Quantal Response Analysis Functions

John Maindonald, Statistics Research Associates

07 June 2021

Contents

Introduction

Quantal responses are counts of all-or-none responses, out of a total n. A dose-response relationship quantifies the response as a function of dose, or more generally as a function of an exposure.

Data are from a broad class of experiments where responses are insect deaths out of some total number exposed. Exposure may be time in coolstorage, or dose of a fumigant, or a concentration-time measure of exposure to a fumigant, or an intensity-time measure of exposure to radiation. See Follett and Neven (2006) for commentary on the regulatory and scientific background. We will use Hawaian fruitfly data that has been supplied by Dr Peter Follett to demonstrate the use of functions in the *qra* package, where the exposure is time in coolstorage,

The following code sets up the data.

```
suppressPackageStartupMessages(library(qra))
HawCon <- qra::HawCon
## Change name "CommonName" to "CN", for more compact output.
CCnum <- match("CommonName", names(HawCon))
names(HawCon)[CCnum] <- "CN"
## trtGp will identify species & lifestage combination
## trtGpRep will identify species, lifestage, and rep
## cTime is centered version of TrtTime
## scTime is centered and scaled version of TrtTime,
## needed to get some mixed model fits to converge
HawCon <- within(HawCon, {
  trtGp <- factor(paste0(CN,LifestageTrt, sep=":"))
  trtGpRep <- paste0(CN,LifestageTrt,":",RepNumber)
  scTime <- scale(TrtTime)
  obs <- factor(1:nrow(HawCon))
})
```

1 Data setup and choice of model

For the data that will be considered here, the exposure measure is time in coolstorage, and the response is mortality of insect pests.

1.1 Graphical Display

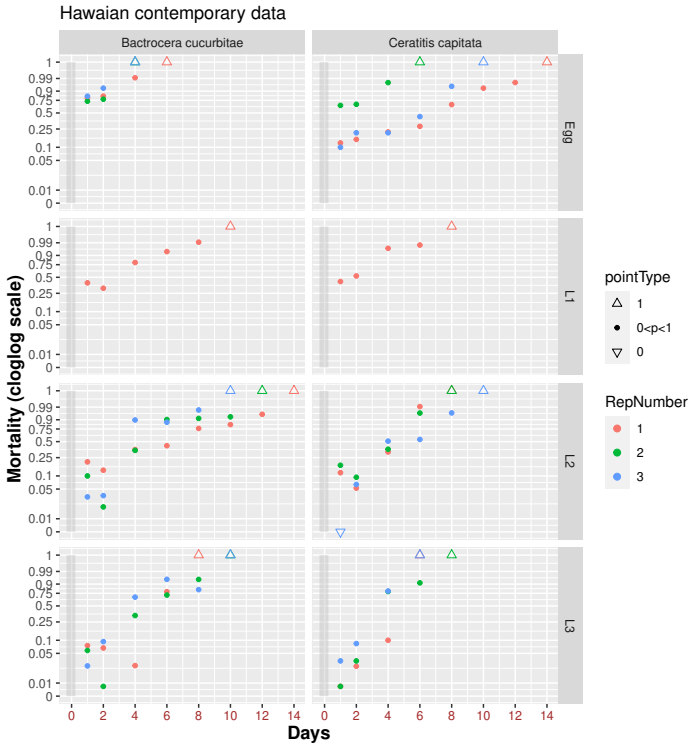

Figure 1: Graphs are designed to give an indication of the pattern, when mortalities are shown on a complementary log-log scale, of mortality response with days in coolstorage.

Figure 1 is designed to give a broad indication of the pattern of response, on a complementary log-log link scale. Responses appear acceptably linear,

at least after the first one or two observations. There are clear systematic differences between replicates, indicative of a strong between replicate component of variation.

1.2 The choice of link function

Commonly used link functions are

- `logit`: this is difficult to distinguish, for any except datasets that have very large numbers of observations at exposures that lead to very low or very high mortality, from the `probit`. With the logit link, the model is predicting the log(odds) of response. The `probit` link can be motivated by assuming the presence of a normally distributed random variable that generates a response whenever its value crosses a threshold.[1]
- `complementary log-log`, abbreviated to `cloglog`. This arises naturally as an extreme value distribution, when (for insect mortality), death arises from the failure of any one of a number of organs.

While later analyses will suggest a mild preference for a complementary log-log link function, as against a logit link, the difference is small. The difference does matter, for extrapolation to mortality values (commonly 99% or greater) that lie close to or beyond the limits of the data. The fitted models should, for this purpose, be treated as providing an indication of the broad pattern of response.

The possible use of a transformation for the exposure variable (or variables), commonly a logarithmic transformation, gives further flexibility. One or other choice within the range of possibilities noted has often been found to work well in practice, at least to the extent that the model survives scrutiny under standard diagnostic checks. Where the model will be used to make predictions beyond the limits of the main body of the data, this adds uncertainty.

