A comparison of the treatment groups:

1. statistically test whether infectivity changes between experiments.

* Mouse to mosquito is no different
* Mosquito to mouse depends on groups but small differences

1. if the answer to 1 is no then combine the data all together and then have a single estimate of the infection probabilities from mouse to mosquito and mosquito to mouse.

* *Combined and ran with 2 estimates detailed in table below*

1. does this change the efficacy estimates of the different interventions.

* Not considerably at all; if the infection probabilities are higher then efficacy estimates are also slightly higher but very little difference (can compare the data in Table 1 and 2)

1. statistically compare controls.

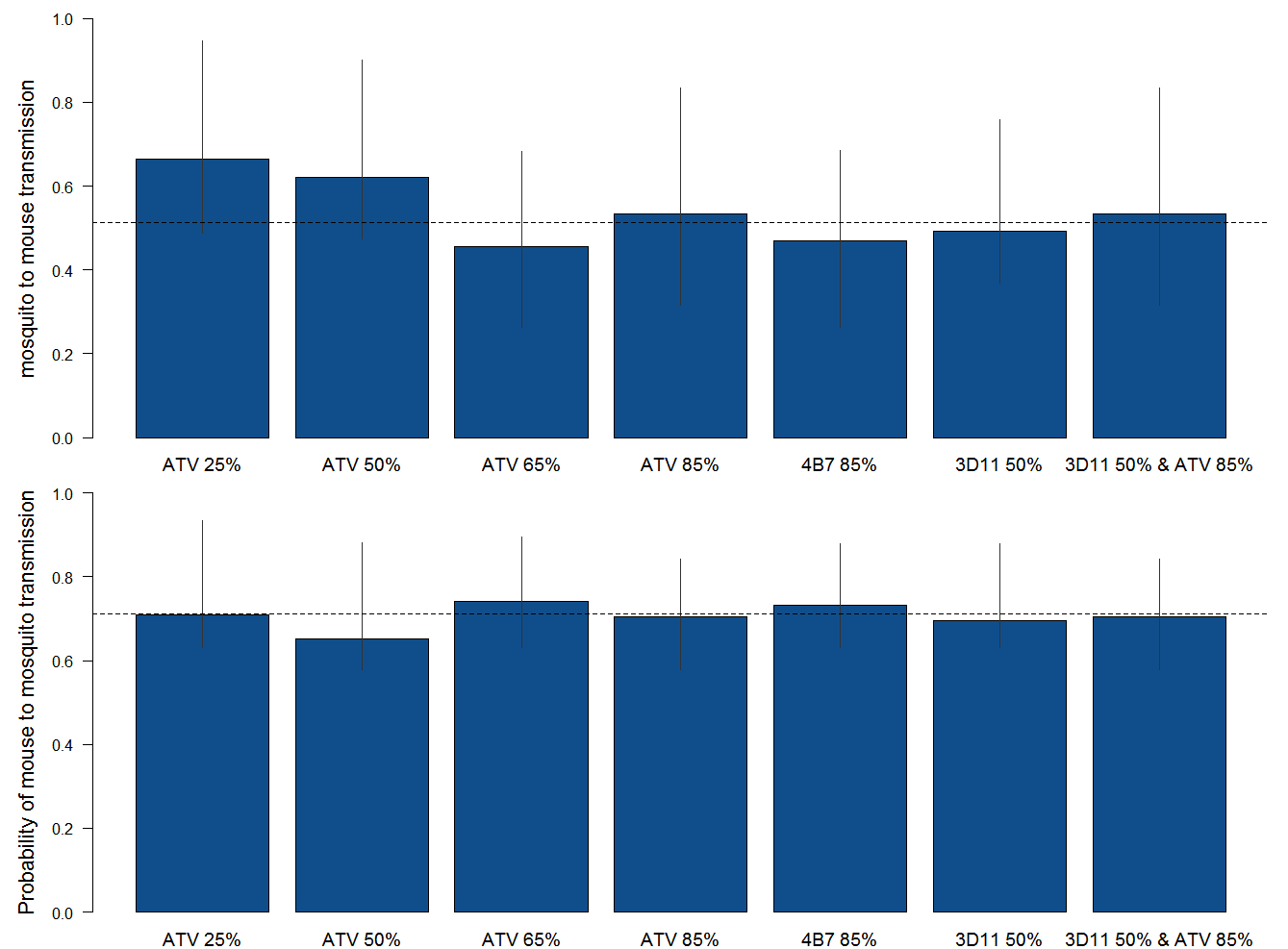
* No significant difference between parasitemia for Nat Comms controls vs new controls. However, borderline differences for gametocytemia – counts from the new controls were always below those for the NatComms data. \*But, difference for 1 bite group but no difference for 2 and 5 bites for prevalence (which is the important bit for the effect size chain binomial estimate)

