursue this here.

A comment should be made concerning the assumption that the nonlinear dependence of the production rate r is on the mean M and not n. If we were dealing with a single individual, this would not be the case, and the differential equations (26) would be genuinely nonlinear. However, because we have a reasonably large number of individuals, which we suppose represents the worm distribution, the correct expression for r in (26) and thus eventually (35) is

$$\bar{r} = \overline{r(n)}, \tag{55}$$

where the overline represents an average over the population. The characteristic equations are the same (e. g., (47), but with R replaced by $\bar{R}$, but the eventual distribution is the same, except that $\bar{R}$ is then the average of R over the distribution.

In view of the asymptotic assumption that $\mu_1 \ll \mu_2$, one might suppose that the direct stochastic equivalent of the first-order Eq. (13) would give similar results. This is not the case, as it is simple to show that the result is a Poisson process for the worm distribution. Nor does the assumption of an exponential distribution of lifetimes affect this. Assumption of a fixed finite lifetime just leads to the renewal equation and again a Poisson process with mean proportional to the product of the effective

egg uptake β and the worm lifetime T. However, if we additionally suppose that this product itself has a gamma distribution, then a negative binomial distribution for the worm population again ensues. Thus, an alternative explanation for such distributions is a simple immigration-death process for the adult worms, together with a gamma distribution for uptake rates, for example. Such distributions are not unreasonable, insofar as very young children are protected, but their uptake will rise sharply when they are toddlers but decrease as they grow up. It remains to be seen whether such an explanation is consistent with observed values of m and k.

6 Conclusions

We have addressed and provided a solution to the question of why human communities subject to endemic infections of the helminth *Ascaris lumbricoides* generally display a negative binomial distribution of adult worm numbers in the human hosts. Our solution method generalises the classic Anderson–May model to a coupled stochastic/deterministic process, and we show that the dynamics of the infection naturally leads to the evolution of negative binomial distributions, providing we include the effect of egg immigration into the model.

If this is the correct explanation, it has important consequences for treatment strategies. Following disinfection, worm recovery would in any case be enabled by ingestion of worms already present in the environment. If, however, these could be removed, then according to the model, treatment would be permanent, and recurrence would be entirely due to immigration. Indeed, this is also true for the classical Anderson–May ordinary differential equation model. Also, observed parameters of the distribution then suggest that immigration is as important as local egg production. If quarantining could be introduced, the immigration rate ν_e and thus also ν would be reduced, and (50) then suggests that the distribution would become much sparser, and eradication more likely.

The alternative is that the negative binomial distribution arises because of the variability in uptake rates in the population, suggesting that immigration may be infrequent. In that case, infection recurrence is most likely due to the continuing presence postinoculation of eggs in the soil.

An interesting question for the Anderson–May model is why the inferred value of the transmission coefficient β is so low. $\beta = \beta' d_3$ is the product of two terms, the uptake rate β' and the survival probability in the body. From our discussion above, the constraint that $R_0 \sim 1$ requires $\beta\lambda \sim 10^{-5}$ d^{-2}, so that even if we take the uptake objects to be infected stools (thus replacing λ by stool production rate, say 1 day^{-1}), we would then need $\beta \sim 10^{-5}$ day^{-1}. It seems difficult to see how we could have stool contact rate $\beta' < 10^{-2}$, say, in which case the immune loss rate would need to be $d_3 \lesssim 10^{-3}$. While that seems entirely reasonable on the basis of an effective immune response, it raises the question why the immune response is not able to completely eradicate the infection; nor is it consistent with our much higher assumed value of d_3 in Table 1. Questions such as this go beyond the Anderson–May model, but are central to the more general question as to why infectious diseases typically have a value of $R_0 \sim 1$, and what mechanism enables this.

Acknowledgments A. C. F. acknowledges the support of the Mathematics Applications Consortium for Science and Industry (www.macsi.ul.ie) funded by the Science Foundation Ireland Grant 12/1A/1683.