1.3 Modeling the error distribution

With data such as here, it is to be expected that the response will show strong extra-binomial variation. This happens because the response varies from insect to insect, and/or because insects do not respond independently. The `glmmTMB` package (Brooks et al. 2017) implements the `betabinomial` error family, with

[1], accessed 29 May 2021

the option to model the scale parameter, and hence the multiplier for the binomial variance, as a function of explanatory variable(s). In the terminology used for R's `glm()` function, and that will be used in the sequel, the multiplier for the binomial variance is referred to as the dispersion factor. For the data considered here, this dispersion factor is high in the middle of the range of data, reducing at the extremes.

An alternative that will be investigated in a later section is the use of binomial errors, with observation level random effects used to account for differences at the level of individual observations. For the present data, this automatically achieves much the same effect as betabinomial errors, without the need for specific attention to the modeling of a dispersion factor. The model fit appears, however to be less satisfactory than the use of betabinomial errors with a dispersion factor adjustment.

Other error models that are described in the literature, and that operate at the level of individual observations, are discussed in the vignette `vignette("countDists", package = "qra")`.

1.4 Choices required for mixed model fits

Fits will require a choice of link functions, modeling of the fixed effects, and modeling of the error distribution. Because there are just three replicates per lifestage, it is necessary to base estimates on a model that brings together components of variance information across lifestages. This inevitably leads to estimates that will sometimes smooth out effects that are specific to an individual lifestage.

The subsection that follows will explore the use of a betabinomial error model, as implemented in the **glmmTMB** package. The parameterization is that described in Morris (1997), with parameters μ and ϕ. Here $\mathrm{E}[x] = n\mu$, and

$$\mathrm{var}[x] = n\mu(1-\mu)\frac{\phi+n}{\phi+1}$$

Setting $\phi = \frac{1-\rho}{\rho}$, where ρ is the intra-class correlation, this equals

$$n\mu(1-\mu)(1+(n-1)\rho)$$

Quasibinomial errors

The `lme4::glmer()` function offers the option of a quasibinomial error, as for `stats::glm()`. It does not offer an equivalent to the `glmmTMB` dispformula.

Specification of a quasibinomial error has the consequence that the model is fitted as for a binomial distribution, with the the binomial variance $n\pi(1-\pi)$ then multiplied by a constant factor Φ that is usually estimated using the Pearson chi-squared statistic. Compare this with the betabinomial, where the multiplier is $\Phi = 1 + (n-1)\rho$, i.e., it increases with n. This is an important difference.

1.5 Complementary log-log versus logit link

Figure 2 shows lines and curves from alternative models that have been fitted to the data.

One has to experiment to get these models to fit. Some of the possibilities are:

- Stay with default control parameters, or try an alternative, such as `ctl <- glmmTMBControl(optimizer=optim, optArgs=list(meth`
- Replace `TrtTime` by `scTime`, which is the centered and scaled version, in some or all of
  - the random effects term(s) in the model formula
  - the fixed effects terms
  - the dispersion formula (`dispformula`).
- As will be shown later, the dispersion formula for the `HawCon` data needs to be able to fit a curve that has a roughly hill shape. Degree 2 normal splines (`splines::ns(x,2)` , where in our case `x=scTime`) appear to work reasonably well.
- Use of `update()` to update a simpler model (e.g., to add the degree 2 term in models that added a curvature term to the straight line in the fixed effect) may avoid warning messages or failures that otherwise result.

Normal spline curves, or polynomial curves, may be used to model the fixed effects, as alternatives to lines. This can be compared with fitting a line, in order to check for curvature in the response.

If `glmer()` (from the `glmer` package) is used in place of `glmmTMB()`, the `lme4`

function `allFit()` can be used to compare results between a range of available optimizers.

Both the complementary log-log model and the logit model appear to fit well, as judged by examining plots of randomized quantile residuals.

Details of the model fitting process

Code that fits the several models is:

```
# Load packages that will be used
suppressMessages(
  {library(lme4); library(glmmTMB); library(splines);
    library(DHARMa)})
form <- cbind(Dead,Live)~0+trtGp/TrtTime+(1|trtGpRep)
form2s <- cbind(Dead,Live)~0+trtGp/TrtTime+ns(scTime,2)+(1|trtGpRep)
HCbb.cll <- glmmTMB(form, dispformula=~trtGp+ns(scTime,2),
                    family=betabinomial(link="cloglog"), data=HawCon)
HCbb2s.cll <- update(HCbb.cll, formula=form2s)
HCbb.logit <- glmmTMB(form, dispformula=~trtGp+ns(TrtTime,2),
                      family=betabinomial(link="logit"),data=HawCon)
HCbb2s.logit <- update(HCbb.logit, formula=form2s)
```

The right hand side of the model formula divides into two parts:

- The first part, i.e., `0+trtGp/TrtTime`, expands to `0 + trtGp + trtGp:TrtTime`. This specifies a different constant term and different slope, and thus a different line, for each different treatment group.
  - If the `0` is omitted, so that this initial part of the formula reduces to `trtGp/TrtTime`, all that changes is the parameterization. There is then an overall constant term, with treatment group effects expressed as differences from the overall constant term.
- The round brackets that enclose the remainder of the right hand side of the formula, i.e., `(scTime|trtGpRep)`, identify it as supplying random effects terms. Here, a different random effects line is added for each replicate (`trtGpRep`). This is achieved by fitting a random intercept and a random slope for each different replicate.
  - Note that `scTime` is by default interpreted as `1 + scTime`.

