# A Meta-analysis of Longevity Estimates of Mosquito Vectors of Disease

Ben Lambert1,2, Ace North1 & H. Charles J. Godfray1

1 Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom

Corresponding author:

Phone: 01865 271176

2 Present address:

# Abstract

# Keywords

# Introduction

Some of the most important infectious diseases afflicting humans are transmitted by humans, including pathogens such as the causative agent of malaria that have been associated with humans throughout our evolutionary history as well recently emergent infections such as the Zika virus. Most mosquito species have a “gonotrophic cycle” involving successive episodes of vertebrate blood feeding, egg maturation and oviposition. In order for a mosquito to transmit a pathogen it must feed on an infectious person and live long enough to complete at least one gonotrophic cycle and feed on an uninfected and susceptible individual. Adult lifespan is thus a critical determinant of the ability of a mosquito population to allow the persistence of an indirectly transmitted infection. Lifespan can of course be straightforwardly assessed in the laboratory, but it is generally accepted that measurements under relatively benign laboratory conditions are likely to have limited relevance in the field, and much effort has been directed at estimating this parameter in the vector’s natural environment. Most work has focused on assessing average daily mortality rates, and the simplest assumption is that these do not vary with mosquito age – in this case longevity is simply the reciprocal of mortality. Testing this assumption and discovering whether mosquitoes senesce or show other types of age-dependent mortality has also been studied in the field.

There are two main strategies to estimate mosquito mortality rates and longevity. The first is through mark-release-recapture (MRR) experiments, a technique that is widely applied to estimate these parameters in many types of animal. As applied to mosquitoes, insects are caught in the field or reared in the laboratory and then marked, typically with fluorescent dust. The mosquitoes are then released into the field and then efforts are made to recapture them, for example using human baits or light traps, usually over an extended period of time. Mortality rates can be statistically estimated from the numbers of recaptures given certain assumptions. The main challenges with MRR is ensuring the marking technique does not affect recapture probability, and distinguishing mortality from mosquitoes dispersing out of range of being recaptured. Also, releasing insects that can transmit disease (especially if this increases ambient population levels) raises important ethical issues.

The second technique is specific to mosquitoes and makes use of their gonotrophic cycle. The mosquito ovary is made up of ovarioles, each of which typically produces one egg every gonotrophic cycle. After the egg passes into the oviduct the distended ovariole does not completely recover its previous form but a discrete dilation remains which can be detected by dissecting the female reproductive organs. Data on the fraction of females that have oviposited provides some information about mortality rates. However, a skilled dissector can distinguish the number of dilations from multiple gonotrophic cycles so providing much richer data on longevity. The challenges of this method include the amount of time and expertise it takes to collect the data, establishing the relationships between physiological and chronological time (though the distribution of the number of gonotrophic cycles wild-caught mosquitoes have gone through is of direct epidemiological relevance) and the fact that it only applies to females.

An issue with both methods is that they require logistically difficult and expensive field campaigns. There is thus value in conducting a meta-analysis of existing data to explore consistency across studies, identify correlates of lifespan and to learn lessons for further studies. Here we analyse data from 232 MRR and # dissection studies using a common statistical methodology. For MRR we make use of a very valuable database of 394 mosquito studies assembled by Guerra et al. (2014) while the dissection studies we extracted from the literature ourselves. We concentrated on the three major genera of mosquito vectors, *Anopheles, Aedes* (in its traditional sense) and *Culex*, which constitute the vast majority of the data.

# Methods

In recent years many important vectors of disease have been shown to be complexes of very closely related species, biotypes or forms that cannot be distinguished morphologically (for example the morphospecies *Anopheles gambiae sensu lato* is now separated into the widespread *gambiae, coluzzii, arabiensis* and a number of more local species). As the majority of studies analysed here took place before molecular techniques allowed these taxa to be separated we work here chiefly with morphospecies.

## Mark-release-recapture

Data from MRR experiments in the Guerra et al. (2014) database were examined and those with fewer than six recaptures and species with only a single MRR study excluded. Of the 232 data sets 177 involved only females, 35 males, and 18 both sex releases. For 102 data sets the age of the released mosquitoes was known while in the other cases it was unknown or unrecorded; in these cases we assumed the mosquitoes were newly emerged at the time of release and return to this assumption later.

