Review

CellPress

Models for measuring anthelmintic drug efficacy for parasitologists

Martin Walker, Thomas S. Churcher, and María-Gloria Basáñez

Department of Infectious Disease Epidemiology, London Centre for Neglected Tropical Disease Research, School of Public Health, Faculty of Medicine (St Mary's Campus), Imperial College London, Norfolk Place, London, W2 1PG, UK

Anthelmintic drug efficacy (ADE) is generally estimated as a population average effect, despite drug responses varying among individuals according to a variety of measurable and non-measurable factors. Model-based and/or individual-level analyses are scarce and often methodologically frail. We propose that wider application of marginal and mixed models would offer benefits to the evaluation of ADE. We demonstrate, with a worked example, how model-based analyses: (i) capture the effects of correlation among hierarchically structured longitudinal data on estimates of ADE; (ii) permit robust inference on the association of measurable factors with ADE; and (iii) enable estimation of variation among individual-level estimates of ADE. The application of modelling approaches is discussed in the context of mass drug administration-based control of human helminthiases.

The imperative to measure anthelmintic drug efficacy

The effectiveness of treating and controlling human and livestock helminthiases critically depends on the efficacy of anthelmintic drugs. In livestock, the utility of anthelmintics has been severely diminished by the rapid evolution and spread of anthelmintic resistance [1–3]. Despite the lessons learnt from 50 years of somewhat indiscriminate livestock treatment strategies [4], there is due concern that resistance, or at least sub-optimal drug efficacy, could derail the burgeoning global onslaught against helminth infections of humans. The current strategic intervention is based principally on anthelmintic mass drug administration (MDA) [5–8], and is endorsed by the World Health Organization (WHO) in their roadmap on the control and elimination of neglected tropical diseases (NTDs) by 2015 and 2020 [9]. Application of appropriate and powerful statistical methods that enable accurate estimation of anthelmintic drug efficacy (ADE) is a high priority, both for monitoring and evaluation (M&E) of control programmes [10] and for analysing outcomes from clinical trials of the next generation of new [11–13] or repurposed anthelmintics [14–16].

In this review, we show that established extensions of generalized linear models (GLMs; see Glossary) [17] offer a versatile and practical way of estimating ADE both at the population level, in terms of an average effect, but also at the level of the individual host. We review the predominant methods currently used for estimating ADE, many of which were developed in a veterinary context and are based on sample statistics. These are contrasted with modelling approaches using previously published data on hookworm egg counts collected from Kenyan schoolchildren before and 1 week after treatment with albendazole [18]. These data are summarized in Table 1 and are referred to as the Kenyan schoolchildren dataset, abbreviated to KSD. We discuss the application of modelling approaches, particularly in the context of the M&E of ADE during MDA-based control of human helminthiases.

Cure rates and intensity reduction rates

Anthelmintic drug efficacy is typically, but not exclusively, expressed as either a cure rate (CR) or an intensity reduction rate (IRR). CRs (the proportion of those positive for parasites that become parasitologically negative after treatment) are calculated from binary data on the presence or absence of infection; IRRs (the proportional reduction of infection load effected by the treatment) are calculated from (typically count) data on the intensity of infection. Both quantify reduction in infection levels after treatment (the drug response) as a percentage of infection levels before treatment using longitudinal data from cohort studies.

CRs and IRRs can be calculated, in principle, using parasitological, molecular, or any other type of data that measure, respectively, infection status or infection intensity. Currently there are few quantitative molecular methods which yield estimates of infection intensity [19], albeit with some notable exceptions. One exception is measurement of circulating filarial antigen (CFA) for *Wuchereria bancrofti* (causing Bancroftian lymphatic filariasis) infection, although because of difficulties in counting adult worms, quantities of CFA have not been calibrated to worm burden, although this has been achieved in animal models [20]. Other examples include measurement of circulating anodic antigen (CAA), which have been correlated with *Schistosoma mansoni* egg output [21], and quantitative polymerase chain reaction (qPCR) for *Ascaris lumbricoides* [22,23] and hookworm [24] infections. Consequently, molecular diagnostics are mostly used for measuring infection status. Indeed, even for infections where molecular diagnostics do exist, or are undergoing field testing, ADE remains overwhelmingly assessed using data on parasite transmission stages (eggs or larvae). Therefore, we focus on

Corresponding author: Walker, M. (m.walker06@imperial.ac.uk).
Keywords: marginal models; mixed models; mass drug administration; longitudinal data; hookworm; albendazole.