In addition, all models use a `dispformula` to allow for change in the betabino-

mial scale parameter ϕ. The `dispformula` used, allowing a different degree 2 function of `scTime` for each different treatment group, was chosen after some experimentation.

1.6 Fitted lines, vs fitted normal spline curves

Figure 2 shows fitted lines, and fitted degree 2 normal spline curves, in Panel A for a complementary log-log (cloglog) link, and in Panel B for a logit link.

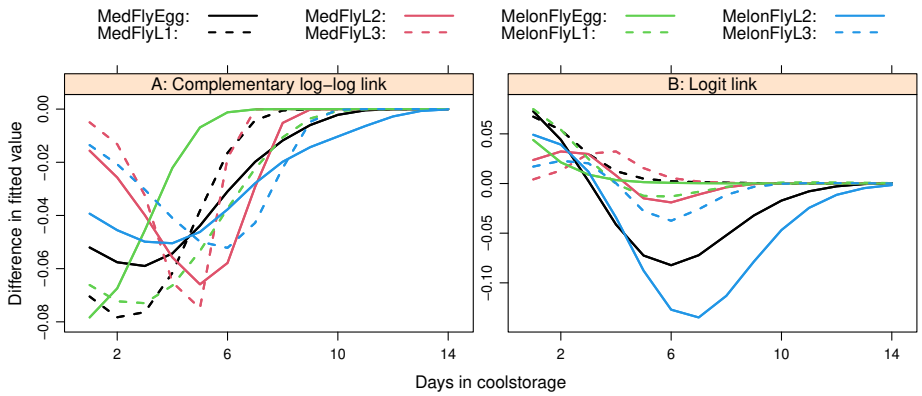

Figure 2: Differences are shown, between fitted degree 2 normal spline curves and fittes lines. Panel A is for the models that use a complementary log-log (cloglog) link, while Panel B is for a logit link.

1.7 Diagnostic checks

Randomized quantile residuals, with the plots that are available in the `DHARMa` provide helpful diagnostic check. For any residual, the corresponding quantile residual is the proportion of residuals expected, under model assumptions, to be less than or equal to it. If the distributional assumptions are satisfied, the quantile residuals should have a distribution that differs only by statistical error from a uniform distribution.

The function `DHARMa::simulateResiduals()` provides a convenient means to simulate enough sets of residuals (by default, 250) to give a good sense, for

each individual observation, of the distribution. These then provide a reference distribution for calculation of *quantile* residuals. Residuals may be calculated allowing only for the fixed effects (the default), or conditional on one or more levels of random effects. If the model is correct, residuals should be uniformly distributed irrespective of the conditioning. See `?DHARMa::simulateResiduals` for details.

- For the data as a whole, the distribution of residuals can be checked by plotting the quantile residuals against the corresponding quantiles.
  - Departures from assumptions show a pattern of difference from the line `y=x` that is different from that for normal distribution quantile-quantile plots.
- A second check plots quantile residuals against quantiles of predicted values. Quantile regression is then used to fit curves at 25%, 50%, and 75% quantiles of the quantile residuals. If the model is correctly specified, these should all be, to within statistical error, horizontal lines.
- Plots against other explanatory variables provide added checks.

Do such deviations from assumptions as are present matter?
A useful device is to simulate new 'observations' from the model, and check whether there is a difference of substance in the fitted values and parameter estimates.

Figure 3 shows the diagnostic plots for the linear model with a complementary log-log link.

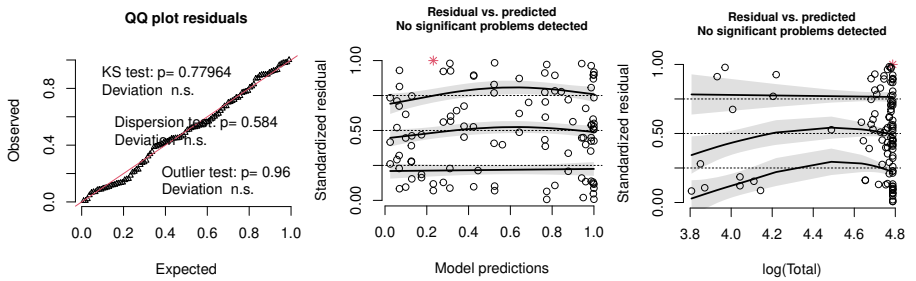

Figure 3: Panel A shows the quantile-quantile plot, for the linear model with a complementary log-log link. Panel B plots estimated quantiles against mortality, while Panel C plots estimated quantiles against total number, on a logarithmic scale.

The quantile-quantile (Q-Q) plot looks fine, The quantile residuals from the data appear, if anything, closer to uniformly distributed than any of the simulated sets of residuals. In Panels B and C, the quartiles of the data are appear satisfactorily close to the relevant quartiles.