We analysed all MRR experiments within the same statistical framework (for full details see the Supplementary Online Material). In the simplest case *NR* mosquitoes are released at day zero and the probability that they remain in the recapture area until day *t* is *S*(*t*) when they are recaptured with probability *ψ*. We model the number of mosquitoes recaptured on day *t* using a negative binomial sampling model with mean , where is cumulative captures before day *t*, and shape parameter *κ*. The negative binomial has been used previously in MRR analyses (ref) because of its ability to represent temporal over-dispersion in recaptures most likely caused by variable weather. A slight modification was required for studies with multiple releases (see SOM).

The simplest model for *S*(*t*) assumes there is a constant probability (*λ*) that a mosquito dies or leaves the recapture area so that the numbers remaining after time *t* are given by the exponential distribution, . We utilised this form extensively but in testing for senescence used five other models where *λ*(*t*) varies with time so that

|  |  |
| --- | --- |
|  | (1) |

Details of the five models (Gompertz, Weibull, Gompertz-Makeham, Logistic and Logistic-Makeham), which vary in their ability to detect different forms o`f age-dependent mortality, are given in the SOM. Using multiple different types of models increased are chances of detecting senescence though, as discussed below, also increases the likelihood of false positives.

Parameters were estimated using Bayesian techniques with relative uninformative priors for *κ* and the parameters of *λ*(*t*), but assuming a prior for indicating a low recapture probability (necessary for convergence). We used a Bayesian hierarchy model to estimate distributions of lifespan at the species and the genus levels, and across the complete data set. This procedure assumes that there is a distribution of lifespan parameters for each species from which those governing individual MRR time series are sampled, and similarly a distribution at the genus level from which those for individual species are derived, etc. (rather akin to random effects in classical statistics). Within this framework we can also allow the parameters for individual time series to be influenced by co-variates such as differences in experimental methodology. As in the estimation of the parameters of the individual experiments, relative uninformative priors were set for the parameters of the hierarchical models except for where again a distribution representing low recapture probabilities was assumed. Posterior distributions were derived using Markov Chain Monte Carlo (MCMC) methods with convergence assessed using the statistic (Gelmin & Rubin 1992). The predictive power of the model was assessed using *K*-fold cross validation which tests the ability of the model fitted to part of the data to predict the rest using multiple different partitions. Further details of the prior specification, fitting and validation are given in the SOM.

Two studies of *Anopheles balabacensis* reported capture rates increasing with time, presumably reflecting a violation of our assumption of constant recapture probabilities. We omitted this species from the analysis.

Add brief Methods for temperature study.

## Dissection

Studies using dissection to estimate mosquito longevity were located in literature databases using relevant keyword, citation and author searches, and by checking previous studies cited by the papers located. The list of studies located with associated metadata is available as a Supplementary Online File.

Most dissection studies recorded the distribution of the number of gonotrophic cycles in time series of mosquito samples. Overall, we found 568 time series in 72 published articles. Because seldom were sufficient ancillary data available data to analyse trends over time we aggregated the data from each time series. We further omitted time series with fewer than 100 mosquitoes and for species with only one data set leaving 131 studies of mosquitoes in the genera *Anopheles, Aedes*, *Culex* and *Mansonia*.

To compare lifespan estimates from dissection and MRR studies we need to convert physiological age (the number of gonotrophic cycles) into chronological age. Using a literature search and a review by Silver (2007) we found 79 estimates in 42 published articles. Most estimates were obtained by dissecting females recaptured in MRR studies or by observations in the laboratory, the latter tending to give longer times. Studies differed greatly in how (if at all) they represented uncertainty in their estimate of the duration of the gonotrophic cycle. Where confidence limits were given we treated these as the relevant quantiles of a normal distribution, where a range was stated (e.g. “4-6 days”) we interpreted the bounds as the 2.5% and 97.5% quantiles of a normal distribution, and where a single figure was quoted we assumed this was the median of a normal distribution with variance estimated from data from related species. Initially we calculated distributions of gonotrophic cycle lengths at the species and then genus levels, but because of the paucity of data for many species and the lack of significant differences we aggregated the data into a single distribution. We converted physiological age to chronological age by sampling from this distribution to obtain a particular gonotrophic cycle length for each mosquito (we also explored sampling from this distribution to obtain the duration of *each* gonotrophic cycle which increased the uncertainty in lifespan estimate but did not affect any of the conclusions).