1471-4922/
 http://dx.doi.org/10.1016/j.pt.2014.08.004

Glossary

Arithmetic mean: the sum of a collection of numbers divided by the number in the collection, often simply called the mean or average.
Cure rate (CR): the proportion of individual hosts positive for parasites who become parasitologically negative after treatment.
Exchangeable correlation: correlation among observations measured from a single unit (e.g., multiple parasite counts measured from a single host) that is assumed constant among units (i.e., among hosts).
Fixed and random effects: definitions of fixed and random effects vary with the specific context [85]. Here, a covariate coefficients (parameters) specified as fixed exert a constant effect among individuals while coefficients specified as random effects exert a variable effect among individuals. Parameters exerting random effects include a fixed component which represents the hypothetical effect on the 'average individual' but not necessarily the average effect among individuals.
Generalized estimating equation (GEE): a technique for estimating the parameters of a marginal model fitted to correlated repeated measures (observations). The GEE approach is semi-parametric because it relies on the first two moments of the observed data, but not on the full likelihood.
Generalized linear model (GLM): an extension of the simple linear regression model that is compatible with error distributions from any of the exponential family of probability distributions, including the normal, Poisson, binomial, and gamma distributions. The simple linear regression model is a GLM with normally distributed errors.
Generalized linear mixed model (GLMM): an extended GLM that includes a linear predictor comprised of covariate coefficients that exert both fixed and random effects.
Geometric mean: a type of mean or average which quantifies the central tendency of a set of numerical observations using the product, rather than the sum, of their values. Typically, geometric means are calculated by first taking the arithmetic mean of the log-transformed values before taking the exponent of the result to transform back onto the original scale.
Hierarchical structure: observations that are nested within units to define a natural hierarchy. Examples are multiple parasite counts measured within a host; multiple hosts living within a single household; multiple households within a single community. Such structure typically produces correlations among repeated measures (observations) made on the same unit. Thus, repeated measures cannot be assumed statistically independent.
Intensity reduction rate (IRR): the infection load after treatment expressed as a proportion of the infection load before treatment.
Linear predictor: the linear combination of covariates and coefficients within a statistical model.
Link function: a function that relates the expected value of a probability distribution to the linear predictor within a statistical model. The natural logarithmic link function is typically used within statistical models for count data. For binomial models, where p is the probability of 'success', the logit link is often used, logit $(p) = \ln[p/(1 - P)]$.
Longitudinal data: measurements or observations made repeatedly on the same unit (repeated measures) through time; for example, multiple hookworm egg counts made from the same host at different times.
Marginal model: an adaptation of a GLM for use with correlated repeated measures (observations). Marginal refers to the marginal mean of observations from individuals (units) sharing a set of covariates. A marginal model comprises three model components; a marginal mean which depends on covariates; a marginal variance which is typically a function of the marginal mean, and a correlation structure for the repeated measures.
Markov chain Monte Carlo (MCMC): a stochastic algorithm central to Bayesian statistical inference which samples parameter values from the posterior probability distribution by combining information from the likelihood of the observed data and the prior probability distribution of the parameters.
Maximum likelihood (ML) estimation: a framework for estimating parameters of a statistical model by conditioning the probability of the observed data (the likelihood) on unknown parameter values using a probability distribution.
Odds ratio (OR): the ratio of the odds that an outcome occurs given a set of covariates compared to the odds that the outcome occurs in their absence. For example, the odds of observing (by Kato–Katz) hookworm eggs 1 week after treatment with albendazole divided by the odds of observing hookworm eggs immediately before treatment.
Overdispersion: the occurrence of variance that is greater than expected based on a simple probability distribution. Extra-Poisson variation is an example of overdispersed count data; where the variance is greater than expected if the data were Poisson distributed (i.e., variance greater than the mean, $v > \mu$).
Posterior probability distribution (posterior): the probability distribution of a random variable conditional on relevant observed data and prior information. The posterior probability is proportional to the likelihood of the data (conditional on a set of parameter values) multiplied by prior probability of the parameters. That is, posterior probability $\propto$ likelihood $\times$ prior probability.
Prior probability distribution (prior): the probability distribution of a random variable that captures one's uncertainty before (prior to) observing relevant data. An uninformative or vague prior expresses a high degree of prior uncertainty. This results in a posterior distribution which is dominated by the likelihood of observed data. Conversely, an informative prior will dominate the posterior if the data holds little information on the variable of interest.
Rate ratio (RR): the ratio of the rate of occurrence of an event given a set of covariates compared to the rate of occurrence in their absence. For example, the average number of hookworm eggs counted (by Kato–Katz) 1 week after treatment with albendazole divided by the average number of hookworm eggs counted before treatment.
Risk ratio (RR): the ratio of the probability of an event occurring given a set of covariates compared to the probability of the event occurring in their absence. For example, the probability of observing hookworm eggs (by Kato–Katz) 1 week after treatment with albendazole divided by the probability of observing hookworm eggs before treatment.
Repeated measures: measurements or observations made repeatedly on the same unit, for example, multiple hookworm counts measured from the same individual host.
Restricted maximum likelihood (REML) estimation: an alternative to ML estimation for models that include random effects. In REML estimation, the dispersion of the random effects is estimated having averaged over some of the uncertainty in the fixed effects. By contrast, in ML estimation, the fixed effects estimates are treated as precisely correct.
Sample statistic: a quantity calculated from a sample of data using simple mathematical functions which are independent of the sample's distribution.
Sampling distribution: the hypothetical expected distribution of a quantity estimated from a random sample of observations.
Sandwich estimator: a standard error (SE) of an estimated quantity that is robust to misspecifications in the variance-covariance of the error distribution in a statistical model. Sandwich estimators are typically used with marginal models so that SEs (and confidence intervals) are invariant to inaccuracies in the specification of the repeated measures correlation structure. In this context, sandwich estimators are based on the empirically observed variation among unit-level statistics rather than on the model-derived variance-covariance matrix which depends on the assumed correlation structure [53].
Standard error (SE): the standard deviation of the sampling distribution of an estimated statistic (e.g., an arithmetic or geometric mean).

