Question 2

The first thing we do is try to built a model that could predict recall depends on cancer age history symptoms menopause and density.However, we found that if we predict that none of them had cancer, we will have an accuracy that performs better than any model we used to predict. But this means that we miss all cancer patients and have a high false negative . If we ask the radiologists to recall all the patients, we will find all cancer patients but we will have a very high false positives. We found that it is hard to minimizing false negatives and false positives at the same time. We think that it is more important to minimizing false negatives than minimizing false positives.

First question：

The models we used get the results(RMSE) that as follows:

result result result result result result result

0.1493401 0.1857107 0.1514213 0.1499492 0.1575127 0.1580203 0.1651269

Thus, we choose model1.

Based on our test, we find that there are some radiologists more clinically conservative than others in recalling patients. We know the problem that the radiologists do not see the same patients. So we build “out of sample test”. Our model simulates radiologists’ judgment. After training, If the model turns out to be valid in raw data, then it should performs the same in 'out-of-sample test' which is a different set of data. We asked all radiologists to look at out of samples to see if their recall rate is still the same as that of raw data.

The higher prob recall means the more conservative the radiologist is

# A tibble: 5 x 2

radiologist Prob\_recall

<chr> <dbl>

1 radiologist13 0.139

2 radiologist34 0.0872

3 radiologist66 0.191

4 radiologist89 0.211 **most conservative**

5 radiologist95 0.0998

We re-run the process many times, it shows that the radiologist89 is the most conservative and radiologist34 is the most not conservative in many repeated results.