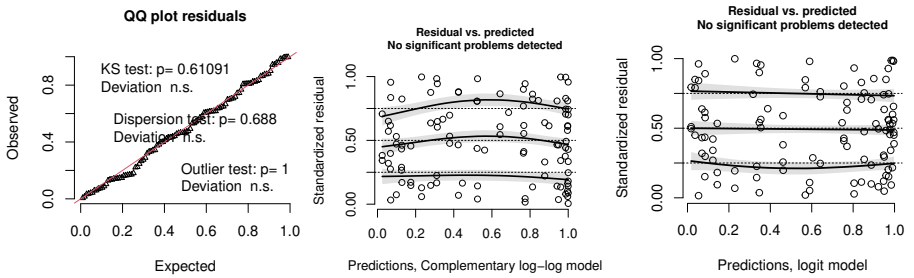

Figure 4: Diagnostic plots, for the model with a logit link.

Now compare (Figure 4) scaled residuals between treatment groups.

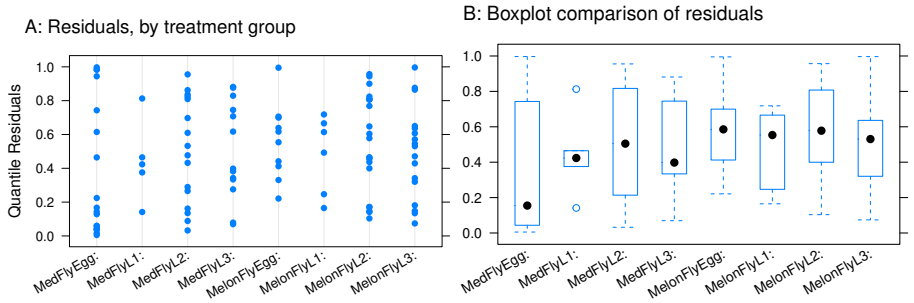

Figure 5: Quantile residuals, by treatment group, for the betabinomial model

The numbers for `MelonEgg`, `MedL1`, and `MelonL1` are too small to give useful boxplot displays. Except for `MedEgg`, where points are concentrated in the tails, the scatter of points in Panel A appears reasonably comparable between treatment groups.

QQ plots look good for all the models. In the sequel, they will be left out.

Uniform quantile-quantile plots — an example.

We will generate data from a highly overdispersed binomial type distribution, then examine the uniform quantile-quantile plot given by `DHARMa::plotResiduals()`. An easy way to generate overdispersed binomial type data is to start with binomial data, then multiply both "successes" and "failures" by a number that is substantially greater than one.

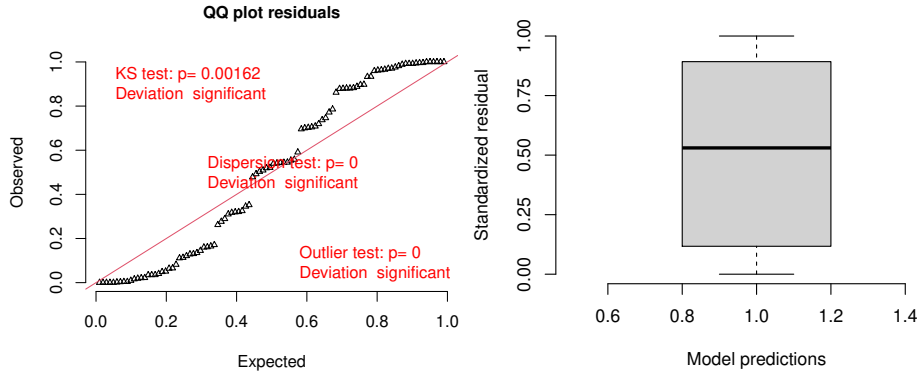

Figure 6: Diagnostics for model fitted to strongly overdispersed binomial type data. Notice that the overdispersion results in an S-shaped distribution of the residuals around the line $y = x$. The boxplot is, in this case, uninformative.

AIC-based model comparisons

Comparisons should be for models with the same outcome variable, and with the same data.[2]

In addition to the models described so far, we will include also models, to be discussed below, that assume binomial errors, with observational level random effects added. The model with the lowest AIC is the preferred model.

The following shows AIC values

```
    bb.cll bb2s.cll bb.logit bb2s.logit biObs.cll biObs.logit
df      27     29.0     27.0         29      20.0         20.0
```

[2]See https://robjhyndman.com/hyndsight/aic/ for a detailed commentary>

```
AIC     710     711.7     721.7          720      719.3         721.3
```

There is a strong preference for a complementary log-log link, with the linear fit slightly better than the degree 2 normal spline fit, arguing for the use of the fitted lines for calculation of LT99s and confidence intervals.

1.8 Estimates of ρ, and of the dispersion factor

oldpar <- par(oma=c(0,0,2,0))

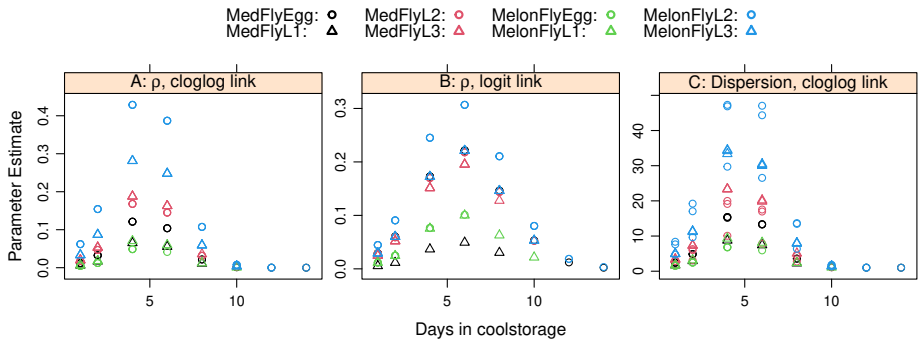

Figure 7: Panels A and B show intra-class correlation estimates for, respectively, a complementary log-log link and a logit link. Both models assume a betabinomial error. Panel C shows, for the complementary log-log model, the dispersion factors that result.