modelling approaches for such parasitological data, although in principle the methods are readily adaptable to other types of data.

Intensity reduction rates are more informative and generally more desirable than CRs, and have been used extensively for assessing anthelmintic efficacy in livestock. Perhaps the most well-known IRR is the faecal egg count reduction (FECR) [25], which is calculated from data on egg counts in faeces. More recently, the WHO has endorsed IRRs for the M&E of human schistosomiasis and soil-transmitted helminthiasis (STH) MDA-based control programmes [26]. CRs are often criticised because some anthelmintics are never truly curative (e.g., ivermectin only affects the microfilarial progeny of adult female *Onchocerca volvulus* and exerts only temporary deleterious effects on worm fertility [27]). In addition, CRs are less relevant to morbidity reduction since morbidity is, by and large, associated with infection intensity [28], and do not adequately reflect the impact of repeated rounds of treatment in the context of MDA interventions in human populations [29]. In a research and development context (R&D; e.g., clinical trials and epidemiological studies), intensity should always be measured, permitting calculation of IRRs. However, for M&E, logistical complexities and the availability of field-ready quantitative diagnostic tools [19] means that data on the presence of absence of infection is common, guaranteeing the continued usefulness of CRs, or other metrics based on binary data.

Contrasting methods of estimation

Anthelmintic drug efficacy can be estimated by two contrasting approaches. Sample estimates (statistics), or sample