Dynamics of individual adherence to mass drug administration in a conditional probability model

Robert J. Hardwick[1,2,3]*, James E. Truscott[1,2,3], William E. Oswald[2,4], Marleen Werkman[1,2,3], Katherine E. Halliday[2,4], Rachel L. Pullan[2,4], Roy M. Anderson[1,2,3]

1 London Centre for Neglected Tropical Disease Research, Department of Infectious Disease Epidemiology, School of Public Health, St. Mary's Campus, Imperial College London, London, United Kingdom
2 The DeWorm3 Project, The Natural History Museum of London, London, United Kingdom
3 MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, United Kingdom
4 Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom

*r.hardwick@imperial.ac.uk

Abstract

We present a comprehensive framework which describes the systematic (binary) choice of individuals to either take treatment, or not for any reason, over the course of multiple rounds of mass drug administration (MDA) — which we here here refer to as 'adherence' and 'non-adherence'. This methodology can be fitted to (or informed by) program data as well as manipulated to reproduce the same adherence behaviours of past analyses, and can go beyond past analyses to describe new behaviours that have yet to be considered in the literature. Our model also has a straightforward interpretation and implementation in simulations of mass drug trials for disease transmission studies and forecasts for control through MDA. We demonstrate how our analysis may be implemented to statistically infer adherence behaviour from a dataset by applying our approach to the recent adherence data from the TUMIKIA project, a recent trial of deworming strategies in Kenya. We stratify our analysis according to age and sex, though the framework which we introduce here may be readily adapted to accomodate other categories. Our findings include the detection of past behaviour dependent non-adherence in all age groups to varying degrees of severity and particularly strong non-adherent behaviour of men of ages 30+. We then demonstrate the use of our model in stochastic individual-based simulations by running two example forecasts for elimination in TUMIKIA with the learned adherence behaviour implemented. Our results demonstrate the impact and utility of including non-adherence from real world datasets in simulations.

Author summary

Mass drug administration (MDA) is an important tool in prevention of morbidity from various neglected tropical diseases (NTDs). Due to a variety of social and medical reasons, many people will either not be offered or refuse such treatment, and if this behaviour is recurring then control measures may face a challenge to achieving their

stated goals. Learning the patterns of individual adherence or non-adherence to MDA control measures for NTDs from real world data followed by their implementation in simulated scenarios is a relatively recent development in the study of NTDs. Past analyses assessing individual adherence have informed the approach we take in this work. However, we have sought to provide a framework which encapsulates as many types of adherence behaviour as possible so that their implementation in modern simulations is streamlined effectively. Our example application to the TUMIKIA data highlights the importance of such a general framework as we find past behaviour dependence that may have been missed by other methods.

1 Introduction

Recent reviews, guidelines and work predicting the outcome of MDA to control the transmission of various NTDs all strongly stress the importance of individual adherence in successfully reaching elimination targets [1–9] (see also Ref [10] for a review on patient adherence to HIV medication). Such analyses have taken a variety of approaches in describing the strength in tendency of participants in a given MDA program with multiple rounds to either passively or actively avoid treatment in a potentially repetitive manner. The precise nomenclature for this behaviour is also debated, where terms such as 'compliance', 'adherence' and 'concordance' were all discussed for their relative merits in a recent review [7]. In this work, we shall refer to the binary choice of individuals to either take treatment, or not for any reason, over the course of multiple rounds of MDA as revealing their 'adherence' or 'non-adherence'. Ultimately, the effect that this behaviour has on the success or failure of control through MDA is equally unambiguous.

In this paper, we develop a general approach to describe individual adherence or non-adherence to MDA, into which past literature approaches (or future ones) may be incorporated. Our principle intention is to provide a framework within which as many behaviours as possible are captured so that computational modelling approaches are more flexible. To illustrate how our methodology may be implemented and interpreted in practice, we apply it to the TUMIKIA project: a recent cluster randomised, controlled trial in Kwale County, Kenya [11–13].

2 Adherence models

2.1 Model definitions

There are a range of studies of treatment adherence in the literature and models of adherence are included in micro-simulations of disease control strategies across a number of diseases. See, e.g., Refs [1, 3–8]. In this section, we will lay out a general model for the treatment adherence across multiple rounds in an MDA intervention program. In Appendix S3, we discuss other implementations of adherence in models and how they fit within our general framework.

At the level of an individual involved in an MDA treatment program, we model adherence as a binary choice, made at each round of MDA, of whether to accept treatment or not. We associate a probability with this choice, making each round a Bernoulli trial for each individual. The probability of accepting treatment is a combination of personal choice and also an individual's access to treatment.

We identify three main ways in which the probability of adherence can vary in a population over the course of an MDA intervention.

1. Dependence on past behaviour: An individual's probability of adhering in the current round may depend on their individual history of adherence in past rounds.

For example, someone may feel that being treated last year makes it less important to receive treatment in the current round. Alternatively, they may have been put off treatment due to side effects from initially taking the drug (e.g. praziquantel to treat schisosomiasis makes some children suffer from headaches, dizziness, stomach pain, nausea, or tiredness)

2. Time dependence: everyone involved in the trial may be subject to global influences that change over time. For example, enthusiasm or funding for the trial may drop as it proceeds, or unforeseen sociological or political events may change people's desire to adhere to the program. This will result in the probability of treatment in a given round for a given individual being explicitly dependent on time and is distinct from dependence on past behaviour, which will also result implicitly in the probability of treatment for an individual changing over time.

3. Population-level heterogeneity: the probability of adherence may vary across the population. That is, individuals may have a personal probability of adherence that they retain across multiple rounds of the intervention. In this case, the probability of adherence will have a distribution across the population. Typically, population-level heterogeneity may be strongly correlated with covariates such as sex or age group, in which case it can be represented by a stratification (or 'binning') of the population into sub-groups, each with their own adherence probability.

In reality, any model of adherence might include one or more of these sources of variability, or none at all in the default case in which the adherence probability is constant across all individuals and all treatment rounds and doesn't depend on past history. For purposes of illustration, we can create a tree of possible model types based on the possible sources of variability (see Fig 1). Models can be stratified into types which have some degree of dependence on the past behaviour of individuals and those that do not. Within each group, there are models with and without population heterogeneity in adherence (heterogeneous and homogeneous populations, respectively) and those with and without time-dependent adherence probabilities (time-dependent and independent, respectively).

The distinctions above are of critical importance as it is possible, e.g., for a treatment program to suffer severely from past behaviour dependent non-adherence without any apparent heterogeneity in adherence within the population. They also allow us to categorise and clarify models of adherence already described in the literature, which include Refs [1, 4–8].

Several models of MDA treatment programs employ an adherence model developed by Plaisier in the context of onchocerciasis (the Plaisier model) [8]. The Plaisier model assigns a probability of adherence to each individual which they then retain for the duration of the MDA program [14, 15]. As such, this model would be characterised by us as a heterogeneous population, time-independent model with no explicit individual dependence on past behaviour. For the interested reader, we discuss the relationship of the Plaisier model (and others [5]) to our categorisation of adherence models with more detail in Appendix S3.

We shall now proceed with a technical overview to the three categories of adherence illustrated by Fig 1 in Secs 2.2, 2.3 and 2.4, respectively. Additional technical details and calculations may also be found in Appendix S1.

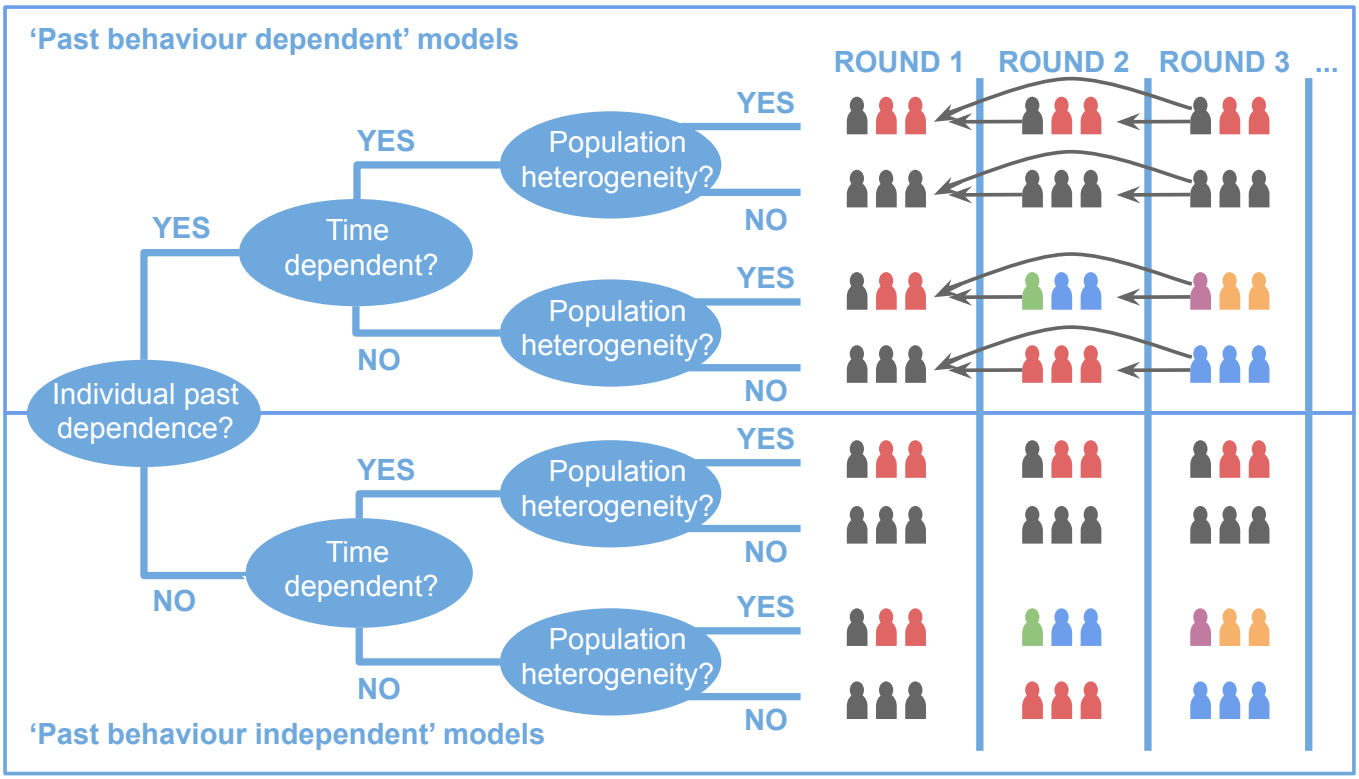

Fig 1. A decision tree illustrating the possible classes of behaviour which may be characterised in adherence models.

2.2 Individual past behaviour dependent adherence

2.2.1 Basic model

Before we introduce the notion of dependence on past behaviour in our model, it shall be instructive to describe the situation where it is absent. When the probability of an individual taking treatment is not dependent upon any of their past behaviour, then it is simply given by the coverage c_n in each round n of MDA. In the absence of population heterogeneity, this probability would then apply to all individuals within a given cohort — a case which corresponds to either the 6th or 8th row of Fig 1, depending on whether the coverage changes over time, i.e., between rounds.

For the case of a homogeneous population with a dependence on past behaviour between successive rounds, let us consider the dynamics of a single individual. In this case, the model becomes a simple Markov chain. The possible patterns of adherence behaviour by an individual after two successive rounds of treatment are TT, TF, FT and FF, where T and F are accepting and avoiding treatment, respectively. Let the probability of accepting treatment in the first round be set to $P(\mathrm{T}) = \alpha$. In round 2, we now fix the conditional probability of getting treated, given treatment in the first round as $P(\mathrm{T}'|\mathrm{T}) = \beta$. Let us also set a corresponding conditional probability for not being treated in the second round given that there was no treatment in the first round as $P(\mathrm{F}'|\mathrm{F}) = 1 - \gamma$. As such, β and γ are now measures of consistent behaviour.