Figures 7A and B (for a logit link) plot the estimates of ρ, based on modeling the logarithm of the scale parameter as a degree 2 natural spline function of `scTime` that is added to a straight line response that is different for each different treatment group. Use of a logarithmic link (the default) has the consequence that effects that are additive on the link scale become multipliers when unlogged.

Figure 7C shows, for the complementary log-log model, the dispersion factors that result — high at midrange times and mortalities, reducing to close to a constant 1.0 at either extreme. Use of a normal spline basis helps ensures that the value for ρ, and hence the dispersion factor, extrapolates to a close to constant value at both extremes.

A dispersion factor that is close to 1.0 at the upper extreme of the data provides

a limited justification for assuming a binomial distribution, for purposes of calculating the sample size needed to demonstrate, at a 95% confidence level, a mandated survival rate of, e.g., no more than one in perhaps 100,000 insects.

2 Binomial errors, plus observation level random effects

The suggestion here is that for the replicates within each species/lifestage combination, the response in any one observation is close to binomial. Added to this is random variation between observations. By comparison with fitting a betabinomial or a quasibinomial error, the effect is to reduce the variance of the observed mortalities at low and high mortality points, relative to variance at midrange values.

The following fits the two models:

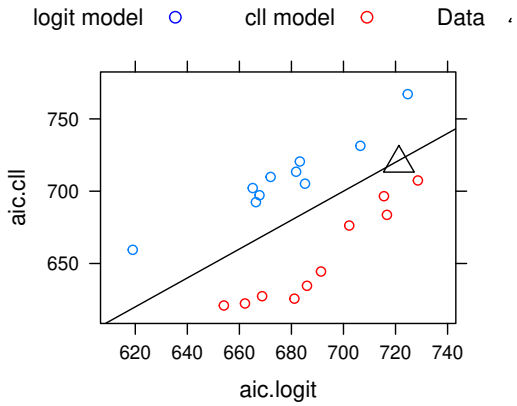

Figure 8: AICs are compared between models with binomial family errors and observation level random effects. The two sets of points make the comparison for, respectively, data that have been simulated from a model with logit link and a model with complementary log-log link. The large triangle makes the comparison for the models fitted to the 'HawCon' data.

Notice that the random effects term `(scTime|trtGpRep)` has been changed to `(1|trtGpRep)`. This avoids convergence failure messages.

The AIC statistic shows a slight preference for the complementary log-log model. The comparison with data that have been simulated from the respective fitted distributions indicates that the distinction is far from clear.

The attempt to repeat a similar comparison with models that assume a betabinomial error failed. Most simulationa generated error messages.

Figure 9 compares the diagnostic plots, between use of a complementary log-log link, and use of a logit link.

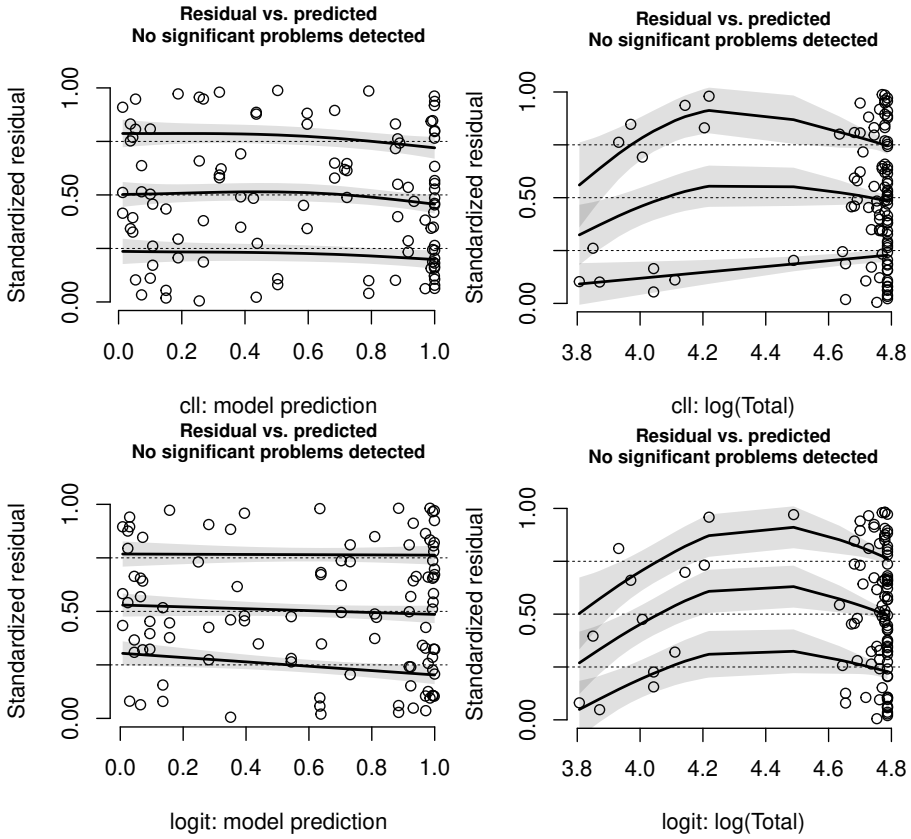

Figure 9: Diagnostics for model with binomial errors and observation level random effects.

