# Logistic regressions on probability of abortion vs. birth

Doing some initial *easier* models of the probability of abortion vs. birth. Doing this so that I don't have to do a ton of different MCMC models.

## Elk\_ID Feedground CollarType GPSCollar  
## Min. : 1 Soda : 57 GPS :239 No :217   
## 1st Qu.:170 Dell : 51 None:109 Yes:239   
## Median :332 Buffalo: 43 Prox: 4   
## Mean :362 Scab : 37 VHF :104   
## 3rd Qu.:570 Greys : 36   
## Max. :769 Bench : 34   
## (Other):198   
## CaptureDate Start\_Interval   
## Min. :2006-02-07 00:00:00 Min. :2006-02-07 00:00:00   
## 1st Qu.:2007-01-18 12:00:00 1st Qu.:2007-01-26 00:00:00   
## Median :2010-01-26 00:00:00 Median :2010-05-07 00:00:00   
## Mean :2009-06-01 20:03:09 Mean :2009-08-20 10:15:47   
## 3rd Qu.:2011-03-12 00:00:00 3rd Qu.:2011-06-05 00:00:00   
## Max. :2013-03-05 00:00:00 Max. :2013-07-02 00:00:00   
##   
## Stop\_Interval TrustRating VIT\_Status PET   
## Min. :1961-06-01 00:00:00 1:156 Abortion : 29 Bad :322   
## 1st Qu.:2007-05-29 00:00:00 2:187 Died : 0 Good:134   
## Median :2010-05-25 00:00:00 4: 5 Parturition:427   
## Mean :2009-08-14 17:53:41 5:108   
## 3rd Qu.:2011-06-10 00:00:00 6: 0   
## Max. :2013-07-10 00:00:00 7: 0   
##   
## CowSero CapDOY StartDOY EndDOY interval1   
## NEG :323 Min. : 5.0 Min. : 15.0 Min. : 45 Min. : 0.0   
## POS :132 1st Qu.:27.0 1st Qu.: 72.8 1st Qu.:146 1st Qu.: 0.0   
## NA's: 1 Median :46.0 Median :146.0 Median :153 Median : 98.0   
## Mean :42.7 Mean :122.2 Mean :153 Mean : 79.6   
## 3rd Qu.:57.0 3rd Qu.:154.0 3rd Qu.:161 3rd Qu.:117.0   
## Max. :81.0 Max. :205.0 Max. :221 Max. :162.0   
##   
## interval2 event left right interval   
## Min. : 1.0 Min. :1 Min. : 11.0 Min. : 41 Min. : 1.0   
## 1st Qu.: 1.0 1st Qu.:1 1st Qu.: 68.8 1st Qu.:142 1st Qu.: 1.0   
## Median : 7.0 Median :1 Median :142.0 Median :149 Median : 7.0   
## Mean : 30.4 Mean :1 Mean :118.2 Mean :149 Mean : 30.4   
## 3rd Qu.: 32.2 3rd Qu.:1 3rd Qu.:150.0 3rd Qu.:157 3rd Qu.: 32.2   
## Max. :206.0 Max. :1 Max. :201.0 Max. :217 Max. :206.0   
##   
## b.type abort   
## Min. :0.0000 Min. :0.0000   
## 1st Qu.:0.0000 1st Qu.:0.0000   
## Median :0.0000 Median :0.0000   
## Mean :0.0636 Mean :0.0636   
## 3rd Qu.:0.0000 3rd Qu.:0.0000   
## Max. :1.0000 Max. :1.0000   
##