To avoid unnecessary repetition, we shall once again use the notation p_n to denote the probability of treatment in the n-th round. Assuming that the conditional probabilities β and γ are constant, the time-independent Markov model may be mapped

to the following recursion relation

$$p_n = \beta p_{n-1} + (1-\gamma)(1-p_{n-1})\,. \tag{1}$$

Defining the system state vector as

$$\mathbf{p}_n = \begin{pmatrix} p_n \\ 1-p_n \end{pmatrix}, \tag{2}$$

we may rewrite Eq (1) above in the form $\mathbf{p}_n = \mathbf{M}\,\mathbf{p}_{n-1}$ where we have defined the following transition matrix

$$\mathbf{M} \equiv \begin{pmatrix} \beta & 1-\gamma \\ 1-\beta & \gamma \end{pmatrix}. \tag{3}$$

The eigenvalues and eigenvectors of $\mathbf{M}$ are given by

$$\lambda' = 1\,, \quad \mathbf{v}' = \frac{1}{2-\gamma-\beta}\begin{pmatrix} 1-\gamma \\ 1-\beta \end{pmatrix}, \tag{4}$$

$$\lambda = \beta+\gamma-1\,, \quad \mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{5}$$

where $\mathbf{v}'$ is normalised to sum to 1. As long as $|\beta+\gamma|<1$, the system will relax to the $\mathbf{v}'$ state, which has a unit eigenvalue, giving a long-term probability of treatment of[1]

$$q \equiv \frac{1-\gamma}{2-\gamma-\beta}\,. \tag{6}$$

In addition, the relaxation will be oscillatory if $\lambda<0$. In Appendix S1, we demonstrate how to obtain the following solution to Eq (1)

$$p_n = \alpha\lambda^{n-1} + q\left(1-\lambda^{n-1}\right). \tag{7}$$

Notice that by matching p_n to the coverage of treatment in a given population, one may directly compare the impact of adherence models such as Eq (1) to those with past behaviour independent adherence behaviour. Furthermore, by setting $\lambda=0$ in Eq (7) one finds a model for past behaviour independent adherence that is time-independent, i.e., $p_n = \beta = 1-\gamma$.

Any sequence of treatments can be seen as a set of alternating adherent and non-adherent runs. A key statistic in the context of preventive chemotherapy is the run length (in rounds) over which an individual complies or fails to adhere. For an adherence run, this is the number of consecutive treatment adherences, given an initial adherence. This can also be thought of as the first passage time to failure. Since the $P(\mathrm{T}'|\mathrm{T}) = \beta$ is constant, the run length is distributed according to a geometric distribution, with

$$P(n_{\mathrm{T}}) = \beta^{n-1}(1-\beta)\,, \quad \mathrm{E}(n_{\mathrm{T}}) = \frac{1}{1-\beta} = \frac{1}{1-\lambda}\frac{1}{1-q}\,. \tag{8}$$

Correspondingly, for a run of failures,

$$P(n_{\mathrm{F}}) = \gamma^{n-1}(1-\gamma)\,, \quad \mathrm{E}(n_{\mathrm{F}}) = \frac{1}{1-\gamma} = \frac{1}{1-\lambda}\frac{1}{q}\,. \tag{9}$$

Any long run of treatment choices by an individual will breakdown into an alternating sequence of F and T runs. Hence, the probability of a round chosen at random being T, $P(\mathrm{T})$, is

$$P(\mathrm{T}) = \frac{\mathrm{E}(n_{\mathrm{T}})}{\mathrm{E}(n_{\mathrm{F}}) + \mathrm{E}(n_{\mathrm{T}})} = q\,, \tag{10}$$

[1] One may confirm this trivially by satisfying $p_n = p_{n-1}$ and Eq (1) simultaneously.

Table 1. Probability table corresponding to two successive rounds of treatment.

Behaviour	Probability
TT	$\alpha\beta$
TF	$\alpha(1-\beta)$
FT	$(1-\alpha)(1-\gamma)$
FF	$(1-\alpha)\gamma$

matching the conclusions drawn earlier from the eigenvalues and eigenvectors of Eq (5). From Eqs (8) and (9), it is clear that as λ approaches 1, the length of both success and failure runs grows as $1/(1-\lambda)$. In the absence of past behaviour dependence, $\lambda = 0$ and the adherent and non-adherent run lengths are given by $1/(1-q)$ and $1/q$, respectively.

2.2.2 Statistical inference from data

Let us now only consider two rounds of treatment to illustrate how we may calculate the important quantities for statistical inference of the time-independent Markov model from a real dataset. Recall that the possible patterns of adherence behaviour by an individual after two successive rounds of treatment are TT, TF, FT and FF, where T and F are accepting and avoiding treatment, respectively. Once again, let: the probability of treatment in the first round be set to $P(\mathrm{T}) = \alpha$; the conditional probability of getting treated in the second round, given treatment in the first round be set to $P(\mathrm{T}'|\mathrm{T}) = \beta$; and the conditional probability for not being treated in the second round given that there was no treatment in the first round be set to $P(\mathrm{F}'|\mathrm{F}) = 1-\gamma$.

In this model, there are effectively 4 types of people with probabilities and behaviours, mapped out in Table 1. Using the probability table, one may infer directly that the the likelihood $\mathcal{L}(D|\boldsymbol{\theta})$ of the data $D = \{N_{\mathrm{T}}, N_{\mathrm{F}}, N_{\mathrm{TT}}, N_{\mathrm{TF}}, N_{\mathrm{FT}}, N_{\mathrm{FF}}\}$ — where N_{T} and N_{F} are the number treated and not treated in the first round and N_{TT} is the number treated in the first and the second rounds, etc. — is a multinomial distribution, where $\boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}$ is now a 3-vector defined over the model parameter space $\boldsymbol{\theta} = (\alpha, \beta, \gamma)$ within the prior domain $\Omega_{\boldsymbol{\theta}} = \{\boldsymbol{\theta} \,|\, \theta_1 \in [0,1]\,, \theta_2 \in [0,1]\,, \theta_3 \in [0,1]\}$. The multinomial can then be factored into independent functions of the three parameters, such that

$$\mathcal{L}(D|\boldsymbol{\theta}) = \alpha^{N_{\mathrm{T}}}(1-\alpha)^{N_{\mathrm{F}}}\beta^{N_{\mathrm{TT}}}(1-\beta)^{N_{\mathrm{TF}}}\gamma^{N_{\mathrm{FF}}}(1-\gamma)^{N_{\mathrm{FT}}}\,. \tag{11}$$

The likelihood above is effectively three independent beta distributions, one in each of the parameters, such that the posterior distribution $\mathcal{P}(\boldsymbol{\theta}|D)$ becomes

$$\mathcal{P}(\boldsymbol{\theta}|D) = \frac{1}{\mathcal{E}}\,\mathrm{Beta}(\alpha; N_{\mathrm{T}}+1, N_{\mathrm{F}}+1)\mathrm{Beta}(\beta; N_{\mathrm{TT}}+1, N_{\mathrm{TF}}+1)\mathrm{Beta}(\gamma; N_{\mathrm{FF}}+1, N_{\mathrm{FT}}+1)\,, \tag{12}$$

where we have assumed a flat prior $\pi(\boldsymbol{\theta}) \propto 1$ to derive the following Bayesian evidence normalisation

$$\mathcal{E} = \frac{\Gamma(N_{\mathrm{TT}}+1)\Gamma(N_{\mathrm{TF}}+1)\Gamma(N_{\mathrm{FT}}+1)\Gamma(N_{\mathrm{FF}}+1)}{(N_{\mathrm{T}}+1)(N_{\mathrm{F}}+1)\Gamma(N+2)}\,, \tag{13}$$

and $N = N_{\mathrm{T}} + N_{\mathrm{F}}$ is defined as the total number of individuals.

Note here that Eqs (11) and (13) may be generalised to the case where n rounds of treatment have taken place. We have provided these expressions in Appendix S1.

2.3 Time-dependent adherence and more general behaviour

2.3.1 Introducing the choice matrices

A significant generalisation of Eq (1) introduces the lower triangular matrices, with elements $C^{\mathrm{T}}_{nn'}$ and $C^{\mathrm{F}}_{nn'}$ corresponding to the conditional probabilities of treatment and non-treatment in round n given treatment and non-treatment in round n', respectively, such that

$$p_n = \sum_{n'=1}^{n-1} \left[C^{\mathrm{T}}_{nn'} p_{n'} + C^{\mathrm{F}}_{nn'} (1 - p_{n'}) \right] . \tag{14}$$

We shall hereafter refer to the above matrices as 'choice matrices'. In the proceeding sections we shall demonstrate that the model parameterisation defined in Eq (14) is extremely general — encapsulating all of the possible adherence behaviours illustrated in Fig 1.

2.3.2 Lower diagonal choice matrices: the time-dependent Markov model

When the only nonzero elements of the choice matrices in Eq (14) are along the their lower diagonals, i.e., such that only $C^{\mathrm{T}}_{n\,n-1} = \beta_{n\,n-1} \neq 0$ and $C^{\mathrm{F}}_{n\,n-1} = 1 - \gamma_{n\,n-1} \neq 0$, the system is described by a time-dependent Markov process with recursion relation

$$p_n = \beta_{n\,n-1} p_{n-1} + (1 - \gamma_{n\,n-1})(1 - p_{n-1}) . \tag{15}$$

Following a similar argument to the one used in solving the homogeneous Markov model (which is provided in detail in Appendix S1), we may obtain a solution to Eq (15), which is given by

$$p_n = \alpha \prod_{n'=2}^{n} \omega_{n'} + \sum_{n''=2}^{n} (1 - \gamma_{n''n''-1}) \prod_{n'=n''}^{n} \omega_{n'} , \tag{16}$$

where we have defined an important new quantity

$$\omega_n \equiv \beta_{n\,n-1} + \gamma_{n\,n-1} - 1 . \tag{17}$$

Notice, firstly, that when $w_n = 0$ the system reverts to a time-dependent past behaviour independent adherence model, i.e., without past behaviour dependence such that $p_n = \beta_{n\,n-1} = 1 - \gamma_{n\,n-1}$. By analogy with the time-independent Markov model (where $\omega_n = \lambda$ in Eq (5)), $|w_n| \neq 0$ signals the presence of some degree of past behaviour dependent adherence behaviour. In more detail, for successive rounds over which $w_n > 0$, the system will relax towards the steady state and when $w_n < 0$ this will be accompanied by oscillatory behaviour. Note also that w_n may act as an indicator for the severity of adherence and non-adherence behaviour in the system — where larger absolute values for w_n approaching a maximum of 1 will indicate increasingly past behaviour dependent behaviour.

At the extrema of: $w_n = 1$, individuals repeat their past behaviour exactly and indefinitely, i.e., TTTTT... and FFFFF...; and $w_n = -1$, individuals repeat the opposite of their past behaviour exactly and indefinitely, i.e., TFTFT... and FTFTF.... The value of w_n is therefore a useful indicator for the type of adherence behaviour in the relatively general description of time-dependent Markov models. We shall use this parameter to illustrate our results from the TUMIKIA project in Sec 3.

2.3.3 General choice matrices: non-Markovian models

The most general set of causal adherence models described by Eq (14) have choice matrices which take the form

$$\boldsymbol{C}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ C^{\mathrm{T}}_{n\,n-1} & 0 & 0 & \dots \\ C^{\mathrm{T}}_{n\,n-2} & C^{\mathrm{T}}_{n-1\,n-2} & 0 & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix} \qquad \boldsymbol{C}^{\mathrm{F}} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ C^{\mathrm{F}}_{n\,n-1} & 0 & 0 & \dots \\ C^{\mathrm{F}}_{n\,n-2} & C^{\mathrm{F}}_{n-1\,n-2} & 0 & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}, \quad (18)$$

where 'non-Markovian' behaviour in the n-th round clearly corresponds to a past behaviour dependence between rounds which exceeds the immediate last round, i.e., $C^{\mathrm{T,F}}_{n\,n-m} \neq 0$ where $m > 1$.

Notice that all of the adherence models that we have identified in this work may be categorised by various constraints on the elements of the choice matrices introduced in Eq (14). For completeness and reference, these are

1. Past behaviour independent adherence that is time-independent: $\forall n > 1$ only $C^{\mathrm{T,F}}_{n\,n-1} \neq 0$, $C^{\mathrm{T}}_{n\,n-1} = C^{\mathrm{F}}_{n\,n-1} = c$ and $p_1 = c$, giving one degree of freedom multiplied by the number of independent bins for population-level heterogeneity.

2. Past behaviour independent adherence that is time-dependent: $\forall n > 1$ only $C^{\mathrm{T,F}}_{n\,n-1} \neq 0$, $C^{\mathrm{T}}_{n\,n-1} = C^{\mathrm{F}}_{n\,n-1} = c_n$ and $p_1 = c_1$, giving n degrees of freedom multiplied by the number of independent bins for population-level heterogeneity.