**Table 1** Estimates for effect size allowing infectivity to vary for each treatment group

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | Efficacy (95% Confidence Interval) | Probability mosquito to mouse (95% CI) | Probability mouse to mosquito (95% CI) |
| ATV\_25% | 0 (0 - 0) | 0.632 (0.474 - 0.790) | 0.679 (0.632 - 0.737) |
| ATV\_50% | 0.025 (0.000 - 0.053) | 0.596 (0.474 - 0.737) | 0.628 (0.579 - 0.684) |
| ATV\_65% | 0.755 (0.684 - 0.842) | 0.443 (0.263 - 0.684) | 0.734 (0.632 - 0.895) |
| ATV\_85% | 0.969 (0.895 - 1.000) | 0.523 (0.316 - 0.842) | 0.698 (0.579 - 0.842) |
| 4B7\_85% | 0.836 (0.737 - 0.895) | 0.460 (0.263 - 0.684) | 0.727 (0.632 - 0.895) |
| 3D11\_50% | 0.233 (0.158 - 0.316) | 0.471 (0.368 - 0.579) | 0.682 (0.632 - 0.789) |
| 3D11\_50% & ATV\_85% | 0.969 (0.895 - 1.000) | 0.523 (0.316 - 0.842) | 0.698 (0.579 - 0.842) |

**COMPARISON OF INFECTIVITY**

Can we use a single estimate for the probability of transmission from an infected mosquito to a susceptible mouse (*r*) and the probability of transmission from an infected mouse to a susceptible mosquito (*s*)?



Performed anova on the estimated infectivity for each treatment group that were used to estimate the 95%CIs.

First for the probability of transmission from mosquito to mouse, using 3D11 50% as the comparative group as it sits in the middle of the range:

Fixed effects:

Estimate Std. Error t value

(Intercept) 0.53312 0.02728 19.55

Df Sum Sq Mean Sq F value Pr(>F)

data\_treatment 6 0.725 0.12090 5.811 1.16e-05 \*\*\*

Residuals 233 4.848 0.02081

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Call:

aov(formula = data\_r ~ data\_treatment, data = tested)

Residuals:

Min 1Q Median 3Q Max

-0.21725 -0.10129 -0.01974 0.07521 0.54386

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.49342 0.02944 16.758 < 2e-16 \*\*\*

data\_treatment3D11 50% & ATV 85% 0.03961 0.03619 1.095 0.27481

data\_treatment4B7 85% -0.02371 0.03549 -0.668 0.50476

data\_treatmentATV 25% 0.17165 0.05252 3.268 0.00125 \*\*

data\_treatmentATV 50% 0.12829 0.04655 2.756 0.00632 \*\*

data\_treatmentATV 65% -0.03728 0.03691 -1.010 0.31352

data\_treatmentATV 85% 0.03961 0.03619 1.095 0.27481

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1442 on 233 degrees of freedom

Multiple R-squared: 0.1302, Adjusted R-squared: 0.1078

F-statistic: 5.811 on 6 and 233 DF, p-value: 1.163e-05

The ATV 25% and 50% are different to the remaining treatment groups. Excluding these I could use the average for the remaining groups as the single estimate: r = 0.4976526