There appears to be substantial heteroscedasticity that is not accounted for. If the observation specific random error could be modeled as a function of treatment time, it appears likely that this model would do just as well, or possibly even better, than the betabinomial model, with a model also for the dispersion factor.

Code is

```
set.seed(29)
simRefcll <- suppressWarnings(
  simulateResiduals(HCbiObs.cll, n=250, seed=FALSE) )
plotResiduals(simRefcll, xlab='cll: model prediction')
plotResiduals(simRefcll, form=log(HawCon$Total),
              xlab="cll: log(Total)")
simReflogit <- suppressWarnings(
  simulateResiduals(HCbiObs.logit, n=250, seed=FALSE) )
plotResiduals(simReflogit, xlab='logit: model prediction')
plotResiduals(simReflogit, form=log(HawCon$Total),
              xlab="logit: log(Total)")
```

3 Confidence intervals for ratios

Fieller's formula provides a methodology for calculating confidence intervals for ratios. Here, it will be turned to use for calculation of confidence intervals for exposures (times, or doses) required to kill 99% of insects.

3.1 99% Lethal time estimates and confidence intervals

The estimated time required to kill 99% of insects (lethal time 99%, or LT99) is commonly used as a starting point for assessing what time might be effective in commercial practice. Thus, for the model that used a complementary log-log link, and setting:

$$y = \log(-\log(1 - 0.99)) = 1.53,$$

one solves for $x = \text{LT99}$ in an equation of the form $y = a + bx$. Thus:

$$\text{LT99} = x = \frac{y - a}{b}$$

The determination of confidence intervals for such ratios is one of a much wider class of problems. The Fieller's formula approach (Morgan (2013)), implemented in the `qra` package, makes the common assumption that $(y - a,\ b)$ has been sampled from a distribution for which the bivariate normal is an adequate approximation. See `?qra::fieller`.

The sampling distribution of the calculated value x is typically, unless var$[b]$ is small relative to b, long-tailed. As a consequence, the *Delta* method, which uses an approximate variance as the basis for inference, is in general unreliable.

The Fieller's formula approach cannot in general be adapted to give confidence intervals for differences of LT99s or other such ratio statistics, unless the denominators of the statistics that are compared happen to be the same. A usable implementation of the simulation approach, which seems needed for the calculation of confidence intervals for LT99 differences, can in principle be handled using the function `lme4::bootMer()`.

3.2 What difference does the choice of model make?

We now proceed to investigate the damage done by assuming a constant dispersion factor, or by specifying a binomial error (there is no allowance for a dispersion factor), or by use of a logit link in place of a betabinomial link. Figure 10 compares confidence intervals, calculated using Fieller's formula, from models whose confidence intervals are identified thus:

- `BB-cll`, noting that `cll=cloglog`, identifies the preferred model, saved as `HCbb.cll`. The dispersion parameter is modeled as a degree 2 normal spline function of treatment time.
- `BIobsRE-cll`, saved as `HCbiObs.cll`, which adds observation level random effects to a binomial error.
- `BB-logit`, saved as `HCbb.logit`, replaces the complementary log-log link of item 1 with a logit link, while retaining the same form of degree 2 normal spline in `trtTime` for the dispersion parameter.
  - All estimates are then shifted upwards, in four cases with a much higher upper limit than for the preferred model.
- `BB-cll, const Disp factor` identifies a model that differs from `HCbb.cll` by fitting a constant dispersion factor;
  - This leads, in 7 out of 8 treatment groups, to a wider confidence interval.

- `Binomial-cll`, saved as `HCbin.LTcll` assumes binomial errors, with overdispersion accounted for in the between treatment component of variance;
  - The transfer of all extra-binomial variance to the between treatments component leads, for the treatment group `MelonL1`, to a dramatic increase in the confidence interval width.

In the two models (`HCbb.cll` and `HCbb.logit`) that assume a betabinomial error and model the variation in the variance, the dispersion factor parameter determines the intra-class correlation ρ.

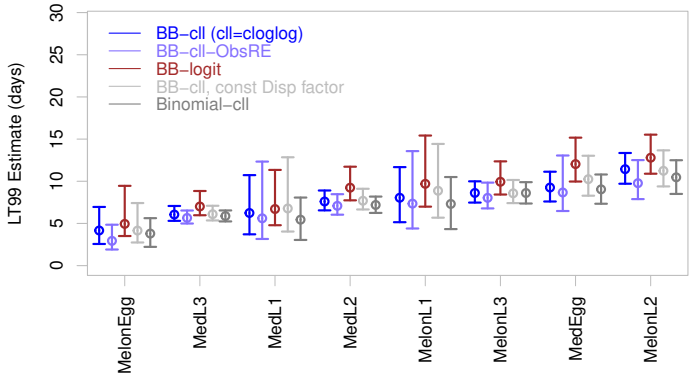

Figure 10: LT99 95% confidence intervals are compared between five different models.