3. Markovian past behaviour dependent adherence that is time-independent: $\forall n > 1$ only $C^{\mathrm{T,F}}_{n\,n-1} \neq 0$, $C^{\mathrm{T}}_{n\,n-1} = \beta$, $C^{\mathrm{F}}_{n\,n-1} = 1 - \gamma$ and $p_1 = \alpha$, giving 3 degrees of freedom multiplied by the number of independent bins for population-level heterogeneity.

4. Markovian past behaviour dependent adherence that is time-dependent: $\forall n > 1$ only $C^{\mathrm{T,F}}_{n\,n-1} \neq 0$, $C^{\mathrm{T}}_{n\,n-1} = \beta_{n\,n-1}$, $C^{\mathrm{F}}_{n\,n-1} = 1 - \gamma_{n\,n-1}$ and $p_1 = \alpha$, giving $2n - 1$ degrees of freedom multiplied by the number of independent bins for population-level heterogeneity.

5. Non-Markovian past behaviour dependent adherence that is time-dependent: $\forall n > 1$ and $\forall n' < n$ only $C^{\mathrm{T,F}}_{nn'} \neq 0$ and $p_1 = \alpha$, giving $1 + n(n-1)$ degrees of freedom multiplied by the number of independent bins for population-level heterogeneity.

2.3.4 Statistical inference from data

The universality of the choice matrix approach suggest that it is an ideal candidate for parameterisation of the inference problem from data and model comparison. Let the data now correspond to a set of n-vectors $D = \{\boldsymbol{X}\}$ where each individual's adherence or non-adherence behaviour in the n-th round is recorded, such that $X_n = \mathrm{T}, \mathrm{F}$. Using Eq (14) the full generalisation of the likelihood (which supports all of the possible adherence models) becomes

$$\mathcal{L}(D|\boldsymbol{\theta}) = \prod_{\forall X_n \in D} \prod_{n'=1}^{n} \left\{ \sum_{n''=1}^{n'-1} \left[C^{\mathrm{T}}_{nn'} \mathbb{1}_{X_{n'}=\mathrm{T}} + C^{\mathrm{F}}_{nn'} \mathbb{1}_{X_{n'}=\mathrm{F}} \right] \right\}, \quad (19)$$

where $\mathbb{1}_A$ denotes an indicator function which takes value unity when condition A is satisfied, else it vanishes.

The large number of available degrees of freedom in Eq (19) motivates a systematic approach to inferring the choice matrix components from a given set of data. We elect to consider models which isolate the many degrees of freedom by constructing scenarios where past behaviour dependent adherence only occurs for a single round and is past behaviour dependent to only one other round — all other degrees of freedom are hence set to those corresponding to time-dependent past behaviour independent adherence, i.e. $C^{\mathrm{T}}_{nn'} = C^{\mathrm{F}}_{nn'} = c_n$. The likelihoods and Bayesian evidence normalisations for this more restricted set of models are calculated in Appendix S1.

2.4 Population heterogeneity in adherence (in brief)

The probability of adherence may vary across a population of individuals. The first possible form that this heterogeneity may take can be attributed to age, gender and other social factors. In such cases, stratification of the population into separate cohorts for study is an appropriate tool to quantify this variation. We primarily take this approach to population heterogeneity in Sec 3 and our analysis of data in Appendix S2.

The second possible form that population heterogeneity could take may not be immediately attributable to social groupings. In such situations, it is intuitive to consider that adherence probability for an individual is drawn from a distribution which applies to the entire population or cohort of study. This approach is the same as used in other models in the literature (see Appendix S3 for more details). We shall now briefly elaborate on how one might include this form of heterogeneity in the formalism we have introduced in this work through a simple, generic example. We leave further specific applications of this approach to a future publication in progress.

To illustrate the generic effect of the population heterogeneity described above on our individual adherence probabilities, let us consider the time-independent Markov model we introduced earlier. The long-term probability of adherence q in Eq (7) may itself be randomly drawn from a population heterogeneity distribution $P_{\mathrm{pop}}(q)$ for an individual within the specified cohort of study, such that $q \sim P_{\mathrm{pop}}(q)$. Note also that λ in Eq (7) need not vary between individuals at the same time. Using the results given in Eqs (8) and (9) for the same model one may deduce that the mean adherent and non-adherent run lengths are generically modified by

$$\mathrm{E}(n_{\mathrm{T}}) = \frac{1}{1-\lambda}\mathrm{E}_{\mathrm{pop}}\left(\frac{1}{1-q}\right) \tag{20}$$

$$\mathrm{E}(n_{\mathrm{F}}) = \frac{1}{1-\lambda}\mathrm{E}_{\mathrm{pop}}\left(\frac{1}{q}\right) , \tag{21}$$

where $\mathrm{E}_{\mathrm{pop}}(\cdot)$ denotes taking an expectation value with the distribution $P_{\mathrm{pop}}(q)$. Hence, depending on the choice for this distribution, one may either shorten or lengthen the mean run lengths across the population accordingly. Note that due to the fact that q is a probability, a natural candidate for $P_{\mathrm{pop}}(q)$ is the beta distribution.

3 Results

3.1 Overview of statistical analysis

In the four rounds of individual adherence data from the TUMIKIA project, we have split the population who traced the different possible behaviours into the standard pre-school-aged children (pre-SAC, ages 0-4), school-aged children (SAC, ages 5-14) and other adult age categories. Importantly, these age categories were assigned at the beginning of the trial and hence — particularly in the case of pre-SAC — the effect of 'ageing-out' of each category must be considered on the overall adherence behaviour.

Table 2. A measure of how past behaviour dependent the adherent and non-adherent behaviour of individuals is in the n-th round of treatment, $\omega_n \equiv \beta_{n\,n-1} + \gamma_{n\,n-1} - 1$, which was introduced in Eq (17). This value is given for each age group and sex inferred from the TUMIKIA project dataset and is computed using the maximum likelihood values for the conditional probabilities.

Age group	ω_2 (Male)	ω_3 (Male)	ω_4 (Male)	ω_2 (Female)	ω_3 (Female)	ω_4 (Female)
Pre-SAC	0.294	0.241	0.051	0.237	0.250	0.046
SAC	0.209	0.141	0.027	0.213	0.226	0.021
15-29	0.228	0.210	0.111	0.223	0.182	0.066
30-49	0.259	0.308	0.268	0.223	0.174	0.118
50+	0.244	0.286	0.259	0.195	0.181	0.139

Using the symbolic representation for behaviours which we introduced in Sec 2 — accepting treatment in a given round is denoted by a 'T', whereas not accepting treatment in a given round is denoted by an 'F'.

It is important to comment here on the validity of interpretating the inferences made as directly due to individual behaviour patterns using the TUMIKIA project adherence data [13]. An important caveat to this interpretation is that, for various reasons, some individuals were not offered treatment and were hence automatically accounted for as 'non-adherent' within the data. The impact of these individuals to the success of the MDA program is the same as if they had directly refused treatment, and hence the practical use of inferring this pattern of adherence for simulation forecasts of MDA outcome is still clear. Despite this fact, however, we cannot fairly discriminate this behaviour pattern from simply not being offered treatment in the present data.

Using the same age categories as before and the likelihoods for the adherence models which have been derived in Appendix S1, in Appendix S2 we demonstrate the applicability of our model for adherence to statistical inference by performing a thorough analysis of the TUMIKIA project dataset. We note that such an analysis has already been performed in Ref [13], hence this analysis does not consistute the novelty of the work presented here but instead is indended to illustrate the application of our mathematical model. We have a detailed description of these findings in Appendix S2 but we provide a short written summary of our general conclusions below.

In Table 2 we have provided the ω_n values, calculated using Eq (17), for each age group and sex inferred from the TUMIKIA project dataset. This value was shown in Sec 2 to be an indicator of how past behaviour dependent the adherent and non-adherent behaviour of individuals is as a response to MDA treatment. We can see quite clearly from Table 2 that a degree of past behaviour dependent non-adherence is indeed present in all the age groups, with the exception of the final round ω_4 values for those in the pre-SAC (which have mostly aged into SAC by this point) and SAC categories — which is to be expected due to the nature of school-based MDA, and is an effect which is found and explained in more detail by Ref [13]. Table 2 also shows that the most past behaviour dependent non-adherent age group and sex appears to be males aged 30+.

In addition to these results, Eq (1) appears to provide a good descriptive model for many of the past behaviour dependent non-adherent age groups and sexes, but this model must be extended to an equivalent time-dependent one — see Eq (15) — in order to describe other cases.

3.2 The impact of adherence on forecasts

In this section we illustrate the impact of adherence, as described by our mathematical model, on the predictions made by simulations for the outcome of MDA on the chances

Table 3. The positive predictive value (PPV) for elimination evaluated by fully age-structured stochastic individual-based simulations of hookworm (with adult worm and eggs/larvae mortality rates set to $\mu_1 = 0.5$ and $\mu_2 = 26.0$ per year, respectively and the density dependent fecundity factor is set to $\gamma = 0.01$, as considered in Ref [16]) with two different clustered community types specified by the TUMIKIA transmission parameters inferred from the baseline epidemiological data in Ref [16]. The parameters quoted are the endemic prevalence P, parasite aggregation parameter k, basic reproduction number R_0 and cluster population number N, where the age profiles are all assumed to be exactly flat for simplicity. The PPVs are evaluated after 100 years post-cessation of MDA and are quoted assuming either past behaviour independent adherence (i.e., simple time-dependent coverage in age groups) or the adherence behaviour inferred from our model in this paper for the TUMIKIA project (see Appendix S2).

Cluster type (see Ref [16])	PPV (Past behaviour independent adherence)	PPV (TUMIKIA adherence)
$(P, k, R_0, N) = (0.15, 0.05, 2.1, 1000)$	0.582	0.148
$(P, k, R_0, N) = (0.4, 0.15, 2.5, 1000)$	0.902	0.672

of elimination by considering the case study of TUMIKIA. The adherence models learned from the statistical analysis in the previous section are applied to stochastic individual-based simulations of hookworm transmission for two typical clustered communities that were treated in the TUMIKIA project [16]. The resulting effect that the known TUMIKIA adherence has on the positive predictive value (PPV) for elimination of hookworm in these two cluster is given in Table 3, where an equivalent PPV assuming past behaviour independent adherence is also provided for direct comparison in each case.

From Table 3 it is immediately clear that although there is relatively high coverage of MDA in the TUMIKIA project [11, 12], the presence of observed past behaviour dependent non-adherence has an important effect on the PPVs for elimination, shifting the chances of hookworm elimination in both clusters lower by 43% and 23%, respectively, when compared to the standard forecasts which assume past behaviour independent adherence.

4 Discussion and conclusions

Despite the causes for non-adherent behaviour being varied, the net effect on the outcome of MDA interventions aiming to eliminate NTDs is much the same, and hence the predicted outcomes from simulation studies should also reflect this. It is for this reason that in this paper we have been able to develop a simple but comprehensive framework which describes the systematic binary choice of individuals to either take treatment, or not for any reason, over the course of multiple rounds of mass drug administration (MDA) — which we have referred to as 'adherence' and 'non-adherence', respectively.

In Sec 2 we introduced our models for adherence which can account for new behaviours that have yet to be considered in the literature. We have also demonstrated that they can also reproduce the same adherence behaviours of past analyses as well in Appendix S3. Our analysis further yielded an interesting parameter, ω_n, given in Eq (17), which can be used as a guide to indicate the strength of adherent or non-adherent behaviour in any given setting.

In order to illustrate our framework in the context of statistical inference from a real adherence dataset, we applied our probability model to the recently collected adherence data from the TUMIKIA project in Kenya in Sec 3. Findings from this dataset extend and support the analysis of recent work [13], which include past behaviour independent adherence/non-adherence for school-aged children (SAC) and the detection of past behaviour dependent non-adherence to treatment in nearly all other age groups and both sexes. A full description of our results and analysis is given in Appendix S2.

In Sec 3 we also commented on the validity of interpretating the inferences made as directly due to individual behaviour patterns using the TUMIKIA project adherence data. As we pointed out, for various reasons, some individuals were not offered treatment and were hence automatically accounted for as 'non-adherent' within the data. However, the impact of these individuals to the success of the MDA program is the same as if they had directly refused treatment and hence the practical use of inferring this pattern of adherence (using the formalism which we have outlined in this work) for simulation forecasts of MDA outcome is still clear. We shall leave the more direct inference of human behaviour in response to control measures through MDA to future work with other sets of data.