Including all the data I could use: r = 0.5135965

For the probability of transmission from an infected mouse to a susceptible mosquito (*s*):

Df Sum Sq Mean Sq F value Pr(>F)

data\_treatment 6 0.1266 0.021099 2.998 0.00769 \*\*

Residuals 233 1.6398 0.007038

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Call:

aov(formula = data\_s ~ data\_treatment, data = tested)

Residuals:

Min 1Q Median 3Q Max

-0.12542 -0.06360 -0.01535 0.04167 0.34868

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.695175 0.017124 40.596 <2e-16 \*\*\*

data\_treatment3D11 50% & ATV 85% 0.009192 0.021047 0.437 0.6627

data\_treatment4B7 85% 0.036701 0.020640 1.778 0.0767 .

data\_treatmentATV 25% 0.012959 0.030546 0.424 0.6718

data\_treatmentATV 50% -0.043860 0.027076 -1.620 0.1066

data\_treatmentATV 65% 0.045426 0.021466 2.116 0.0354 \*

data\_treatmentATV 85% 0.009192 0.021047 0.437 0.6627

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08389 on 233 degrees of freedom

Multiple R-squared: 0.07167, Adjusted R-squared: 0.04776

F-statistic: 2.998 on 6 and 233 DF, p-value: 0.007688

Small differences between the data. Using all data I can estimate *s* = 0.7125, excluding ATV 65% as the anomalous group *s* = 0.7065391

Estimating the effect sizes for each of the treatment groups with these fixed estimates I get:

**Table 2** Estimates for efficacy when infectivity is fixed

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | *r* estimate (mosquito to mouse) | *s* estimate (mouse to mosquito) | New effect size |
| ATV 25% | 0.4976526 | 0.7125 | 0.0075 (0.0000 – 0.0151) |
| ATV 50% | 0.4976526 | 0.7125 | 0.0628 (0.0050 – 0.1206) |
| ATV 65% | 0.4976526 | 0.7125 | 0.7714 (0.7186 – 0.8241) |
| ATV 85% | 0.4976526 | 0.7125 | 0.9673 (0.9347 – 1.0000) |
| 4B7 85% | 0.4976526 | 0.7125 | 0.8442 (0.7839 – 0.9045) |
| 3D11 50% | 0.4976526 | 0.7125 | 0.2864 (0.2261 – 0.3467) |
| 3D11 50% & ATV 85% | 0.4976526 | 0.7125 | 0.9673 (0.9347 – 1.0000) |
| ATV 25% | 0.5135965 | 0.7125 | 0.0075 (0.0000 – 0.0151) |
| ATV 50% | 0.5135965 | 0.7125 | 0.0729 (0.0151 – 0.1307) |
| ATV 65% | 0.5135965 | 0.7125 | 0.7764 (0.7236 – 0.8291) |
| ATV 85% | 0.5135965 | 0.7125 | 0.9673 (0.9347 – 1.0000) |
| 4B7 85% | 0.5135965 | 0.7125 | 0.8492 (0.7889 – 0.9095) |
| 3D11 50% | 0.5135965 | 0.7125 | 0.2990 (0.2412 – 0.3568) |
| 3D11 50% & ATV 85% | 0.5135965 | 0.7125 | 0.9673 (0.9347 – 1.0000) |
| ATV 25% | 0.4976526 | 0.7065391 | 0.0075 (0.0000 – 0.0151) |
| ATV 50% | 0.4976526 | 0.7065391 | 0.0578 (0.0000 – 0.1156) |
| ATV 65% | 0.4976526 | 0.7065391 | 0.7714 (0.7186 – 0.8241) |
| ATV 85% | 0.4976526 | 0.7065391 | 0.9673 (0.9347 – 1.0000) |
| 4B7 85% | 0.4976526 | 0.7065391 | 0.8618 (0.8090 – 0.9146) |
| 3D11 50% | 0.4976526 | 0.7065391 | 0.2814 (0.2211 – 0.3417) |
| 3D11 50% & ATV 85% | 0.4976526 | 0.7065391 | 0.9673 (0.9347 – 1.0000) |
| ATV 25% | 0.5135965 | 0.7065391 | 0.0075 (0.0000 – 0.0151) |
| ATV 50% | 0.5135965 | 0.7065391 | 0.0653 (0.0050 – 01256) |
| ATV 65% | 0.5135965 | 0.7065391 | 0.7764 (0.7236 – 0.8291) |
| ATV 85% | 0.5135965 | 0.7065391 | 0.9673 (0.9347 – 1.0000) |
| 4B7 85% | 0.5135965 | 0.7065391 | 0.8467 (0.7892 – 0.9095) |
| 3D11 50% | 0.5135965 | 0.7065391 | 0.2915 (0.2312 – 0.3518) |
| 3D11 50% & ATV 85% | 0.5135965 | 0.7065391 | 0.9673 (0.9347 – 1.0000) |