Appendix

Solution of (1)

We write (1) in characteristic form:

$$\begin{aligned} \dot{m} &= 1, \quad \dot{E} = -\mu_{\mathrm{e}} E, \\ \dot{m} &= 1, \quad \dot{H} = -\mu_{\mathrm{h}} H, \\ \dot{m} &= 1, \quad \dot{I} = -\mu_i I, \end{aligned} \tag{56}$$

with the boundary conditions (2) taking the parametric form

$$t = \tau, \quad E = E_0(\tau), \quad H = \beta' L(\tau), \quad I = H(\tau, \tau_3) \quad \text{at} \quad m = 0 \tag{57}$$

(these give the solutions for $t > m$; for $t < m$ we would use an initial condition at $t = 0$, but since m is finite, this just produces a transient which washes through the system). The solution of (56) and (57) is given parametrically by

$$\begin{aligned} &m = t - \tau, \quad E = E_0(\tau)\mathrm{e}^{-\mu_{\mathrm{e}}(t-\tau)}, \quad H = \beta' L(\tau)\mathrm{e}^{-\mu_h(t-\tau)}, \\ &I = H(\tau, \tau_3)\mathrm{e}^{-\mu_i(t-\tau)}, \end{aligned} \tag{58}$$

whence we obtain (3).

References

Adler FR, Kretzschmar M (1992) Aggregation and stability in parasite-host models. Parasitology 104:199–205

Anderson RM (1978) Regulation and stability of host-parasite population interactions. I. Regulatory processes. J Anim Ecol 47:219–249

Anderson RM, May RM (1985) Helminth infections of humans: mathematical models, population dynamics, and control. Adv Parasitol 24:1–101

Anderson RM, May RM (1991) Infectious diseases of humans: dynamics and control. Oxford University Press, Oxford

Anderson RM, Medley GF (1985) Community control of helminth infections of man by mass and selective chemotherapy. Parasitology 90:629–660

Anderson R, Hollingsworth TD, Truscott J, Brooker S (2012) Optimisation of mass chemotherapy to control soil-transmitted helminth infection. Lancet 379:289–290

Anderson TJC, Zizza CA, Leche GM, Scott ME, Solomons NW (1993) The distribution of intestinal helminth infections in a rural village in Guatemala. Mem Inst Oswaldo Cruz 88:53–65

Anderson TJC, Romero-Abal ME, Jaenike J (1995) Mitochondrial DNA and Ascaris microepidemiology: the composition of parasite populations from individual hosts, families and villages. Parasitology 110:221–229

Barbour AD, Kafetzaki M (1991) Modelling the overdispersion of parasite loads. Math Biosci 107:249–253

Bartlett MS (1960) Stochastic population models in ecology and epidemiology. Methuen, London

Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, Hotez PJ (2006) Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet 367:1521–1532

Betson M, Nejsum P, Llewellyn-Hughes J, Griffin C, Atuhaire A, Arinaitwe M, Adriko M, Ruggiana A, Turyakira G, Kabatereine NB, Stothard JR (2012) Genetic diversity of Ascaris in southwestern Uganda. Trans R Soc Trop Med Hyg 106:75–83

Bottomley C, Isham V, Basáñez M-G (2005) Population biology of multispecies helminth infection: interspecific interactions and parasite distribution. Parasitology 131:417–433

Cairncross S, Blumenthal U, Kolsky P, Moraes L, Tayeh A (1996) The public and domestic domains in the transmission of disease. Trop Med Int Health 1:27–34

Cornell SJ (2010) Modelling parasite transmission and control. In: Michael E, Spear RC (eds) Modelling stochastic transmission processes in helminth infections, vol 673. Adv. Exper. Med. Biol. Landes Bioscience, Austin, Texas, pp 66–78

Cort WW, Stoll NR (1931) Studies on Ascaria lumbricoides and Tricuris trichiura in China. Am J Epidemiol 14:655–689