Using the learned adherence behaviour from the TUMIKIA dataset, we then demonstrated the use of our stochastic individual-based simulation model for STH transmission and control by MDA by running two example forecasts for the likelihood and time to the elimination of hookworm transmission with the adherence behaviour recorded in Kenya by comparison with runs that assume random adherence at each round of treatment for any given treatment coverage level. The difference between the two assumptions are striking. And show clearly how important it is to measure adherence if the outcome of any given MDA based control program is to be correctly predicted.

WHO recommendations on how best to measure the impact of MDA programs to control NTDs only advise recording patterns of treatment coverage round by round with some rough stratification by age groupings treated. No advice to Ministries of Health is given on trying to record adherence patterns. As we have demonstrated, the precise form of the adherence pattern can greatly influence the extent of needed coverage and the number of treatment rounds necessary to eliminate transmission. In part, lack of guidance is understandable, given the costs and time involved in longitudinal studies to record adherence of individuals within any given MDA program. However, given the importance of these pattens in determining control progamme impact and outcome, collecting such data should be given a higher priority even if just focused on a few sentinel sites to broadly capture the prevailing behaviours in defined settings. It is likely, the social, environmental and other influences will create some heterogeneity in adherence patterns within countries and health implementation units. Additional background research on what degree of heterogeneity exists in a given country would also be of great value. In the coming few years more data on adherence patterns will emerge from detailed research studies of MDA impact to add to the information provided by the Tumikia study [11]. These include the ongoing DW3 trial studies in India, Benin and Malawi for the control of STH [17] and the Geshiyaro study in Ethiopia for the control of STH and schistosome infections by MDA [18].

Supporting information

S1 Appendix. Extended mathematical notes.

S2 Appendix. TUMIKIA project analysis and figures.

S3 Appendix. A comparison with existing models in the literature.

Acknowledgements

The authors would like to sincerely thank Benjamin Collyer for careful reading of, and useful comments on, the manuscript. RJH, JET, MW and RMA gratefully thank the

Bill and Melinda Gates Foundation for research grant support via the DeWorm3 (OPP1129535) award to the Natural History Museum in London (http://www.gatesfoundation.org/). The views, opinions, assumptions or any other information set out in this article are solely those of the authors. All authors acknowledge joint Centre funding from the UK Medical Research Council and Department for International Development.

S1 Appendix.

Summary. In this supplementary information we derive many of the key mathematical expressions which are used and referred to in the main text.

Time-independent Markov model

Assuming that the conditional probabilities β and γ are constant, the time-independent Markov model may be mapped to the following recursion relation

$$p_n = \beta p_{n-1} + (1-\gamma)(1-p_{n-1})\,. \tag{22}$$

As in the main text, defining the system state vector as

$$\mathbf{p}_n = \begin{pmatrix} p_n \\ 1-p_n \end{pmatrix}\,, \tag{23}$$

we may rewrite Eq (22) above in the form $\mathbf{p}_n = \mathbf{M}\,\mathbf{p}_{n-1}$ where we have defined the following transition matrix

$$\mathbf{M} \equiv \begin{pmatrix} \beta & 1-\gamma \\ 1-\beta & \gamma \end{pmatrix}\,. \tag{24}$$

The eigenvalues and eigenvectors of $\mathbf{M}$ are given by

$$\lambda' = 1\,, \quad \mathbf{v}' = \frac{1}{2-\gamma-\beta}\begin{pmatrix} 1-\gamma \\ 1-\beta \end{pmatrix} \equiv \begin{pmatrix} q \\ 1-q \end{pmatrix}\,, \tag{25}$$

$$\lambda = \beta+\gamma-1\,, \quad \mathbf{v} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\,. \tag{26}$$

where $\mathbf{v}'$ is normalised to sum to 1. Given that $|\lambda| < 1$ in all realistic circumstances, it is clear from this description that $\mathbf{v}$ represents the equilibrium of the system over multiple rounds with λ defining the rate of relaxation towards it. When $\lambda = 0$, the model becomes a history-independent model in which the next round is dictated solely by its probability at that round.

In order to study the dynamics in more detail, we apply the following transformation

$$p_n \to \tilde{p}_n = p_n(\beta+\gamma-1)^{1-n}\,, \tag{27}$$

to the relation given by Eq (22), such that

$$\tilde{p}_n = \tilde{p}_{n-1} + (1-\gamma)(\beta+\gamma-1)^{1-n}\,. \tag{28}$$

Through explicit summation, Eq (28) is solved by

$$\tilde{p}_n - \tilde{p}_1 = \sum_{n'=2}^{n}(\tilde{p}_{n'} - \tilde{p}_{n'-1}) = \sum_{n'=2}^{n}(1-\gamma)(\beta+\gamma-1)^{1-n'}\,. \tag{29}$$

By reapplying the inverse transformation $\tilde{p}_n \to p_n$ to Eq (29) and identifying $\tilde{p}_1 = p_1 = \alpha$, we obtain the following solution to Eq (22)

$$\begin{aligned} p_n &= \alpha(\beta+\gamma-1)^{n-1} + \sum_{n'=2}^{n}(1-\gamma)(\beta+\gamma-1)^{n-n'} \\ &= \alpha(\beta+\gamma-1)^{n-1} + \frac{1-\gamma}{\beta+\gamma-2}\left[(\beta+\gamma-1)^{n-1}-1\right]\,. \end{aligned} \tag{30}$$

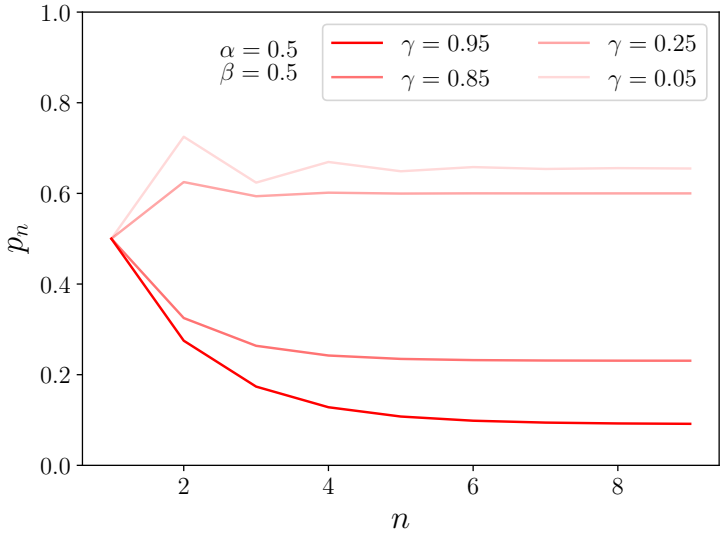

Fig 2. The probability of accepting treatment in the n-th round given by the Markovian model solution in Eq (30) for a range of γ values. The other probabilities have been fixed to $\alpha = 0.5$ and $\beta = 0.5$.

Equivalently, satisfying the dual to Eq (22) in terms of the probability of non-treatment in the n-th round $1 - p_n$, solutions to Eq (30) must also satisfy

$$1 - p_n = (1-\alpha)(\beta+\gamma-1)^{n-1} + \frac{1-\beta}{\beta+\gamma-2}\left[(\beta+\gamma-1)^{n-1} - 1\right] . \tag{31}$$

In Fig 2 we illustrate the dynamics of the system using Eq (30) with range of parameter values chosen for γ. Notice, in particular, that the system exhibits oscillation before relaxing to a steady state when γ is chosen such that the eigenvalue $\lambda = \beta + \gamma - 1 < 0$.

For another way of calculating the expected lengths of repeat adherence $\mathrm{E}(n_{\mathrm{T}})$ or non-adherence $\mathrm{E}(n_{\mathrm{F}})$ of an individual (as computed in the main text), given that they begin with the same choice in the first round, one need only fix ($\alpha = \beta$, $\gamma = 1$) or ($\alpha = 1 - \gamma$, $\beta = 1$) and take moments with Eq (30), respectively, such that

$$\begin{aligned}(\alpha=\beta,\,\gamma=1) \;\Rightarrow\; \mathrm{E}(n_{\mathrm{T}}) &= \sum_{n=0}^{\infty} n\left(1 - \frac{p_n}{p_{n-1}}\right) p_{n-1} \\ &= \sum_{n=0}^{\infty} n(1-\beta)\beta^{n-1} = \frac{1}{1-\beta} \end{aligned} \tag{32}$$

$$\begin{aligned}(\alpha=1-\gamma,\,\beta=1) \;\Rightarrow\; \mathrm{E}(n_{\mathrm{F}}) &= \sum_{n=0}^{\infty} n\left(1 - \frac{1-p_n}{1-p_{n-1}}\right)(1-p_{n-1}) \\ &= \sum_{n=0}^{\infty} n(1-\gamma)\gamma^{n-1} = \frac{1}{1-\gamma} . \end{aligned} \tag{33}$$

Time-dependent Markov model

Consider the choice matrices with elements $C^{\mathrm{T}}_{nn'}$ and $C^{\mathrm{F}}_{nn'}$ corresponding to the conditional probabilities of treatment and non-treatment in round n given treatment

and non-treatment in round n', respectively, such that

$$p_n = \sum_{n'=1}^{n-1} \left[C^{\mathrm{T}}_{nn'} p_{n'} + C^{\mathrm{F}}_{nn'} (1 - p_{n'}) \right] . \tag{34}$$

When the only nonzero elements of the choice matrices in Eq (34) are along the their lower diagonals, i.e., such that only $C^{\mathrm{T}}_{n\,n-1} = \beta_{n\,n-1} \neq 0$ and $C^{\mathrm{F}}_{n\,n-1} = 1 - \gamma_{n\,n-1} \neq 0$, the system is described by a time-dependent Markov process with recursion relation

$$p_n = \beta_{n\,n-1} p_{n-1} + (1 - \gamma_{n\,n-1})(1 - p_{n-1}) \, . \tag{35}$$

Following a similar argument to the one used in solving the homogeneous Markov case, we may obtain an implicit solution to Eq (35). Using the transformation

$$p_n \to \tilde{p}_n = \frac{p_n}{\prod_{n'=2}^{n} (\beta_{n'n'-1} + \gamma_{n'n'-1} - 1)} \, , \tag{36}$$

we once again substitute into the relation given by Eq (35), yielding

$$\tilde{p}_n = \tilde{p}_{n-1} + \frac{1 - \gamma_{n\,n-1}}{\prod_{n'=2}^{n} (\beta_{n'n'-1} + \gamma_{n'n'-1} - 1)} \, , \tag{37}$$

where Eq (37) is solved by the explicit summation

$$\tilde{p}_n - \tilde{p}_1 = \sum_{n''=2}^{n} (\tilde{p}_{n''} - \tilde{p}_{n''-1}) = \sum_{n''=2}^{n} \frac{1 - \gamma_{n''n''-1}}{\prod_{n'=2}^{n''} (\beta_{n'n'-1} + \gamma_{n'n'-1} - 1)} \, . \tag{38}$$

Using the corresponding inverse transformation to Eq (37) we hence obtain a solution to Eq (35), which is given by

$$p_n = \alpha \prod_{n'=2}^{n} (\beta_{n'n'-1} + \gamma_{n'n'-1} - 1) + \sum_{n''=2}^{n} (1 - \gamma_{n''n''-1}) \prod_{n'=n''}^{n} (\beta_{n'n'-1} + \gamma_{n'n'-1} - 1) \, . \tag{39}$$

Likelihoods and Bayesian evidence

Let the data now correspond to a set of n-vectors $D = \{\boldsymbol{X}\}$ where each individual's adherence or non-adherence behaviour in the n-th round is recorded, such that $X_n = \mathrm{T}, \mathrm{F}$. Using Eq (34) the full generalisation of the likelihood (which supports all of the possible adherence models, becomes

$$\mathcal{L}(D|\boldsymbol{\theta}) = \prod_{\forall X_n \in D} \prod_{n'=1}^{n} \left\{ \sum_{n''=1}^{n'-1} \left[C^{\mathrm{T}}_{nn'} \mathbb{1}_{X_{n'}=\mathrm{T}} + C^{\mathrm{F}}_{nn'} \mathbb{1}_{X_{n'}=\mathrm{F}} \right] \right\} , \tag{40}$$

where $\mathbb{1}_A$ denotes an indicator function which takes value unity when condition A is satisfied, else it vanishes.