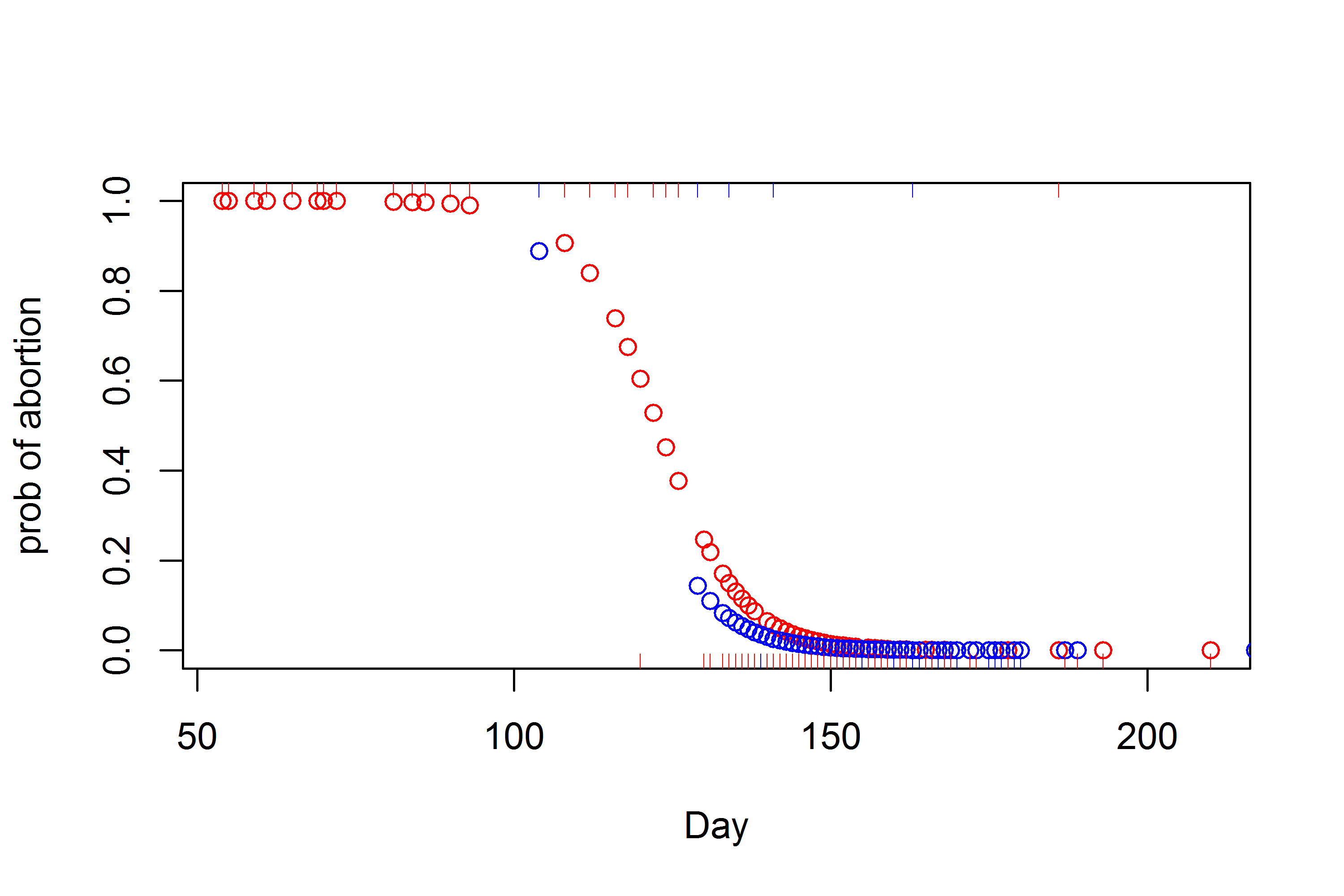
##   
## Abortion Died Parturition  
## NEG 7 0 316  
## POS 22 0 110

For simplicity we'll just assume that all events occur at the end of the interval

##   
## Call:  
## glm(formula = abort ~ right + CowSero, family = "binomial", data = Type)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.362 -0.212 -0.115 -0.053 4.417   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 18.1098 4.4372 4.08 4.5e-05 \*\*\*  
## right -0.1542 0.0314 -4.91 9.3e-07 \*\*\*  
## CowSeroPOS 0.8169 0.7042 1.16 0.25   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 215.785 on 454 degrees of freedom  
## Residual deviance: 76.121 on 452 degrees of freedom  
## AIC: 82.12  
##   
## Number of Fisher Scoring iterations: 8

This is interesting. Seropositives are more likely to abort, but once you account for the timing, you may not need that variable.