**COMPARSION OF CONTROL DATA**

Are the original controls synonymous with the new controls? Compared the individual biting rates 1, 2, and 5 where both datasets were complete

**OOCYSTS (for mosquito infection)** There is a significant difference in the number of oocysts and prevalence of oocysts for each biting rate

1 Bite

Model:

Oocyst ~ DATASOURCE

Df Deviance AIC LRT Pr(>Chi)

<none> 177.06 1165.1

DATASOURCE 1 321.81 1307.8 144.75 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

summary.lm(glm(OocPrev~DATASOURCE,data=con1))

Call:

glm(formula = OocPrev ~ DATASOURCE, data = con1)

Residuals:

Min 1Q Median 3Q Max

-0.3667 -0.3667 -0.1333 0.6333 0.8667

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.36667 0.03165 11.586 < 2e-16 \*\*\*

DATASOURCENEW -0.23333 0.04694 -4.971 1.08e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4246 on 328 degrees of freedom

Multiple R-squared: 0.07005, Adjusted R-squared: 0.06722

F-statistic: 24.71 on 1 and 328 DF, p-value: 1.077e-06

2 Bites

Model:

Oocyst ~ DATASOURCE

Df Deviance AIC LRT Pr(>Chi)

<none> 194.10 1337.9

DATASOURCE 1 197.36 1339.2 3.26 0.07099 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> summary.lm(glm(OocPrev~DATASOURCE,data=con2))

Call:

glm(formula = OocPrev ~ DATASOURCE, data = con2)

Residuals:

Min 1Q Median 3Q Max

-0.3473 -0.3473 -0.2600 0.6527 0.7400

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.34731 0.03561 9.752 <2e-16 \*\*\*

DATASOURCENEW -0.08731 0.05177 -1.686 0.0927 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4602 on 315 degrees of freedom

Multiple R-squared: 0.008947, Adjusted R-squared: 0.005801

F-statistic: 2.844 on 1 and 315 DF, p-value: 0.09271

5 Bites

Model:

Oocyst ~ DATASOURCE

Df Deviance AIC LRT Pr(>Chi)

<none> 259.21 1741.2

DATASOURCE 1 265.51 1745.5 6.2988 0.01208 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> summary.lm(glm(OocPrev~DATASOURCE,data=con5))

Call:

glm(formula = OocPrev ~ DATASOURCE, data = con5)

Residuals:

Min 1Q Median 3Q Max

-0.5778 -0.2064 -0.2064 0.4222 0.7935

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.57778 0.03400 16.992 < 2e-16 \*\*\*

DATASOURCENEW -0.37133 0.04999 -7.428 9.34e-13 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4562 on 333 degrees of freedom

Multiple R-squared: 0.1421, Adjusted R-squared: 0.1396

F-statistic: 55.18 on 1 and 333 DF, p-value: 9.336e-13

**BLOOD STAGE INFECTIONS (for mouse infection)** There was no significant difference between parasitemia for either control (p > 0.2 in all cases).