The large number of available degrees of freedom in Eq (40) motivates a systematic approach to inferring the choice matrix components from a given set of data. We elect to consider models which isolate the many degrees of freedom by constructing scenarios where past behaviour dependent adherence only occurs for a single round and is temporally dependent on only one other round — all other degrees of freedom are hence set to those corresponding to time-dependent past behaviour independent adherence, i.e. $C^{\mathrm{T}}_{nn'} = C^{\mathrm{F}}_{nn'} = c_n$. The likelihood for this more restricted set of models — which we denote as $\mathcal{L}_{nn'}(D|\boldsymbol{\theta})$, where nn' corresponds to the pair of rounds chosen to be

dependent on each other in time — may be obtained by rewriting Eq (40) in the following form

$$\mathcal{L}_{nn'}(D|\boldsymbol{\theta}) = (1-C^{\mathrm{T}}_{nn'})^{Z^{n'n}_{\mathrm{TF}}}(C^{\mathrm{T}}_{nn'})^{Z^{n'n}_{\mathrm{TT}}}(1-C^{\mathrm{F}}_{nn'})^{Z^{n'n}_{\mathrm{FF}}}(C^{\mathrm{F}}_{nn'})^{Z^{n'n}_{\mathrm{FT}}}\prod_{\forall n''\neq n} c_{n''}^{N_{n''}}(1-c_{n''})^{N-N_{n''}}\,, \quad (41)$$

where we have defined

$$Z^{n'n}_{\mathrm{AB}} \equiv \sum_{\{\forall \boldsymbol{X}\,|\,X_{n'}=\mathrm{A},\,X_n=\mathrm{B}\}} N_{\boldsymbol{X}}\,, \quad (42)$$

where the data $D=\{N_{\boldsymbol{X}}\}$ has now been compressed into the set of numbers of people who track the same behaviour as $\boldsymbol{X}$, i.e., for 3 rounds, this forms the set of the following numbers of people: $N_{\mathrm{TTT}}, N_{\mathrm{TTF}}, N_{\mathrm{TFT}}$, etc. The Bayesian evidence integral corresponding to Eq (41) with a choice of flat prior $\pi(\boldsymbol{\theta})\propto 1$ is therefore

$$\begin{aligned}\mathcal{E}_{nn'} &= \int_0^1 (1-C^{\mathrm{T}}_{nn'})^{Z^{n'n}_{\mathrm{TF}}}(C^{\mathrm{T}}_{nn'})^{Z^{n'n}_{\mathrm{TT}}}\int_0^1 (1-C^{\mathrm{F}}_{nn'})^{Z^{n'n}_{\mathrm{FF}}}(C^{\mathrm{F}}_{nn'})^{Z^{n'n}_{\mathrm{FT}}}\mathrm{d}C^{\mathrm{T}}_{nn'}\mathrm{d}C^{\mathrm{F}}_{nn'} \\ &\quad\times\prod_{\forall n''\neq n}\left[\int_0^1 c_{n''}^{N_{n''}}(1-c_{n''})^{N-N_{n''}}\mathrm{d}c_{n''}\right] \\ &= \frac{\Gamma(Z^{n'n}_{\mathrm{TF}}+1)\Gamma(Z^{n'n}_{\mathrm{TT}}+1)}{\Gamma(Z^{n'n}_{\mathrm{TT}}+Z^{n'n}_{\mathrm{TF}}+2)}\frac{\Gamma(Z^{n'n}_{\mathrm{FF}}+1)\Gamma(Z^{n'n}_{\mathrm{FT}}+1)}{\Gamma(Z^{n'n}_{\mathrm{FF}}+Z^{n'n}_{\mathrm{FT}}+2)} \\ &\quad\times\prod_{\forall n''\neq n}\frac{\Gamma(N_{n''}+1)\Gamma(N-N_{n''}+1)}{\Gamma(N+2)}\,. \end{aligned} \quad (43)$$

Some non-Markovian past dependence may be captured by the likelihood defined in Eq (41), however their Bayesian evidence may need to be compared with equivalent Markovian models which also generate decaying long-term correlations of a particular form. Using the same formalism as Eq (41), the time-dependent Markov model has the following likelihood

$$\mathcal{L}(D|\boldsymbol{\theta}) = \alpha^{N_{\mathrm{T}}}(1-\alpha)^{N_{\mathrm{F}}}\prod_{\forall n\geq 2}(1-C^{\mathrm{T}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{TF}}}(C^{\mathrm{T}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{TT}}}(1-C^{\mathrm{F}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{FF}}}(C^{\mathrm{F}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{FT}}}\,, \quad (44)$$

and, hence, yields the following Bayesian evidence

$$\begin{aligned}\mathcal{E} &= \int_0^1 \alpha^{N_{\mathrm{T}}}(1-\alpha)^{N_{\mathrm{F}}}\mathrm{d}\alpha\prod_{\forall n\geq 2}\int_0^1 (1-C^{\mathrm{T}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{TF}}}(C^{\mathrm{T}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{TT}}} \\ &\quad\times\int_0^1 (1-C^{\mathrm{F}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{FF}}}(C^{\mathrm{F}}_{n\,n-1})^{Z^{n-1\,n}_{\mathrm{FT}}}\mathrm{d}C^{\mathrm{T}}_{n\,n-1}\mathrm{d}C^{\mathrm{F}}_{n\,n-1} \\ &= \frac{\Gamma(N_{\mathrm{T}}+1)\Gamma(N_{\mathrm{F}}+1)}{\Gamma(N+2)}\prod_{\forall n\geq 2}\frac{\Gamma(Z^{n-1\,n}_{\mathrm{TF}}+1)\Gamma(Z^{n-1\,n}_{\mathrm{TT}}+1)}{\Gamma(Z^{n-1\,n}_{\mathrm{TT}}+Z^{n-1\,n}_{\mathrm{TF}}+2)}\frac{\Gamma(Z^{n-1\,n}_{\mathrm{FF}}+1)\Gamma(Z^{n-1\,n}_{\mathrm{FT}}+1)}{\Gamma(Z^{n-1\,n}_{\mathrm{FF}}+Z^{n-1\,n}_{\mathrm{FT}}+2)}\,. \end{aligned} \quad (45)$$

Eqs (44) and (45) may also be used to obtain the likelihood of the time-independent Markov model

$$\mathcal{L}(D|\boldsymbol{\theta}) = \alpha^{N_{\mathrm{T}}}(1-\alpha)^{N_{\mathrm{F}}}\beta^{\sum_{\forall n\geq 2}Z^{n-1\,n}_{\mathrm{TT}}}(1-\beta)^{\sum_{\forall n\geq 2}Z^{n-1\,n}_{\mathrm{TF}}}\gamma^{\sum_{\forall n\geq 2}Z^{n-1\,n}_{\mathrm{FF}}}(1-\gamma)^{\sum_{\forall n\geq 2}Z^{n-1\,n}_{\mathrm{FT}}}\,, \quad (46)$$

and the Bayesian evidence of the same model

$$\begin{aligned}\mathcal{E} &= \int_0^1 \alpha^{N_{\mathrm{T}}}(1-\alpha)^{N_{\mathrm{F}}}\mathrm{d}\alpha \int_0^1 \beta^{\sum_{\forall n\geq 2} Z_{\mathrm{TT}}^{n-1\,n}}(1-\beta)^{\sum_{\forall n\geq 2} Z_{\mathrm{TF}}^{n-1\,n}}\mathrm{d}\beta \\ &\qquad \times \int_0^1 \gamma^{\sum_{\forall n\geq 2} Z_{\mathrm{FF}}^{n-1\,n}}(1-\gamma)^{\sum_{\forall n\geq 2} Z_{\mathrm{FT}}^{n-1\,n}}\mathrm{d}\gamma \\ &= \frac{\Gamma(N_{\mathrm{T}}+1)\Gamma(N_{\mathrm{F}}+1)}{\Gamma(N+2)} \frac{\Gamma\left(\sum_{\forall n\geq 2} Z_{\mathrm{TF}}^{n-1\,n}+1\right)\Gamma\left(\sum_{\forall n\geq 2} Z_{\mathrm{TT}}^{n-1\,n}+1\right)}{\Gamma\left[\sum_{\forall n\geq 2}\left(Z_{\mathrm{TT}}^{n-1\,n}+Z_{\mathrm{TF}}^{n-1\,n}\right)+2\right]} \\ &\qquad \times \frac{\Gamma\left(\sum_{\forall n\geq 2} Z_{\mathrm{FF}}^{n-1\,n}+1\right)\Gamma\left(\sum_{\forall n\geq 2} Z_{\mathrm{FT}}^{n-1\,n}+1\right)}{\Gamma\left[\sum_{\forall n\geq 2}\left(Z_{\mathrm{FF}}^{n-1\,n}+Z_{\mathrm{FT}}^{n-1\,n}\right)+2\right]}\,. \end{aligned} \tag{47}$$

S2 Appendix.

Summary. In this supplementary information we apply the framework of our mathematical model for adherence to the TUMIKIA project [11–13] and write a brief analysis description for each age group and sex.

Introduction

In Figs 3, 4, 5, 6 and 7 we plot the maximum likelihood as well as the limits of the marginalised 95% credible region for the conditional probabilities given treatment (filled points) or non-treatment (hollow points) in a previous round of the overall, male and female participants in the top, middle and bottom rows, respectively. In the left column the constant conditional probabilities between any given sequential pair of rounds have been inferred, which corresponds to the time-independent Markov model of the main text and Appendix S1. In the right column all possible round pair dependencies are considered (indicated by the arrows on the horizontal axis), where in each case the components corresponding to a given round were measured assuming all other respective rounds were inferred to be from past behaviour independent adherence. In all plots, above each pair of components we have also provided the log-Bayes factors [19], defined by

$$\ln \mathrm{B}_{nn'} = \ln\left(\frac{\mathcal{E}_{nn'}}{\mathcal{E}_{\mathrm{ref}}}\right) , \tag{48}$$

where the evidence for each pair $\mathcal{E}_{nn'}$ has been evaluated using the relations provided in Appendix S1 and the reference model evidence $\mathcal{E}_{\mathrm{ref}}$ has been set to that of time-dependent past behaviour independent adherence for all components.

Results

In Figs 3, 4 and 5 we present our results for the pre-SAC, SAC and 15-29 age groups of individuals in the TUMIKIA project. These age groups appear to be well-described by a time-dependent Markov model so past behaviour dependent non-adherence is clearly present. This may be identified by the largest log-Bayes factor values being given in the red-coloured right column plots for all three sets of plots. However, the conditional probabilities in all groups appear to drift closer together by round 4 of treatment, which signals a gradual transition from past behaviour dependent to independent adherence.

In Figs 6 and 7 we present our results for the 30-49 and 50+ age groups of individuals in the TUMIKIA project. The overall cohort, as well as the males and females in both age groups, appear to exhibit strong evidence of past behaviour dependent non-adherence — in particular, they are all apparently well-described by a time-independent Markov model. These conclusions may be drawn both by the consistent distance between all of the values for the inferred conditional probabilities with the red points of the right column of plots, as well as the largest evidence (as measured by the log-Bayes factor in the top row of the plots) for a difference in conditional probabilities in the left column in both plots.

In all of the cohorts studied in Figs 3, 4, 5, 6 and 7, we report no evidence for the existence of dependencies between rounds that depart from a Markovian description (as can be inferred from the comparatively small log-Bayes factors for the blue and green conditional probabilities in the right column of all plots). This is an interesting, and perhaps surprising, result regarding the nature of human behaviour in response to mass drug administration.