Cox FEG (2002) History of human parasitology. Clin Microbiol Revs 15:595–612

Crompton DWT (2001) Ascaris and ascariasis. Adv Parasitol 48:285–375

Forrester JE, Scott ME, Bundy DAP, Golden MHN (1988) Clustering of Ascaris lumbricoides and Trichuris trichiura infections within households. Trans R Soc Trop Med Hyg 82:282–288

Gaba S, Cabaret J, Ginot V, Silvestre A (2006) The early drug selection of nematodes to anthelmintics: stochastic transmission and population in refuge. Parasitology 133:345–356

Hadeler KP, Dietz K (1983) Nonlinear hyperbolic partial differential equations for the dynamics of parasite populations. Comp Math Appls 9:415–430

Hall A, Holland C (2000) Geographical variation in Ascaris lumbricoides fecundity and its implications for helminth control. Parasitol Today 16:540–544

Herbert J, Isham V (2000) Stochastic host-parasite interaction models. J Math Biol 40:343–371

Holland C (ed) (2013) Ascaris: the neglected parasite. Elsevier, The Amsterdam

Hollingsworth TD, Pulliam JRC, Funk S, Truscott JE, Isham V, Lloyd AL (2015) Seven challenges for modelling indirect transmission: vector-borne diseases, macroparasites and neglected tropical diseases. Epidemics 10:16–20

Isham V (1995) Stochastic models of host-macroparasite interaction. Ann Appl Prob 5:720–740

Jacquez JA, O'Neill P (1991) Reproduction numbers and thresholds in stochastic epidemic models I. Homogeneous populations. Math Biosci 107:161–186

Jia T-W, Melville S, Utzinger J, King CH, Zhou X-N (2012) Soil-transmitted helminth reinfection after drug treatment: a systematic review and meta-analysis. PLoS Negl Trop Dis 6(5):e1621

Kostitzin VA (1934) Symbiose, parasitisme et évolution. Hermann et Cie, Paris

Kretschmar M, Adler FR (1993) Aggregated distributions in models for patchy populations. Theor Popul Biol 43:1–30

May RM, Anderson RM (1978) Regulation and stability of host-parasite population interactions. II. Regulatory processes. J Anim Ecol 47:249–267

Otto GF, Cort WW, Keller AE (1931) Environmental studies of families in Tennessee infested with Ascaris, Tricuris and hookworm. Am J Epidemiol 14:156–193

Pullan RL, Smith JL, Jasrasaria R, Brooker SJ (2014) Global numbers of infection and disease burden of soil transmitted helminth infections in 2010. Parasit Vectors 7:37

Scott ME (2008) Ascaris lumbricoides: a review of its epidemiology and relationship to other infections. Ann Nestlé (Engl) 66:7–22

Stoll NR (1947) This wormy world. J Parasitol 33:1–18

Tallis GM, Leyton M (1966) A stochastic approach to the study of parasite populations. J Theor Biol 13:251–260

Tallis GM, Leyton MK (1969) Stochastic models of populations of helminthic parasites in the definitive host. I Math Biosci 4:39–48

Truscott JE, Hollingsworth TD, Brooker SJ, Anderson RM (2014) Can chemotherapy alone eliminate the transmission of soil transmitted helminths? Parasit Vectors 7:266

Tyson E (1683) Lumbricus teres, or some anatomical observations on the roundworm bred in human bodies. Phil Trans R Soc Lond 13:154–161

Walker M, Hall A, Basáñez M-G (2010) Trickle or clumped infection process? A stochastic model for the infection process of the parasitic roundworm of humans, Ascaris lumbricoides. Int J Parasitol 40:1381–1388

WHO (1996) Report of the WHO Informal Consultation on the use of chemotherapy for the control of morbidity due to soil-transmitted nematodes in humans. World Health Organization (WHO/CTD/SIP/96.2), Geneva

Williams D, Burke G, Hendley JO (1974) Ascariasis: a family disease. J Pediatr 84:853–854