Code to extract the 95% confidence interval for the LT99 is:

```
LTbb.cll <- qra::extractLT(p=0.99, obj=HCbb.cll, link="cloglog",
                                a=1:8, b=9:16, eps=0, df.t=NULL)[,-2]
rownames(LTbb.cll) <- shorten(rownames(LTbb.cll))
```

Other confidence intervals are calculated by modifying the call to `extractLT()` as required.

The message from Figure 10 is that, for purposes of estimating the LT99 or other high lethal time point, assumptions for the error family as well as for the link can make a large difference. Where those choices are made casually, without careful checking, serious biases may result. Use of a model that is unsatisfactory within the range of the available data is not a good starting point for extrapolation into

high mortality regions, with the additional uncertainties that then result.

4 Further models and model fitting functions

4.1 Binomial errors, observation level random effects, and more

This adds to the earlier observation level random effects model by allowing the associated variance to differ between species/lifestage/replicate combinations. Contributions to the variance at all levels other than the observation add to the binomial variance.

The code that follows implements such a model. It does not improve on the observation level random effects model demonstrated earlier.

4.2 Fits using `lme4::glmer()`

This is an alternative to `glmmTMB()`, for fits that assume a binomial error, plus observation level random effects. The AIC statistics differ slightly from those given using `glmmTMB()`, for reasons that are unclear.

```
fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```
    TMB:cll mer:cll TMB:lgt mer:lgt mer:cllCurve mer:lgtCurve
df     20.0    18.0    20.0    18.0         19.0         19.0
AIC   719.3   717.4   721.3   721.5        721.4        722.1
```

AIC statistics

The first two AIC values, for the models with a complementary log-log link, are 719.33 (for the `glmmTMB` model), and 717.38 (for the `glmer` model). The third and fourth values, both for the logit link, are 721.29 (`glmmTMB`) and 721.54 (`logit`). In both cases, `glmer` models that allow for overall curvature give an increased AIC statistic. The reason for the small difference between `glmer()` fits and `glmmTMB()` fits is unclear.

Control settings

Use of `glmer()` with `nAGQ=1` results in an AIC statistic that is closer to that from the `glmmTMB()` fit. The `glmer()` default has `optimizer = c("bobyqa", "Nelder_Mead")`, which has the effect of using the first optimizer for the first part of the calculation, then moving to use the second optimizer. See `?glmerControl`. Setting `control=glmerControl(optimizer='bobyqa')`, i.e., stay with "bobya" right through, avoids a "failed to converge" warning that for this fit occur with the default.

This choice was made after using the function `allFit()` to run the fit with a range of range of methods and optimizers, then checking for possible differences in the loglikelihoods and in the parameter estimates.

```
check <- (requireNamespace('dfoptim',quietly=TRUE)&
    requireNamespace('optimx',quietly=TRUE))
if(check)
ss<-suppressWarnings(summary(allFit(HCglmerBIobs.cll)))
```

```
stopifnot(check)
names(ss)
```

```
[1] "which.OK" "msgs"     "fixef"    "llik"     "sdcor"    "theta"
[7] "times"    "feval"
```

```
ss$msgs
```

```
$bobyqa
NULL

$Nelder_Mead
[1] "Model failed to converge with max|grad| = 0.155721 (tol = 0.002, component 1

$nlminbwrap
NULL

$nmkbw
[1] "Model failed to converge with max|grad| = 0.00217456 (tol = 0.002, component

$`optimx.L-BFGS-B`
NULL
```

```
$nloptwrap.NLOPT_LN_NELDERMEAD
NULL

$nloptwrap.NLOPT_LN_BOBYQA
NULL
```

```
ss$llik
```

```
                      bobyqa                      Nelder_Mead
                      -340.7                          -340.7
                  nlminbwrap                           nmkbw
                      -340.7                          -340.7
             optimx.L-BFGS-B nloptwrap.NLOPT_LN_NELDERMEAD
                      -340.7                          -340.7
   nloptwrap.NLOPT_LN_BOBYQA
                      -340.7
```

4.3 Gaussian errors, on a complementary log-log scale

A further model that might be tried is a linear mixed model, with $log(1-log((p+0.002)/(1+0.004)))$ as the dependent variable (complementary log-log link), and gaussian error. Figure 11 shows the plot of residuals versus predicted quantile deviations.

Table 1 compares the LT99 95% confidence interval estimates and bounds. Note the big differences for the Egg and L1 (larval stage 1) results.

The issue here is that all points where the data show 100% mortality transform to the same y-ordinate in the plot of $log(1-log((p+0.002)/(1+0.004)))$ against `TrtTime` .