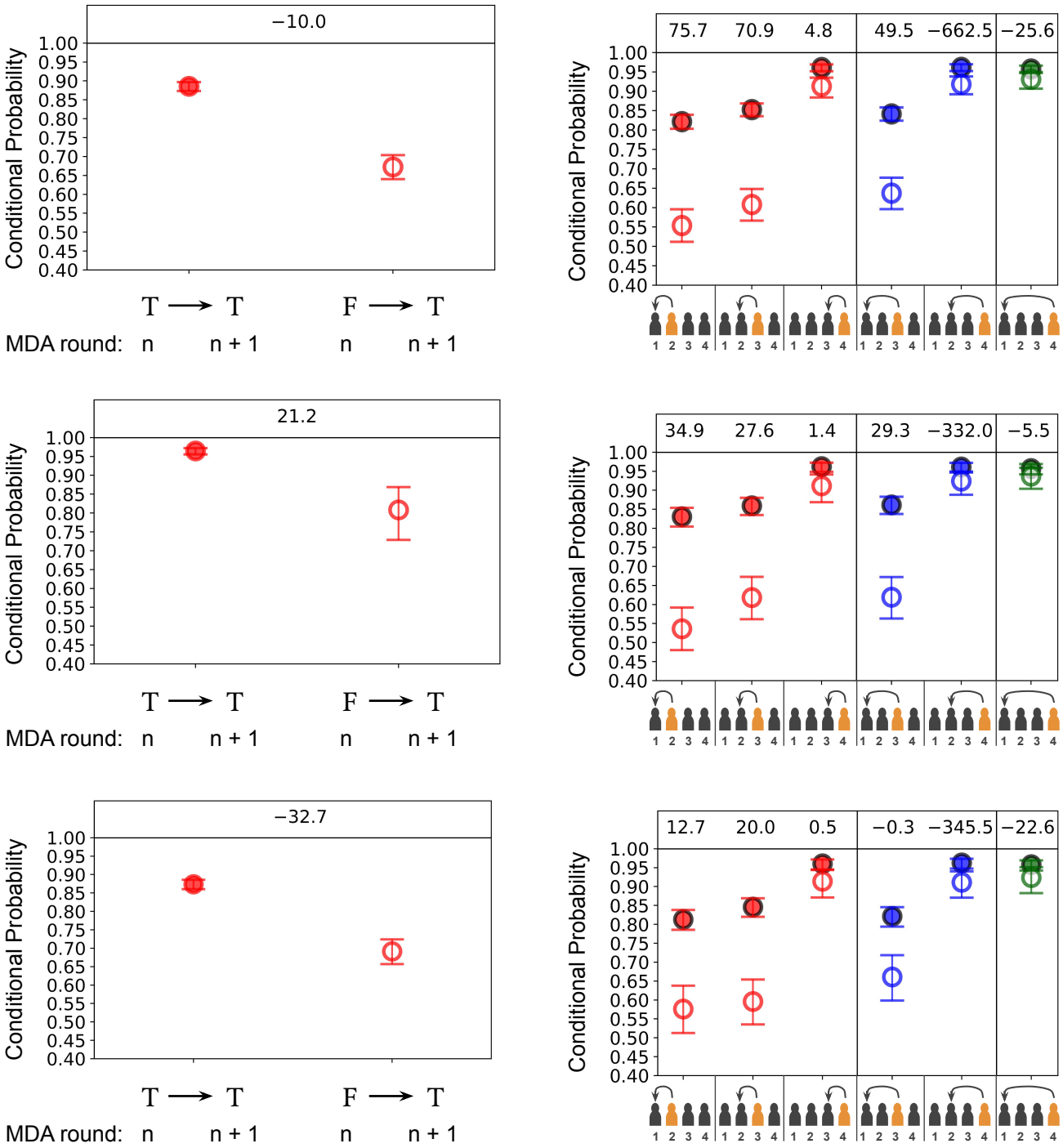

Fig 3. *Left column:* The maximum likelihood as well as the limits of the marginalised 95% credible region for the conditional probabilities of accepting treatment for any given pair of sequential rounds (these are hence homogeneous in time and the process is Markovian) given treatment (filled points) or non-treatment (hollow points) in a previous round. *Right column:* The same as the left column but with allowed time dependence in the conditional probabilities of accepting treatment in each respective round (highlighted in orange on the horizontal axes). In each case the components corresponding to a given round were measured assuming all other respective rounds were inferred to be from time-dependent past behaviour independent adherence and hence the likelihood is given in Appendix S1. Different colours for each point correspond to different lengths in time for the dependencies in behaviour. The datasets used are from the standard pre-SAC (0-4) age category from the TUMIKIA project where the: top row corresponds to the overall group; middle row corresponds to the male sub-group; and bottom row corresponds to the female sub-group.

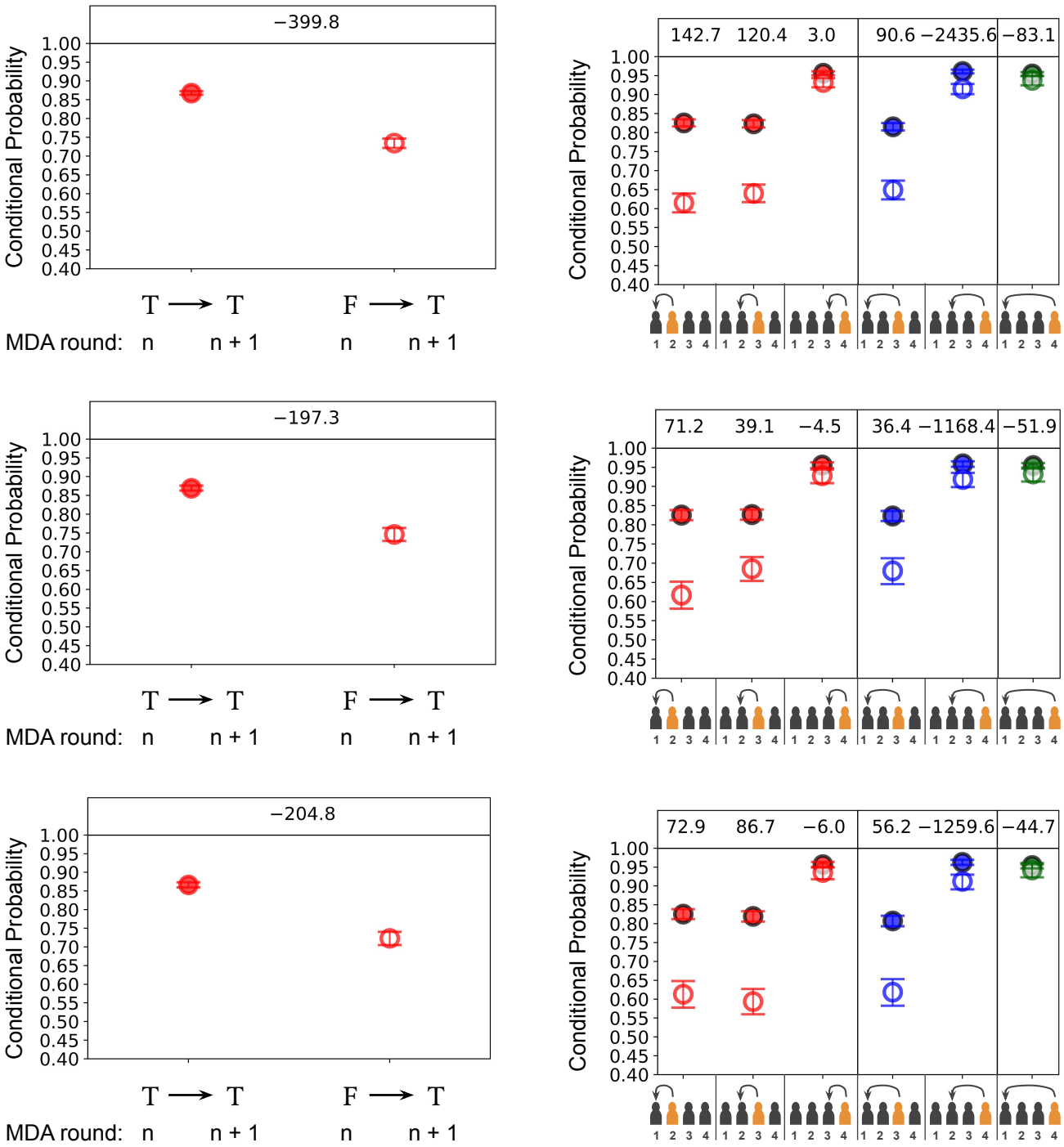

Fig 4. *Left column:* The maximum likelihood as well as the limits of the marginalised 95% credible region for the conditional probabilities of accepting treatment for any given pair of sequential rounds (these are hence homogeneous in time and the process is Markovian) given treatment (filled points) or non-treatment (hollow points) in a previous round. *Right column:* The same as the left column but with allowed time-dependent in the conditional probabilities of accepting treatment in each respective round (highlighted in orange on the horizontal axes). In each case the components corresponding to a given round were measured assuming all other respective rounds were inferred to be from time-dependent past behaviour independent adherence and hence the likelihood is given in Appendix S1. Different colours for each point correspond to different lengths in time for the dependencies in behaviour. The datasets used are from the standard SAC (4-15) age category from the TUMIKIA project where the: top row corresponds to the overall group; middle row corresponds to the male sub-group; and bottom row corresponds to the female sub-group.

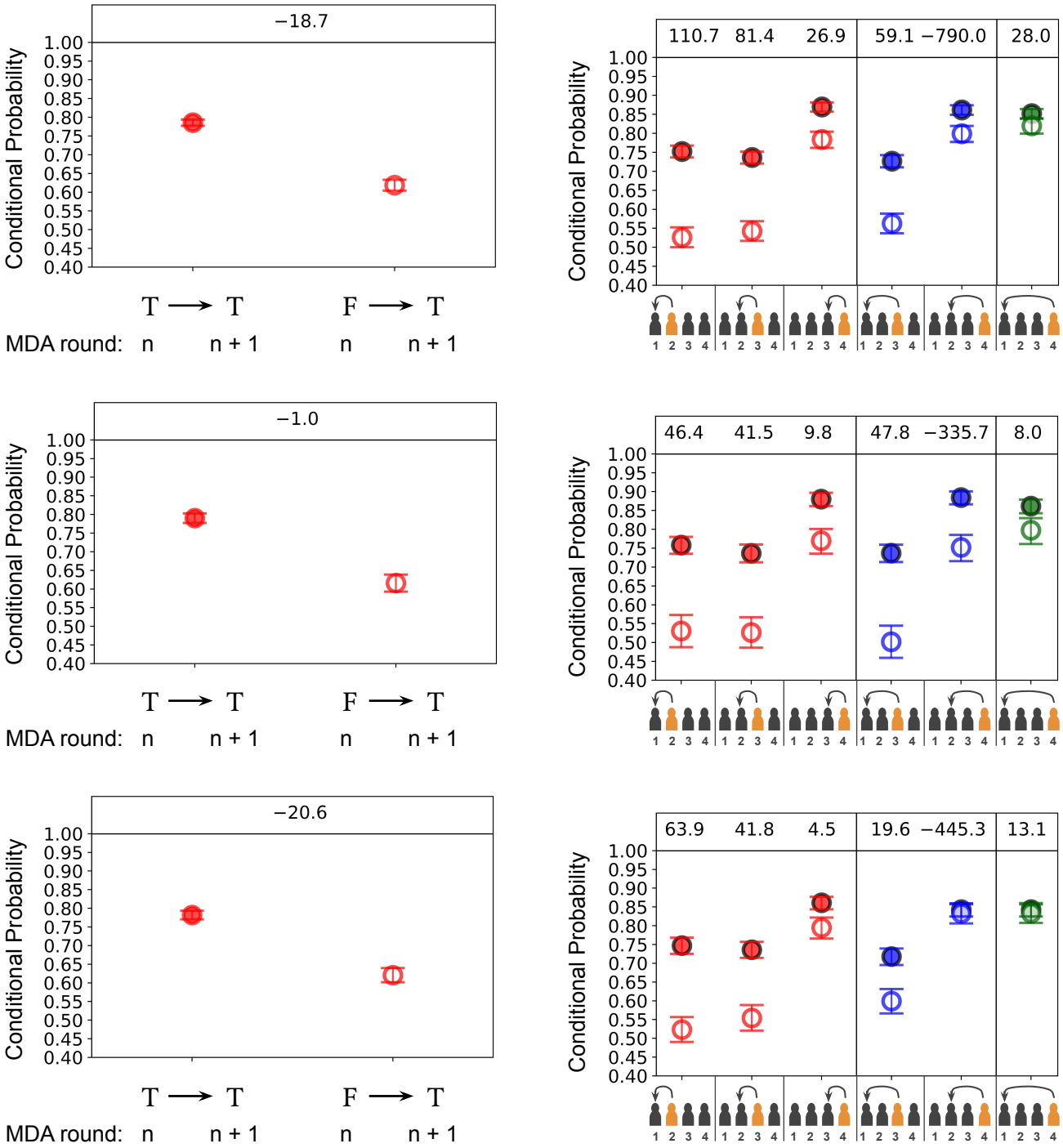

Fig 5. *Left column:* The maximum likelihood as well as the limits of the marginalised 95% credible region for the conditional probabilities of accepting treatment for any given pair of sequential rounds (these are hence homogeneous in time and the process is Markovian) given treatment (filled points) or non-treatment (hollow points) in a previous round. *Right column:* The same as the left column but with allowed time-dependent in the conditional probabilities of accepting treatment in each respective round (highlighted in orange on the horizontal axes). In each case the components corresponding to a given round were measured assuming all other respective rounds were inferred to be from time-dependent past behaviour independent adherence and hence the likelihood is given in Appendix S1. Different colours for each point correspond to different lengths in time for the dependencies in behaviour. The datasets used are from the 15-29 age category from the TUMIKIA project where the: top row corresponds to the overall group; middle row corresponds to the male sub-group; and bottom row corresponds to the female sub-group.