## Warning: some values will be clipped  
## Warning: some values will be clipped



plot of chunk plot-model1

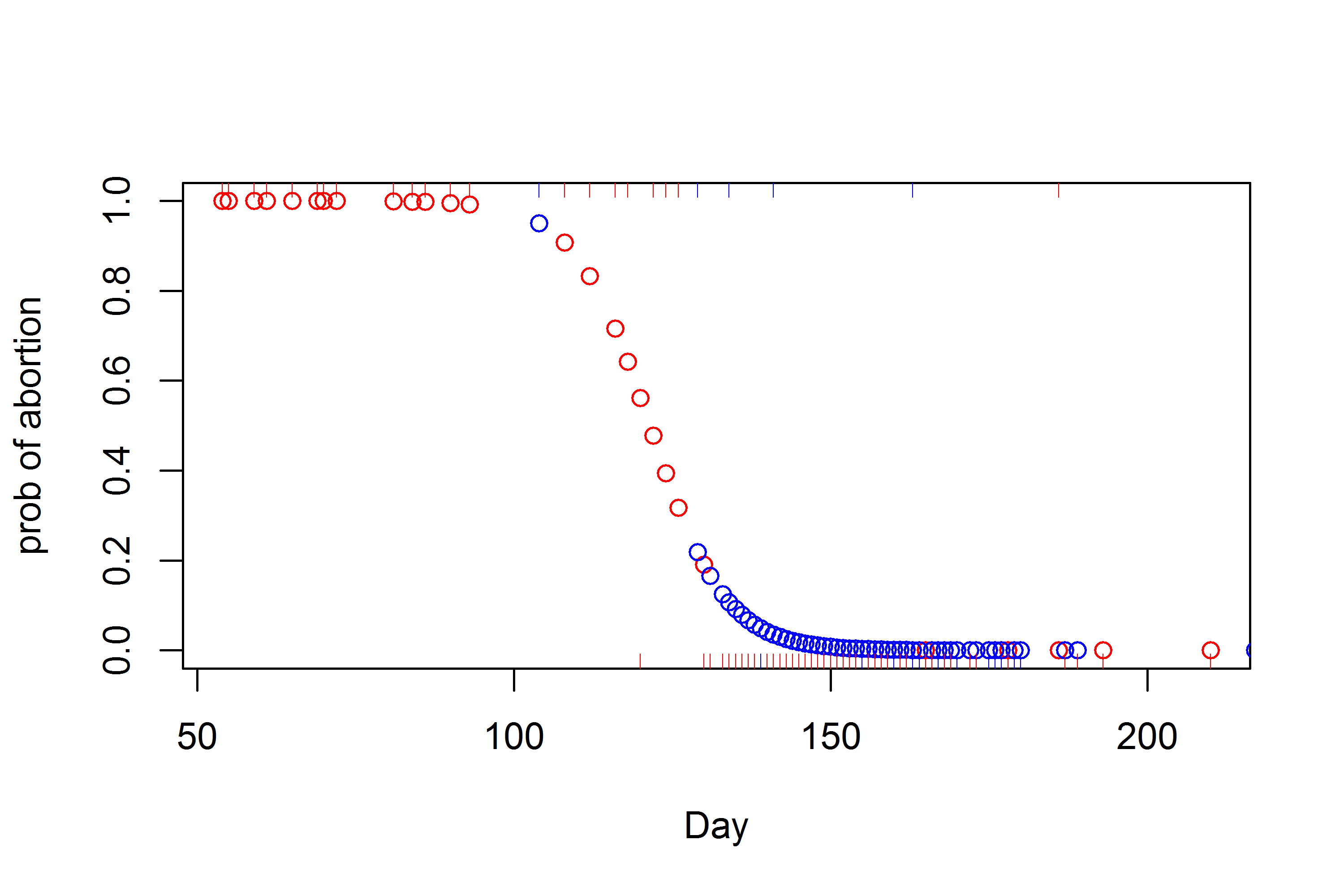
##   
## Call:  
## glm(formula = abort ~ right + CowSero + right \* CowSero, family = "binomial",   
## data = Type)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.353 -0.211 -0.113 -0.051 4.399   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 18.59464 8.23278 2.26 0.0239 \*   
## right -0.15765 0.05861 -2.69 0.0071 \*\*  
## CowSeroPOS 0.13852 9.62315 0.01 0.9885   
## right:CowSeroPOS 0.00491 0.06947 0.07 0.9436   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 215.785 on 454 degrees of freedom  
## Residual deviance: 76.116 on 451 degrees of freedom  
## AIC: 84.12  
##   
## Number of Fisher Scoring iterations: 8