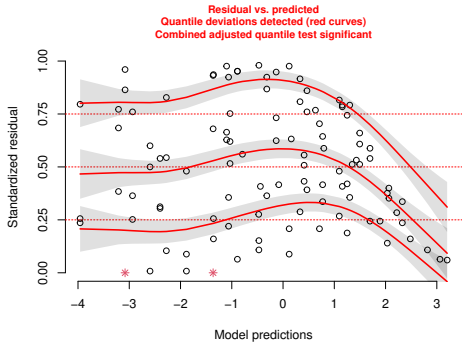

Figure 11: Residuals versus predicted quantile deviations, for the linear mixed model, with $log(1 - log((p + 0.002)/(1 + 0.004)))$ as the dependent variable, complementary log-log link, and gaussian error.

Table 1: Comparison of estimates and upper and lower 95% confidence limits, between the preferred betabinomial complementary log-log model and this model.

	Estimate		Lower		Upper	
	bb	gauss	bb	gauss	bb	gauss
MedEgg	9.3	10.5	7.6	8.0	**11.1**	**16.6**
MedL1	6.2	6.3	3.7	3.1	**10.7**	**16.9**
MedL2	7.6	7.6	6.6	6.7	**8.9**	**8.9**
MedL3	6.1	6.3	5.3	5.6	**7.1**	**7.2**
MelonEgg	4.2	3.9	2.6	1.7	7.0	15.8
MelonL1	8.0	8.1	5.1	4.9	11.7	15.9
MelonL2	11.4	10.5	9.7	8.9	13.4	12.8
MelonL3	8.6	8.6	7.5	7.5	10.0	10.0

5 Parting comments

5.1 At best, we have rough approximations to the dose-response

Quite strong assumptions, which could be checked only to a limited extent, have been made in order to get the results given. The checks that were performed made it clear that that the model that waqs finally chosen was not quite correct. The model that was chosen as "best" assumed that

- The pattern of mortality response is, on a complementary log-log scale, linear with time, consistently across the eight species/lifestage combinations.
  - We did check whether it was linear as opposed to a degree 2 normal spline, but the difficulties involved in getting a model to fit the limited data did not allow a check for a degree 3, or more complex, response pattern.

- Within replicate errors follow a betabinomial distribution, albeit with a dispersion parameter that varies with time in coolstorage.
  - The modeling of variation in the dispersion parameter ensures that, whether or not the betabinomial assumption is strictly correct, some limited account is taken of the pattern of change of variance with time. It is much less crude than assuming a binomial within replicate error, and using between replicate variation to account for extra-binomial variation, with no adjustment for variation with time in coolstorage.

- Results can be extrapolated to give the times needed to give, with some limited confidence, 99% or higher mortalities.
  - In order to gain any reasonable level of confidence that the model continues to give acceptably accurate predictions at the times required for 99% mortality levels and beyond, huge numbers of insects, spread across a vastly more replicates than in a study such as generated this dataset, would be required.

- Tolerance to treatment is constant within a lifestage.
  - This is unlikely to be strictly true.

Additionally, in the absence of provision for calculation of the Kenward-Roger

or other well-validated degrees of freedom approximation for models fitted using `glmmTMB()` or `glmer()`, the default that is used by `extractLT()` has to treated as a rough approximation. See Halekoh and Højsgaard (2014) for details of the Kenward-Roger approximation.

The results given should, accordingly, be used with caution. Common practice is to use such results to identify a "most tolerant" lifestage, and to suggest a treatment protocol, which is then be checked out in a large-scale trial.

One could have better confidence in models that had been checked out against a wide range of relevant data. For a given response probability π, does the multiplier for the binomial variance $n\pi(1-\pi)$ increase with n, as use of a betabinomial error assumes?

A useful first step would be creation of an open database into which researchers would be required, or at least strongly encouraged, to place their data. This would allow checking of model predictions for each specific treatment type, and for each class of pathogen, against the times found to give high mortality in large-scale trials. While there is an extensive literature that presents results of analyses from relevant trial data, and a very limited literature that makes comparisons across a number of different datasets, few of the relevant datasets are available in the public domain.

References

Brooks, Mollie E., Kasper Kristensen, Koen J van Benthem, Arni Magnusson, Casper W Berg, Anders Nielsen, Hans J Skaug, Martin Maechler, and Benjamin M Bolker. 2017. "glmmTMB Balances Speed and Flexibility Among Packages for Zero-Inflated Generalized Linear Mixed Modeling." *The R Journal* 9 (2): 378–400. https://journal.r-project.org/archive/2017/RJ-2017-066/index.html.

Follett, Peter A, and Lisa G Neven. 2006. "Current Trends in Quarantine Entomology." *Annu. Rev. Entomol.* 51. Annual Reviews: 359–85.

Halekoh, Ulrich, and Søren Højsgaard. 2014. "A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models – the R Package pbkrtest." *Journal of Statistical Software* 59 (9): 1–30. http://www.

jstatsoft.org/v59/i09/.

Morgan, Byron JT. 2013. *Analysis of Quantal Response Data*. Vol. 46. Springer.

Morris, William F. 1997. "Disentangling Effects of Induced Plant Defenses and Food Quantity on Herbivores by Fitting Nonlinear Models." *The American Naturalist* 150 (3). The University of Chicago Press: 299–327.