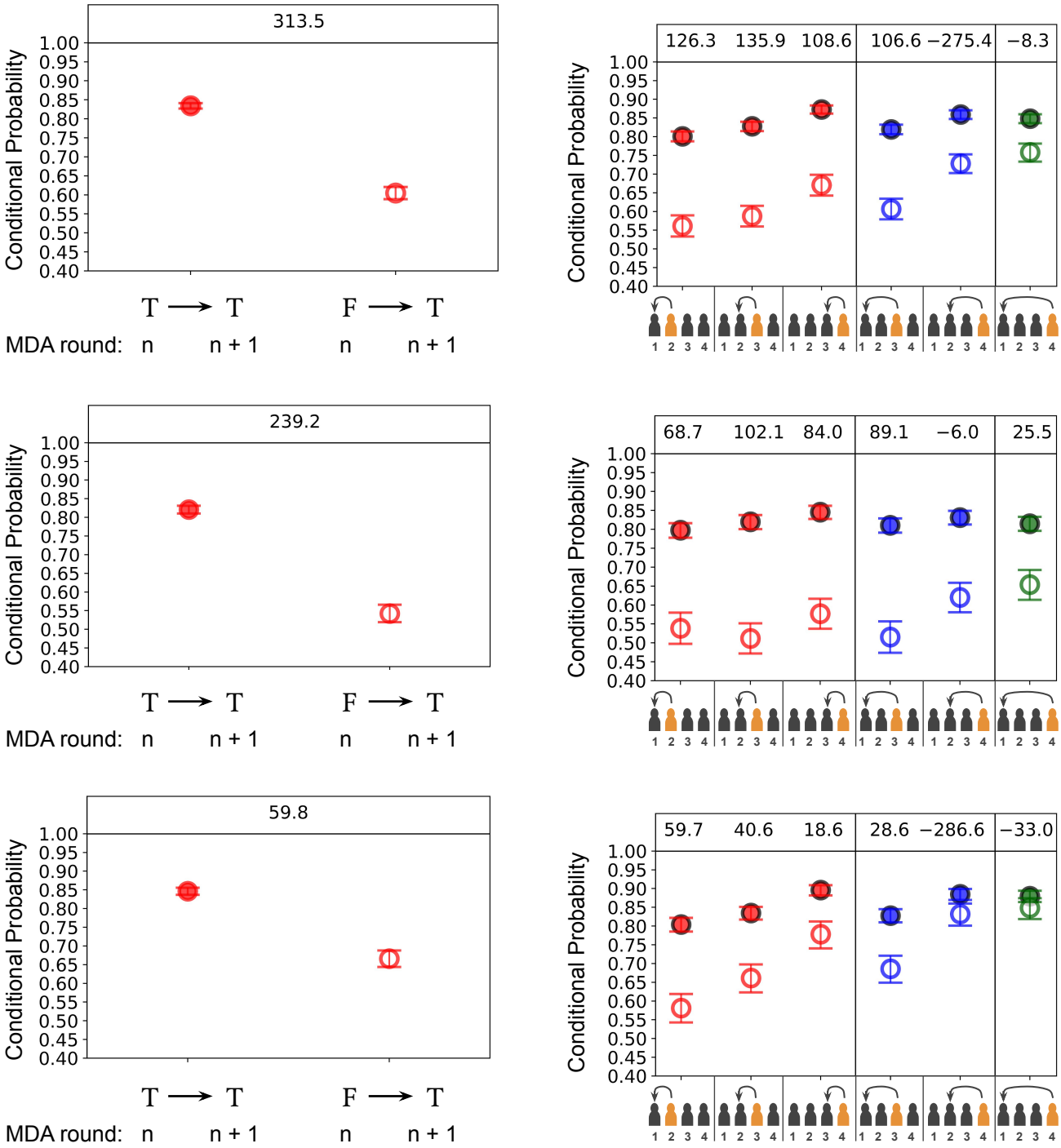

Fig 6. *Left column:* The maximum likelihood as well as the limits of the marginalised 95% credible region for the conditional probabilities of accepting treatment for any given pair of sequential rounds (these are hence homogeneous in time and the process is Markovian) given treatment (filled points) or non-treatment (hollow points) in a previous round. *Right column:* The same as the left column but with allowed time dependence in the conditional probabilities of accepting treatment in each respective round (highlighted in orange on the horizontal axes). In each case the components corresponding to a given round were measured assuming all other respective rounds were inferred to be from time-dependent past behaviour independent adherence and hence the likelihood is given in Appendix S1. Different colours for each point correspond to different lengths in time for the dependencies in behaviour. The datasets used are from the 30-49 age category from the TUMIKIA project where the: top row corresponds to the overall group; middle row corresponds to the male sub-group; and bottom row corresponds to the female sub-group.

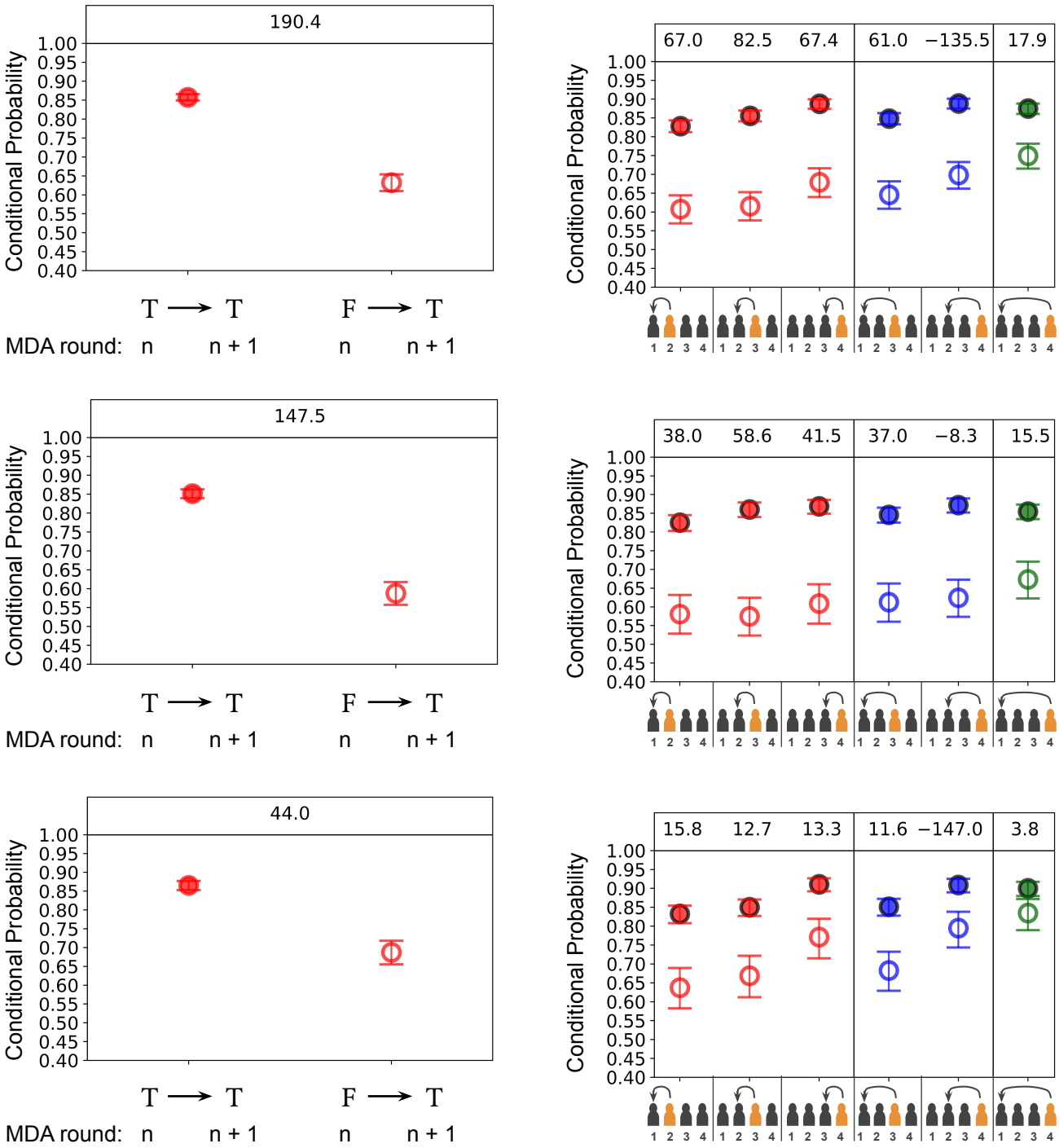

Fig 7. *Left column:* The maximum likelihood as well as the limits of the marginalised 95% credible region for the conditional probabilities of accepting treatment for any given pair of sequential rounds (these are hence homogeneous in time and the process is Markovian) given treatment (filled points) or non-treatment (hollow points) in a previous round. *Right column:* The same as the left column but with allowed time dependence in the conditional probabilities of accepting treatment in each respective round (highlighted in orange on the horizontal axes). In each case the components corresponding to a given round were measured assuming all other respective rounds were inferred to be from time-dependent past behaviour independent adherence and hence the likelihood is given in Appendix S1. Different colours for each point correspond to different lengths in time for the dependencies in behaviour. The datasets used are from the 50+ age category from the TUMIKIA project where the: top row corresponds to the overall group; middle row corresponds to the male sub-group; and bottom row corresponds to the female sub-group.

S3 Appendix.

Summary. In this supplementary information, we analyse some of the existing models of adherence from the literature in the context of our proposed framework.

The Plaisier model

Several models of MDA treatment programs employ an adherence model developed by Plaisier in the context of onchocerciasis control [8, 20]. The Plaisier model assigns a probability of adherence to each individual which they then retain for the duration of the MDA program [14, 15]. As such, this model would be characterised by us as a heterogeneous population, time-independent model with no explicit individual dependence on past behaviour. The individual probability of adherence is given by $U^{(1-c)/c}$, where U is a uniform random number and c is expected probability of treatment and hence the expected coverage. The model is therefore completely parameterized by the overall expected coverage. The PDF for the adherence probability for this process is given by

$$\pi(p) = \frac{c}{1-c} p^{(2c-1)/(1-c)} \,. \tag{49}$$

The PDF of p rises monotonically from zero to one for all values of $c > 0.5$ and falls monotonically for $c < 0.5$ (for $c = 0.5$, it is flat). Note that $\pi(p)$ is a beta distribution: $\pi(p) = \mathrm{Beta}[p; c/(1-c), 1]$. For this distribution, the mean failure run length is hence given by

$$\mathrm{E}(n_{\mathrm{F}}) = \frac{c}{2c-1} \,. \tag{50}$$

Note that in this model, adherence failure run length becomes undefined at a coverage of 50% or less. Additionally, one can show that the variance of this random variable becomes undefined for values of coverage below 66%, suggesting that failure run lengths in finite populations drawn from this distribution will exhibit extreme variability.

The probability of an individual being untreated across N rounds of MDA in this model can also be calculated, giving

$$\pi_{\mathrm{un}} = \int_0^1 (1-p)^N \mathrm{Beta}[p; c/(1-c), 1] \, \mathrm{d}p = \frac{c}{1-c} B[c/(1-c), N+1] \,, \tag{51}$$

where $B(\cdot, \cdot)$ is the beta function. Fig 8 shows the distribution of adherence probabilities for 2 different coverage values and also the probability of an individual not adhering with treatment across a 4-round MDA program.