There was borderline significance for gametocytemia at the 95% level for each bite rate:

summary.lm(glm(Gametocytemia~DATASOURCE,data=conb1))

Call:

glm(formula = Gametocytemia ~ DATASOURCE, data = conb1)

Residuals:

Min 1Q Median 3Q Max

-0.1645 -0.1645 0.0000 0.0000 0.8755

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.16450 0.05325 3.089 0.00406 \*\*

DATASOURCENEW -0.16450 0.08134 -2.022 0.05131 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2381 on 33 degrees of freedom

Multiple R-squared: 0.1103, Adjusted R-squared: 0.08331

F-statistic: 4.09 on 1 and 33 DF, p-value: 0.05131

> summary.lm(glm(Gametocytemia~DATASOURCE,data=conb2))

Call:

glm(formula = Gametocytemia ~ DATASOURCE, data = conb2)

Residuals:

Min 1Q Median 3Q Max

-0.305 -0.305 -0.030 -0.030 1.655

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.30500 0.09562 3.190 0.00312 \*\*

DATASOURCENEW -0.27500 0.14606 -1.883 0.06857 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4276 on 33 degrees of freedom

Multiple R-squared: 0.097, Adjusted R-squared: 0.06964

F-statistic: 3.545 on 1 and 33 DF, p-value: 0.06857

> summary.lm(glm(Gametocytemia~DATASOURCE,data=conb5))

Call:

glm(formula = Gametocytemia ~ DATASOURCE, data = conb5)

Residuals:

Min 1Q Median 3Q Max

-0.6525 -0.6525 -0.0460 -0.0460 4.1575

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.6525 0.2160 3.021 0.00484 \*\*

DATASOURCENEW -0.6065 0.3299 -1.838 0.07505 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.966 on 33 degrees of freedom

Multiple R-squared: 0.09288, Adjusted R-squared: 0.0654

F-statistic: 3.379 on 1 and 33 DF, p-value: 0.07505

And contrasting differences for prevalence of blood stage infection (significant at 95% level for 1 bite but not 2 or 5 bites):

summary.lm(glm(bloodstage~DATASOURCE,data=conb1))

Call:

glm(formula = bloodstage ~ DATASOURCE, data = conb1)

Residuals:

Min 1Q Median 3Q Max

-0.4500 -0.4500 -0.1333 0.5500 0.8667

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.4500 0.1006 4.472 8.66e-05 \*\*\*

DATASOURCENEW -0.3167 0.1537 -2.060 0.0473 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.45 on 33 degrees of freedom

Multiple R-squared: 0.114, Adjusted R-squared: 0.0871

F-statistic: 4.244 on 1 and 33 DF, p-value: 0.04735

> summary.lm(glm(bloodstage~DATASOURCE,data=conb2))

Call:

glm(formula = bloodstage ~ DATASOURCE, data = conb2)

Residuals:

Min 1Q Median 3Q Max

-0.5000 -0.5000 -0.2667 0.5000 0.7333

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.5000 0.1096 4.561 6.69e-05 \*\*\*

DATASOURCENEW -0.2333 0.1675 -1.393 0.173

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4903 on 33 degrees of freedom

Multiple R-squared: 0.05556, Adjusted R-squared: 0.02694

F-statistic: 1.941 on 1 and 33 DF, p-value: 0.1729

> summary.lm(glm(bloodstage~DATASOURCE,data=conb5))

Call:

glm(formula = bloodstage ~ DATASOURCE, data = conb5)

Residuals:

Min 1Q Median 3Q Max

-0.7500 -0.4667 0.2500 0.2500 0.5333

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.7500 0.1065 7.043 4.64e-08 \*\*\*

DATASOURCENEW -0.2833 0.1627 -1.742 0.0908 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4762 on 33 degrees of freedom

Multiple R-squared: 0.08421, Adjusted R-squared: 0.05646

F-statistic: 3.034 on 1 and 33 DF, p-value: 0.09083