##   
## Call:  
## glm(formula = abort ~ right, family = "binomial", data = Type)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.284 -0.209 -0.116 -0.052 4.672   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 20.5395 4.2476 4.84 1.3e-06 \*\*\*  
## right -0.1691 0.0311 -5.44 5.4e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 215.785 on 454 degrees of freedom  
## Residual deviance: 77.434 on 453 degrees of freedom  
## AIC: 81.43  
##   
## Number of Fisher Scoring iterations: 8

## df AIC  
## model0 2 81.43  
## model1 3 82.12  
## model2 4 84.12

So it looks like the interaction term is unnecessary. and even CowSero is unnecessary

## Warning: some values will be clipped  
## Warning: some values will be clipped



plot of chunk plot-model0

But if you don't have the data on when the event occurred. Then Seropositive status is important:

##   
## Call:  
## glm(formula = abort ~ CowSero, family = "binomial", data = Type)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.604 -0.209 -0.209 -0.209 2.768   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.810 0.382 -9.97 <2e-16 \*\*\*  
## CowSeroPOS 2.200 0.448 4.91 9e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 215.79 on 454 degrees of freedom  
## Residual deviance: 186.44 on 453 degrees of freedom  
## AIC: 190.4  
##   
## Number of Fisher Scoring iterations: 6

So the probability ofa seropositive aborting is 0.17 plus or minus a SE of 0.03. Compared to a seronegative: 0.02 plus or minus a SE of 0.01.

This would be on the order of reducing an 80% birth rate to 74% if the overall prevalence is 30%. (so not too big).

Any evidence of a feedground effect?

##   
## Call:  
## glm(formula = abort ~ CowSero + Feedground, family = "binomial",   
## data = Type)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.9005 -0.3390 -0.0001 0.0000 2.6277   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.03e+00 1.11e+00 -2.74 0.00612 \*\*   
## CowSeroPOS 1.92e+00 5.31e-01 3.61 0.00031 \*\*\*  
## FeedgroundBench 2.05e-01 1.25e+00 0.16 0.87003   
## FeedgroundBlack -1.75e+01 7.93e+03 0.00 0.99824   
## FeedgroundBuffalo -1.82e+01 2.58e+03 -0.01 0.99437   
## FeedgroundCamp 2.34e+00 1.65e+00 1.42 0.15634   
## FeedgroundDell -8.15e-02 1.17e+00 -0.07 0.94465   
## FeedgroundDog -1.88e+01 1.19e+04 0.00 0.99874   
## FeedgroundFall -1.75e+01 4.74e+03 0.00 0.99705   
## FeedgroundFinnegan -1.80e+01 6.02e+03 0.00 0.99762   
## FeedgroundForest 2.45e-01 1.36e+00 0.18 0.85638   
## FeedgroundFranz 5.74e-02 1.37e+00 0.04 0.96660   
## FeedgroundGreys -3.88e-01 1.23e+00 -0.31 0.75291   
## FeedgroundHA99 -1.77e+01 3.36e+03 -0.01 0.99579   
## FeedgroundHorse -1.88e+01 8.44e+03 0.00 0.99822   
## FeedgroundJewett -1.81e+01 4.10e+03 0.00 0.99647   
## FeedgroundMcNeel -1.29e+00 1.52e+00 -0.85 0.39520   
## FeedgroundMuddy -1.78e+01 4.20e+03 0.00 0.99662   
## FeedgroundPatrol -1.83e+01 4.85e+03 0.00 0.99699   
## FeedgroundScab -1.81e+01 2.79e+03 -0.01 0.99481   
## FeedgroundSoda 3.33e-02 1.18e+00 0.03 0.97740   
## FeedgroundSouth -1.87e+01 5.61e+03 0.00 0.99734   
## FeedgroundSpring -1.75e+01 5.91e+03 0.00 0.99763   
## FeedgroundUG -1.80e+01 6.02e+03 0.00 0.99762   
## FeedgroundUGNWR -1.75e+01 6.70e+03 0.00 0.99791   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 215.79 on 454 degrees of freedom  
## Residual deviance: 155.96 on 430 degrees of freedom  
## AIC: 206  
##   
## Number of Fisher Scoring iterations: 19

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: abort  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev  
## NULL 454 216  
## CowSero 1 29.4 453 186  
## Feedground 23 30.5 430 156

Not really.