The Griffin Model

The adherence model used by Irvine et al [5] to model MDA adherence in the treatment of lymphatic filariasis was originally created by Griffin et al in the context of intevention strategies against malaria transmission [21]. The original Griffin model is quite broad and deals with multiple simultaneous interventions and the correlations in their uptake. It does not include conditional dependencies for an individual's behaviour and is therefore a heterogeneous population, time-independent, individually past behaviour independent model in its simplest form. Each individual in the population is assigned a

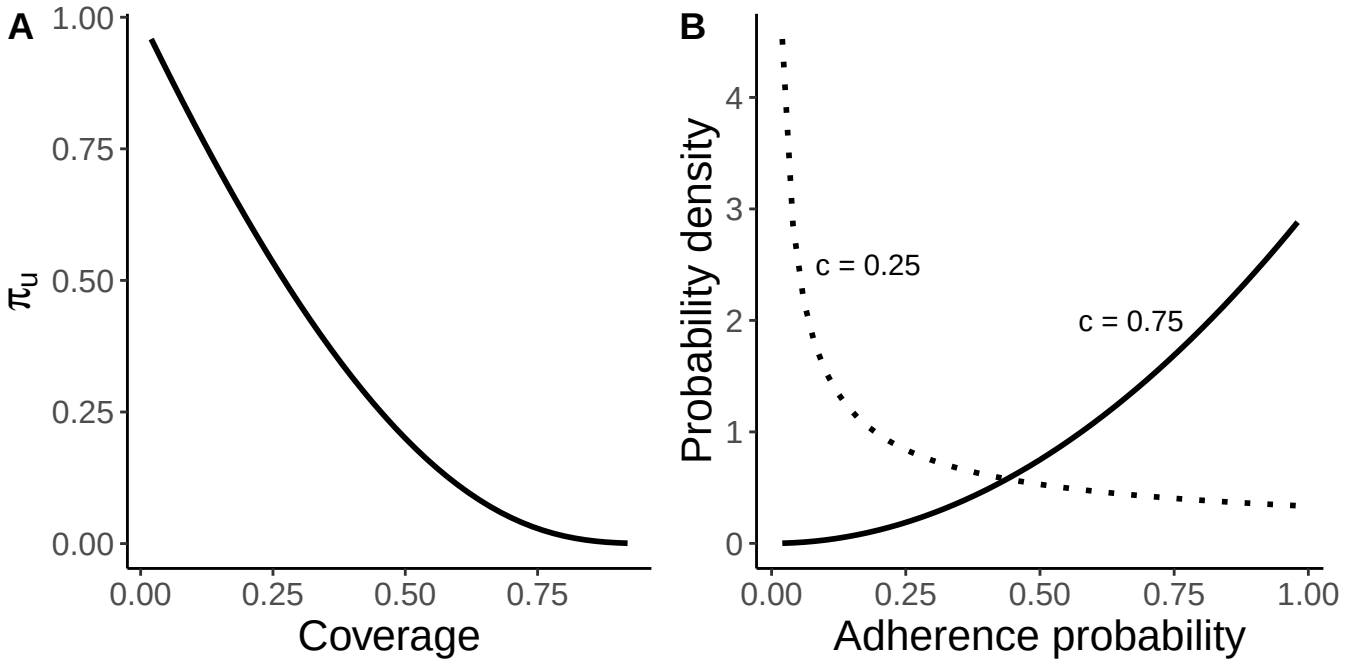

Fig 8. A) Probability of an individual with adherence drawn from the Plaisier distribution of not adhering with treatment during a 4 round MDA program. B) The probability distribution for adherence for coverages of 25% and 75%.

correlation parameter, u_i, drawn from a normal distribution with mean u_0 and variance σ^2. These parameters are retained throughout the MDA program. At each round a MDA round, each individual draws a unit-variance normal deviate with mean u_i, z. Treatment is accepted if $z < 0$. The expected coverage is given by $\phi(-u_0/\sqrt{1+\sigma^2})$, where ϕ is the standard normal cumulative probability function. This leaves one free parameter to control the distribution of adherence probabilities across the population.

The cumulative distribution of adherence probability, p, is given by

$$\pi(p) = \phi[\phi^{-1}(p;0,1) + u_0; 0, \sigma^2]\,, \tag{52}$$

giving a PDF

$$P(p) \propto \exp\left[-\frac{1-\sigma^2}{2\sigma^2}\left(\phi^{-1}(p) + \frac{u_0}{1-\sigma^2}\right)^2\right]. \tag{53}$$

The function $\phi^{-1}(p;0,1)$ varies monotonically in the range $(-\infty,\infty)$ with p. In Eq (53), the parameter $\sigma = 1$ acts to discriminate between two functional forms. For $\sigma < 1$, the distribution has a 'normal' shape with a single local maximum, while for $\sigma > 1$, the distribution has asymptotes with local maxima at the $p = 0$ and/or 1. In this, it is very similar, qualitatively, to the beta distribution (see Fig 9).

References

1. Boatin BA, Basáñez MG, Prichard RK, Awadzi K, Barakat RM, García HH, et al. A research agenda for helminth diseases of humans: towards control and elimination. PLoS neglected tropical diseases. 2012;6(4):e1547.

2. Krentel A, Fischer PU, Weil GJ. A review of factors that influence individual compliance with mass drug administration for elimination of lymphatic filariasis. PLoS neglected tropical diseases. 2013;7(11):e2447.

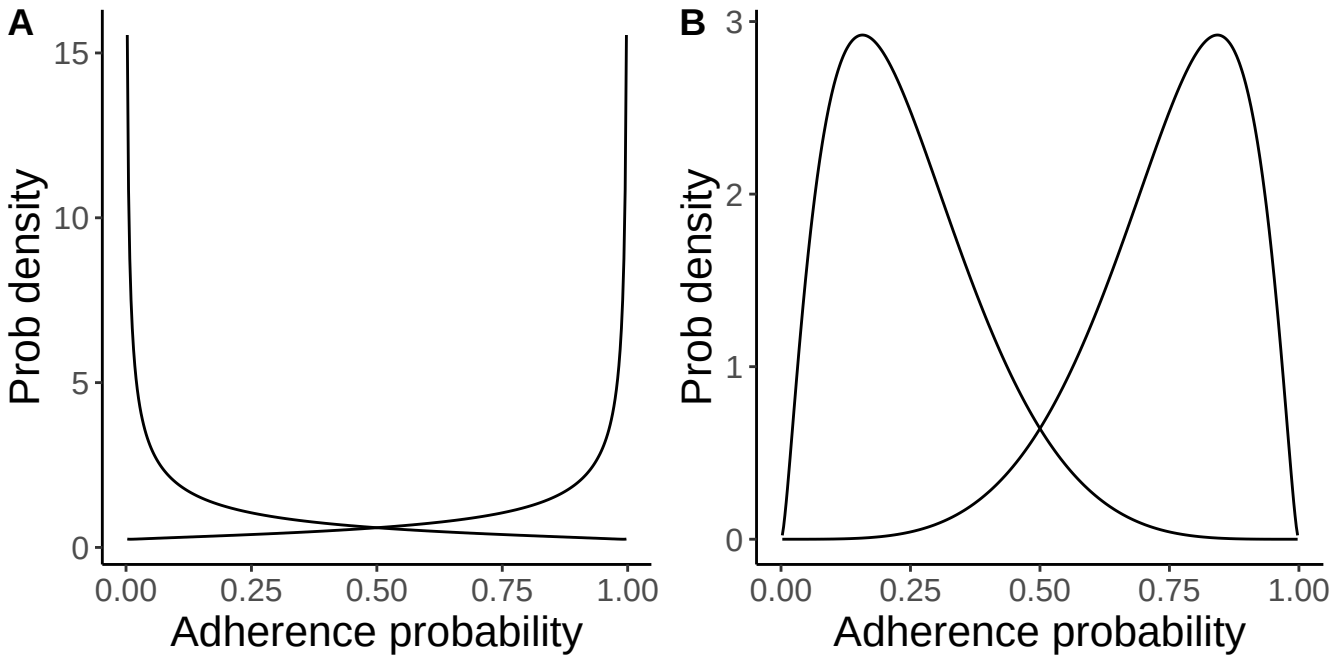

Fig 9. Adherence probability distributions with A) $\sigma = 1.2$ and B) $\sigma = 0.5$ for mean coverages of 25% and 75%. The probability distribution for adherence for coverages of 25% and 75%.

3. Truscott JE, Hollingsworth TD, Brooker SJ, Anderson RM. Can chemotherapy alone eliminate the transmission of soil transmitted helminths? Parasites & vectors. 2014;7(1):266.

4. Babu BV, Babu GR. Coverage of, and compliance with, mass drug administration under the programme to eliminate lymphatic filariasis in India: a systematic review. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2014;108(9):538–549.

5. Irvine MA, Reimer LJ, Njenga SM, Gunawardena S, Kelly-Hope L, Bockarie M, et al. Modelling strategies to break transmission of lymphatic filariasis - aggregation, adherence and vector competence greatly alter elimination. Parasites & Vectors. 2015 Oct;8(1):547. Available from: https://doi.org/10.1186/s13071-015-1152-3.

6. Coffeng LE, Bakker R, Montresor A, de Vlas SJ. Feasibility of controlling hookworm infection through preventive chemotherapy: a simulation study using the individual-based WORMSIM modelling framework. Parasites & vectors. 2015;8(1):541.

7. Shuford KV, Turner HC, Anderson RM. Compliance with anthelmintic treatment in the neglected tropical diseases control programmes: a systematic review. Parasites & vectors. 2016;9(1):29.

8. Farrell SH, Truscott JE, Anderson RM. The importance of patient compliance in repeated rounds of mass drug administration (MDA) for the elimination of intestinal helminth transmission. Parasites & Vectors. 2017 Jun;10(1):291. Available from: https://doi.org/10.1186/s13071-017-2206-5.

9. Organization WH, et al. Guideline: preventive chemotherapy to control soil-transmitted helminth infections in at-risk population groups. World Health Organization; 2017.

10. Fogarty L, Roter D, Larson S, Burke J, Gillespie J, Levy R. Patient adherence to HIV medication regimens: a review of published and abstract reports. Patient Education and Counseling. 2002;46(2):93 – 108. Available from: http://www.sciencedirect.com/science/article/pii/S0738399101002191.

11. Pullan RL, Halliday KE, Oswald WE, Mcharo C, Beaumont E, Kepha S, et al. Effects, equity, and cost of school-based and community-wide treatment strategies for soil-transmitted helminths in Kenya: a cluster-randomised controlled trial. The Lancet. 2019;393(10185):2039–2050.

12. Halliday KE, Oswald WE, Mcharo C, Beaumont E, Gichuki PM, Kepha S, et al. Community-level epidemiology of soil-transmitted helminths in the context of school-based deworming: Baseline results of a cluster randomised trial on the coast of Kenya. PLoS neglected tropical diseases. 2019;13(8):e0007427.

13. Oswald WE, Kepha S, Halliday KE, Mcharo C, Safari T, Witek-McManus S, et al. Systematic non-compliance in four rounds of community-wide treatment for soil-transmitted helminths in the TUMIKIA trial, Kwale County, Kenya. Submitted to Parasites & Vectors; 2019.

14. Plaisier AP. Modelling onchocerciasis transmission and control. Erasmus University Rotterdam; 1996.

15. Plaisier AP, Stolk WA, van Oortmarssen GJ, Habbema JDF. Effectiveness of annual ivermectin treatment for Wuchereria bancrofti infection. Parasitology Today. 2000;16(7):298–302.

16. Truscott JE, Ower AK, Werkman M, Halliday K, Oswald WE, Gichuki PM, et al. Heterogeneity in transmission parameters of hookworm infection within the baseline data from the TUMIKIA study in Kenya. Parasites & vectors. 2019;12(1):442.

17. Ásbjörnsdóttir KH, Ajjampur SSR, Anderson RM, Bailey R, Gardiner I, Halliday KE, et al. Assessing the feasibility of interrupting the transmission of soil-transmitted helminths through mass drug administration: The DeWorm3 cluster randomized trial protocol. PLOS Neglected Tropical Diseases. 2018 01;12(1):1–16. Available from: https://doi.org/10.1371/journal.pntd.0006166.

18. Mekete K, Ower A, Dunn J, Sime H, Tadesse G, Abate E, et al. The Geshiyaro Project: a study protocol for developing a scalable model of interventions for moving towards the interruption of the transmission of soil-transmitted helminths and schistosome infections in the Wolaita zone of Ethiopia. Parasites & vectors. 2019;12(1):1–12.

19. Jeffreys H. The theory of probability. OUP Oxford; 1998.

20. Dyson L, Stolk WA, Farrell SH, Hollingsworth TD. Measuring and modelling the effects of systematic non-adherence to mass drug administration. Epidemics. 2017;18:56–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28279457http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5340860.

21. Griffin JT, Hollingsworth TD, Okell LC, Churcher TS, White M, Hinsley W, et al. Reducing Plasmodium falciparum Malaria Transmission in Africa: A Model-Based Evaluation of Intervention Strategies. PLoS Medicine. 2010 aug;7(8):e1000324. Available from: `http://dx.plos.org/10.1371/journal.pmed.1000324`.