

Precision and recall are evaluating measures for models and are functions of a confusion matrix. Confusion matrices have 4 labels:

1. true positives (when output is that patient has cancer and he actually does)
2. true negatives (output is patient doesn't have cancer, and they actually don't)
3. false positives (output says patient has cancer, but they actually don't)
4. false negatives (output says patient doesn't have cancer, but they actually do)

The model's precision is the following ratio:

$\text{true positives} / (\text{true positives} + \text{false positives})$

What precision tells us therefore is how often, when the model says a patient has cancer, do they actually have it.

The model's recall is the following ratio:

$\text{true positives} / (\text{true positives} + \text{false negatives})$

What recall tells us is how often, when the patient actually has cancer, does the model correctly predict them as having it.

In case of a fatal cancer, an incorrect classification could mean death. The differentiating factor among the 2 evaluating measures is false positives in precision against false negatives in recall. False positives indicate the model predicting cancer when the patient does not actually have it, while false negatives indicate the patient having cancer but our model incorrectly predicting them as not having it. The latter is of more significance to us as it is important to know how often our model might misclassify a patient as not having cancer when they actually do. Therefore, recall is more important here.