We modelled the number of mosquitoes found by dissection to be of age *a* using the negative binomial distribution with mean and shape parameter *κ*, where *Ψ* is the product of the recruitment rate of adult mosquitoes, which we assume is constant over time, and the probability of being captured for dissection, and *S*(*a*) is the probability of surviving until age *a*. We used the number of females that have yet to lay eggs (nulliparous) to estimate the recruitment rate as described further in the SOM. Initial examination revealed that in some data sets the number of nulliparous females was anomalously low, something that has been noticed before (Gillies & Wilkes 1965). As some studies have suggested that the first gonotrophic cycle tends to be longer than the subsequent ones, this is probably due to differences in capture probability. In data sets where the fraction of nulliparous females was less than 90% the uniparous (completed on gonotrophic cycle) we excluded the nulliparous observation.

Data was analysed using a Bayesian framework as with the MRR data with minor differences in the specification of the priors (see SOM). Some published studies do not distinguish the number of gonotrophic studies beyond a threshold (the more ovariole dilations there are the harder it is to count them) which is akin to censoring the data. Because this censoring involves only a relatively small number of mosquitoes in each time series (median = 2%), and because of the technical difficulties of allowing for censoring in our Bayesian estimation procedures, we assume the females died at the threshold age.

# Results

## Lifespan estimates from MRR

MRR estimates the length of time a mosquito remains alive and is still in the area available for recapture. It thus provides a lower bound to lifespan which we shall refer to as LBL. In 211 of the 230 MRR time series the estimated LBL was less than 10 days (Fig. #). The smallest estimate was 1 day for the Asian malaria vector *Anopheles subpictus* which is unfeasibly short and almost certainly reflects dispersal out of the recapture zone or a violation of the assumptions of our analyses. The longest estimate was 14.5 days for the temperate species *Aedes cantans* which is not a human disease vector. There are multiple data sets for the most important vector species such as *Anopheles gambiae, Aedes aegypti* and *albopictus* and *Culex tarsalis* all of which show considerable variation. For example, there are # estimates of LBL for *A. aegypti* which range from # to # with a mean of # and coefficient of variation of #. There are significant differences in LBL amongst species (ANOVA on median LBL: *F*31 = 2.74, *p* <0.01; the non-parametric Kruskal Wallace was also significant).

The estimated LBL for *Culex, Anopheles* and *Aedes* were 2.4, 3.6 and 4.5 days respectively with an overall estimate of 3.7 days (Fig. #). Differences between genera were significant (ANOVA on median LBL: *F*1,2 = 4.74, *p* <0.01; Kruskal Wallace also significant). *K*-fold cross validation suggests that after the effect of genus is taken into account the incorporation of a species term provides little more predictive power (in part explained by the latter model over-fitting the data where there are few time series per species).

We reasoned that if dispersal out of the recapture area was reducing the LBL below the true lifespan then there should be a positive correlation between the spatial extent of the recapture zone and LBL. We found no such pattern (Fig. #). To pursue this it would be desirable to have further measures of recapture effort such as the density of traps or collection sites but this was not recorded in enough sites to make this analysis feasible. Guerra *et al.* (2014) record whether the mosquitoes in the different MRR studies were fed prior to release, and if so whether with blood, sugar or both. Analysing the # datasets with this information we found no significant association between lifespan and type of feeding.

The MRR experiments included a mixture of male-only and female-only releases, and releases of both sexes [am I right that in these cases you didn’t try to separate male and female recaptures to get sex-specific estimates? If so was this because the data wasn't there or because it would have been too hard to extract?]. We estimated average male and female LBL at the genus level (there were too few studies to make comparisons at the species level). There was a consistent trend for females to live longer than males (by about 0.8 days overall) but this was not significant.

To access whether temperature is associated with LBL we used weather records to calculate average temperatures at the MRR sites (see Methods). We found no significant relationship between study-site temperature and LBL (overall or within genus) for the # datasets we analysed (Fig. #).

## Number of gonotrophic cycles estimates from dissection

Dissection allows the number of completed gonotrophic cycles to be counted and from this the mean number of cycles before death was estimated. The average number of cycles completed in a lifetime across 131 studies is # with 95% less than 3 (Fig. #). The estimated greatest number of cycles was for *Anopheles sargentii* (2.5) which is adapted to desert conditions (it is known as the “oasis vector” of malaria) and may have evolved greater longevity. The important African malaria vector *An. gambiae s.l.* second (#). The smallest estimated mean number of gonotrophic cycles is for *Anopheles* *bellator* (0.5) which transmits malaria in Brazil’s Atlantic Forest. There are significant differences in estimated lifetime gonotrophic c amongst species (ANOVA: *F*24 = 2.24, *p* <0.01; the non-parametric Kruskal Wallace was also significant).

