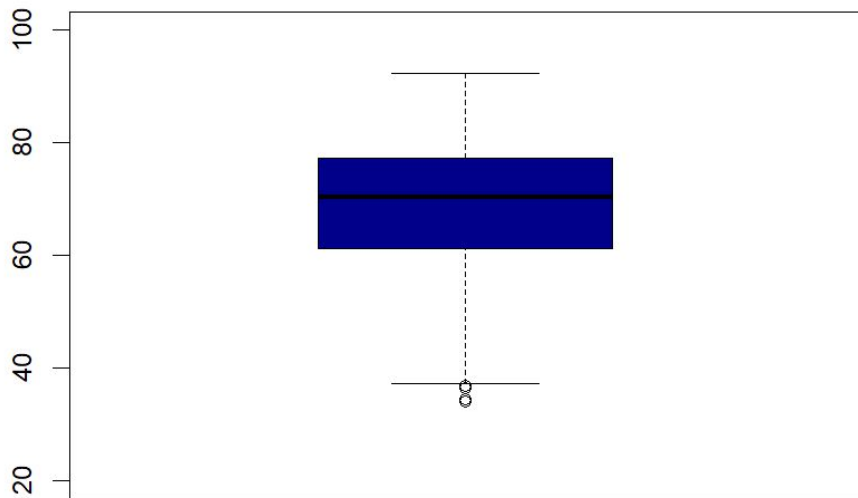


Inverse Probability Treatment Weighting

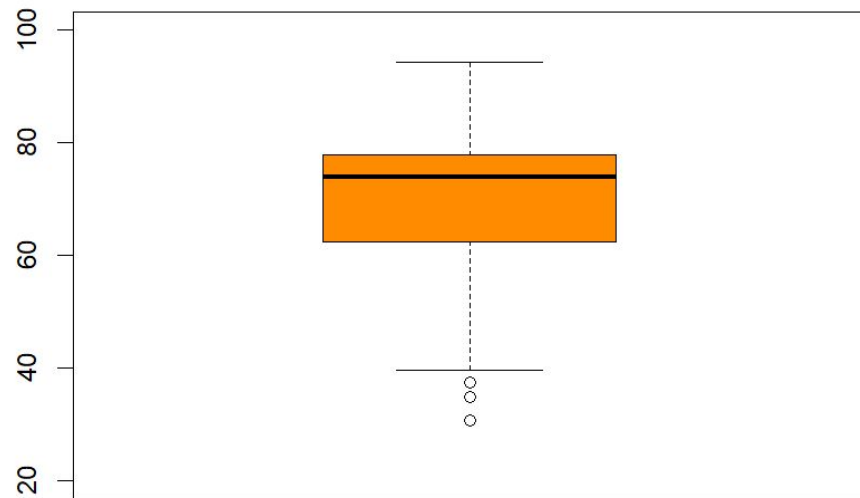
We Are Interested In Addressing Sampling Bias Among Those Who Had Samples Taken For Amyloidosis

It was initially thought that age would be a confounding factor

Boxplot of Sampled Age



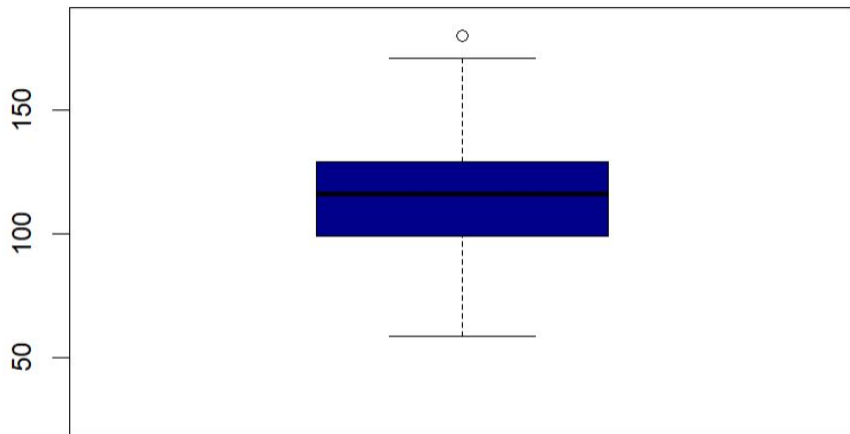
Boxplot of Unsampled Age



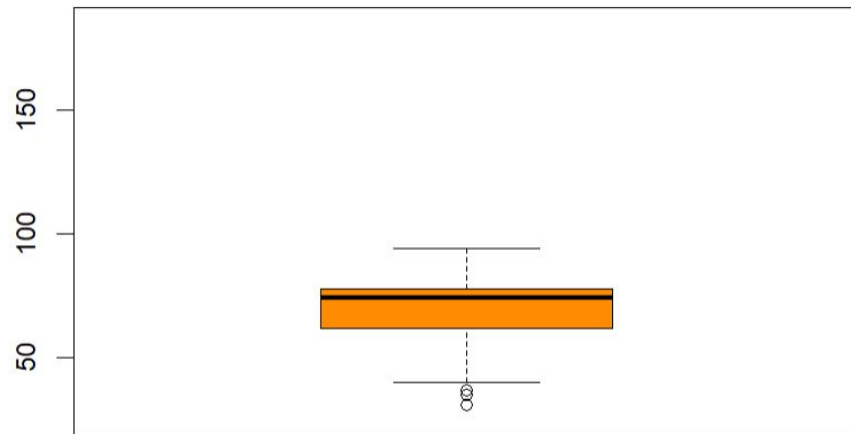
We Are Interested In Addressing Sampling Bias Among Those Who Had Samples Taken For Amyloidosis

Then it points became of interest

Boxplot of Sampled Points



Boxplot of Unsampled Points



Clearly there is bias, those sampled are at much higher risk

To correct for this we will use

Inverse Probability Treatment Weighting

Calculates weights used in model as inverse of propensity score

In our model we used the formula $\text{weights} = 1/p(\text{sampled} \mid \text{points})$

We fit two models

Model 1: $\text{Sample} \sim \text{Points}$, this model was used on the large data set

Model 2: $\text{Amyloid Status} \sim \text{Points} + \text{Severe}$

- Given weights found from model 1
- Fit this model to the dataset containing only those who were sampled for amyloidosis to correct for the sampling bias

Model 1: Sample ~ Points

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-15.61627	1.89884	-8.224	< 2e-16	***
Points	0.18007	0.02224	8.095	5.71e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 468.70 on 340 degrees of freedom
Residual deviance: 151.13 on 339 degrees of freedom
AIC: 155.13

Model 2: Amyloid Status ~ Points + Severe with weights

Call:

```
glm(formula = amyloid_status ~ Points + Severity, family = quasibinomial,  
     data = amyloid, weights = weight)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-8.07338	1.56068	-5.173	5.93e-07	***
Points	0.04271	0.01285	3.325	0.001066	**
SeveritySevere	2.19651	0.55614	3.950	0.000111	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1