The estimated lifetime gonotrophic cycles for the different genera were *Anopheles,* 1.4; *Mansonia*, 1.1; *Culex,* 1.0; and *Aedes* 0.8 (Fig. #). The differences between the genera were significant (ANOVA on median number of cycles: *F*1,3 = 5.23, *p* <0.01; the non-parametric Kruskal Wallace was also significant, , *p* <0.01).

## Comparison of longevity estimates from two methods

To compare the two methods we converted numbers of gonotrophic cycle (physiological age) into lifespan (chronological age) as described in the Methods. For ten species we had enough data from both species to make a comparison and in # cases there was a significant difference, in all cases the dissection method suggesting a longer lifespan than MRR. I’m guessing this from looking at your figure 3.17 and will rewrite. My suggestion would be to plot Dissection against MRR lifespans for these ten species using two dimensional box-whisker plots. This has the added advantage that it is not duplicating the figures in the earlier sections. Given the few species I’m not sure looking at the genus level worth doing.

## Evidence for age-dependent mortality

The survival model upon which the above analyses are based is the single-parameter exponential model which assumes an age-invariant mortality hazard. We also fitted five multi-parameter models that allow, in different ways, mortality to vary with age. We did this to maximise our chance of detecting age-varying mortality (though aware of the risks of false positives with multiple estimations).

In Fig. # we compare the performance of the six models for describing lifespan in MRR studies of # species using K-fold cross-validation. In no cases is there a large difference in the performance of the multi-parameter model compared with the exponential. The equivalent analysis for the dissection studies is shown in Fig. #. In 14 of the 25 cases one of the more flexible models performed better (in the others the exponential was best or there was no clear winner). Examination of the 14 cases showed that the age-dependence always took the form of an increase in the risk of mortality with age.

We conclude that there is weak evidence at best for age-dependent mortality from studies of mosquitoes in the field.

## Estimates of the fraction mosquitoes capable of transmitting disease

We can use the posterior parameter estimates from our Bayesian analysis to estimate the fraction of mosquitoes that live beyond a certain age. In order to transmit a disease, a mosquito must live longer than the length of the intrinsic incubation period (the time taken for a pathogen ingested in one blood meal to be ready to be transmitted during a future feeding event). This is a lower bound as it does not include the waiting time to find a host after feeding or egg maturation. In Fig. # we plot the fraction of the mosquito population that pass this threshold using estimates from both MRR and dissection studies for the major vector species and their most significant diseases.

For malaria, estimates of the minimum fraction of the population that can transmit the disease vary from <0.1% for *An. subpictus* (as noted above likely to be due to the LBL substantially underestimating lifespan) to 46% for the drought-adapted and long-lived *An sergentii.* The figures for the two major African malaria vectors were *An. gambiae s.l.*:2%, MRR; and *An. funestus s.l.*: 4%, MRR. There were significant differences amongst species (ANOVA: *F*16,88 = 32.91, *p* <0.01; Kruskal Wallace , *p* <0.01).

*Aedes aegypti* and *Aedes albopictus* are the main vectors of dengue, chikungunya and Zika viruses. Because of their short intrinsic incubation periods a greater fraction of mosquito potentially live long enough to transmit diseases (Fig. #), rising to a maximum of 76% for *Ae albopictus* transmitting chikungunya.

# Discussion

* Summary
  + Uniform analysis framework
  + Mosquitoes live a relatively short time so normally only a small % transmit disease
  + MRR < dissection though difference not great
  + No association with other factors
  + Little evidence of age-dependent mortality
* Criticism of our study
  + Could we have missed patterns by using the same analysis technique on all datasets (i.e. should we have taken a more bespoke approach) Ben: could you think about this?
* Criticism of MRR
  + Marking
  + Conflation of mortality and dispersal (evidence from *Aedes*)
  + Statistical power [probably include Ben’s simulation]
  + Determinants of recapture rate [not sure about including Table 2.5]
* Criticism of dissection
  + Females only
  + Difficulty of method
  + New v Old School
  + Fluctuating population sizes
  + Sampling issues and nulliparous deficit
  + Phys & chron age and conversion
* Comparison of the two
  + UBL from laboratory studies
  + New methods (e.g. NIR)
* Age-dependent mortality
  + Lack of statistical power [Ben’s simulation]
  + Comparison with lab results
* Lack of association with other factors
  + Experimental design
  + Sex
  + Temperature

# Acknowledgements

Text

# References

# Figure Legends

Figure 1: