ROC Curve analysis of manually classified cellular masks

Training of the counting algorithm.

The limits for acceptance of positivity were selected after a ROC analysis (Feinstein 2002) of the measurements from ~300 random cellular masks from random images. We measured Mean grey value, Integrated density and thresholded Area fraction within the masks and compared them to the classification by an experienced researcher on the original and SME-projected images, which was used as "ground truth" ("golden standard").

```
#loading needed libraries
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.1.1       v purrr   0.3.2
## v tibble  2.1.1       v dplyr   0.8.0.1
## v tidyr   0.8.3       v stringr 1.4.0
## v readr   1.3.1       v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------ tid
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
require(pROC)
```

```
## Loading required package: pROC
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
##
##     cov, smooth, var
```

```
require(Epi)
```

```
## Loading required package: Epi
```

1. Measurements and correspnding manual classification is loaded in three separate dataframes: for mRNA FISH, Cytoplasmic antigen and Nuclear antigen channel. The assigned by the operator class is in column "logic...".

```
load("training_data_june2018.RData", verbose = T)
```

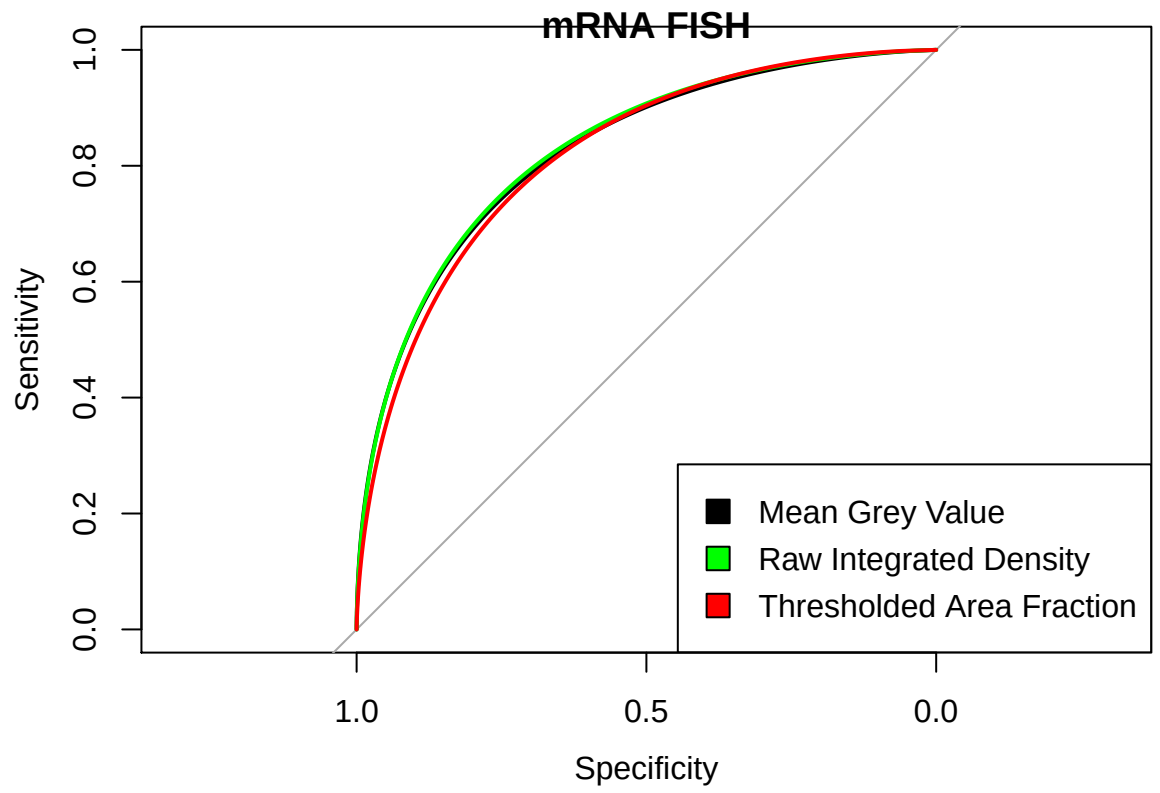
```
## Loading objects:
##   Z_score_FISH_pattern_jun2018_only_man
##   Z_score_IHC_pattern_jun2018_only_man
##   Z_score_BrdU_pattern_jun2018_only_man
```

2. Constructing ROC-curves With the help of the pROC library we construct the Receiver Operating Characteristics curves for the selected measurements in each channel separately: Mean grey value, Raw Integrated Density and Thresholded Area fraction:

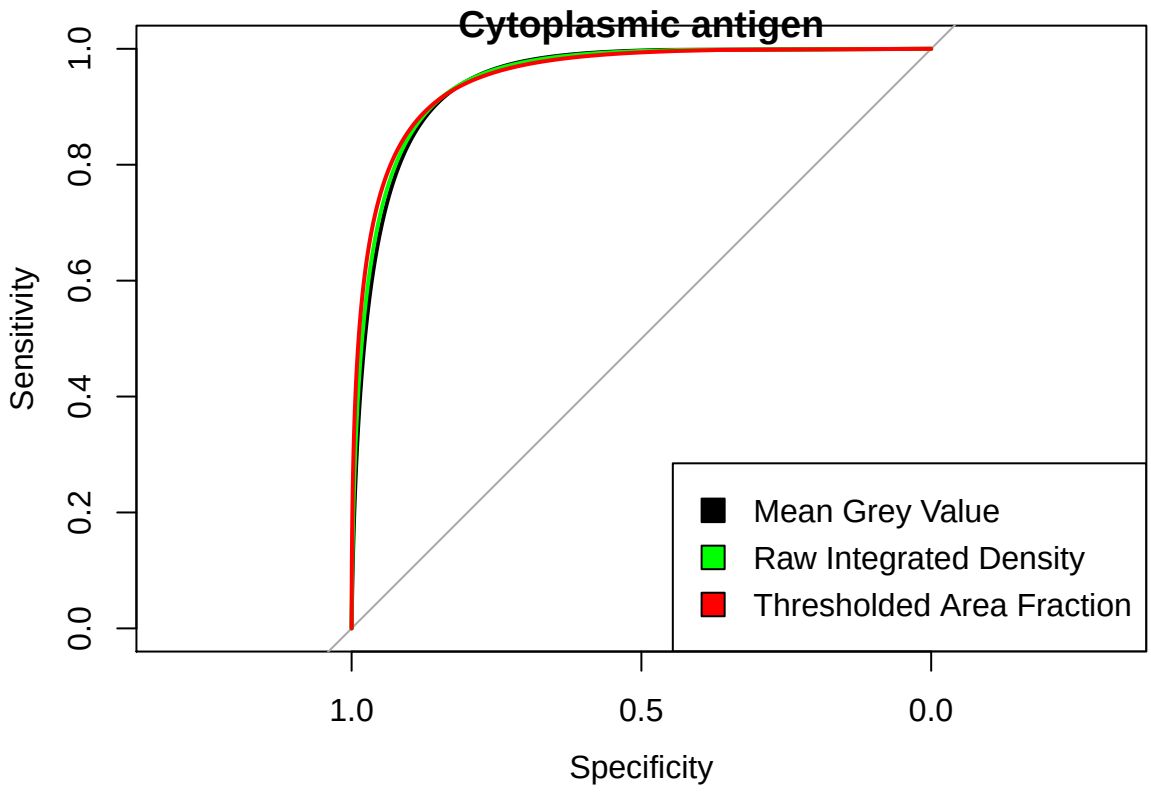
```
roc_Zscore_fish_jun2018 <- roc(logicFISH ~ Mean+RawIntDen+frArea+li_thr_frArea, data = Z_score_FISH_patt
roc_Zscore_ihc_jun2018 <- roc(logicIHC ~ Mean+RawIntDen+frArea+li_thr_frArea, data = Z_score_IHC_patter
roc_Zscore_brdu_jun2018 <- roc(logicBrdU ~ Mean+RawIntDen+frArea+li_thr_frArea, data = Z_score_BrdU_patt
```

3. Evaluation of the ROC curves First we plotted the three ROC curves (specificity vs sensitivity) for each channel type. The curves are smoothed by binormal smoothing implemented in the pROC library:

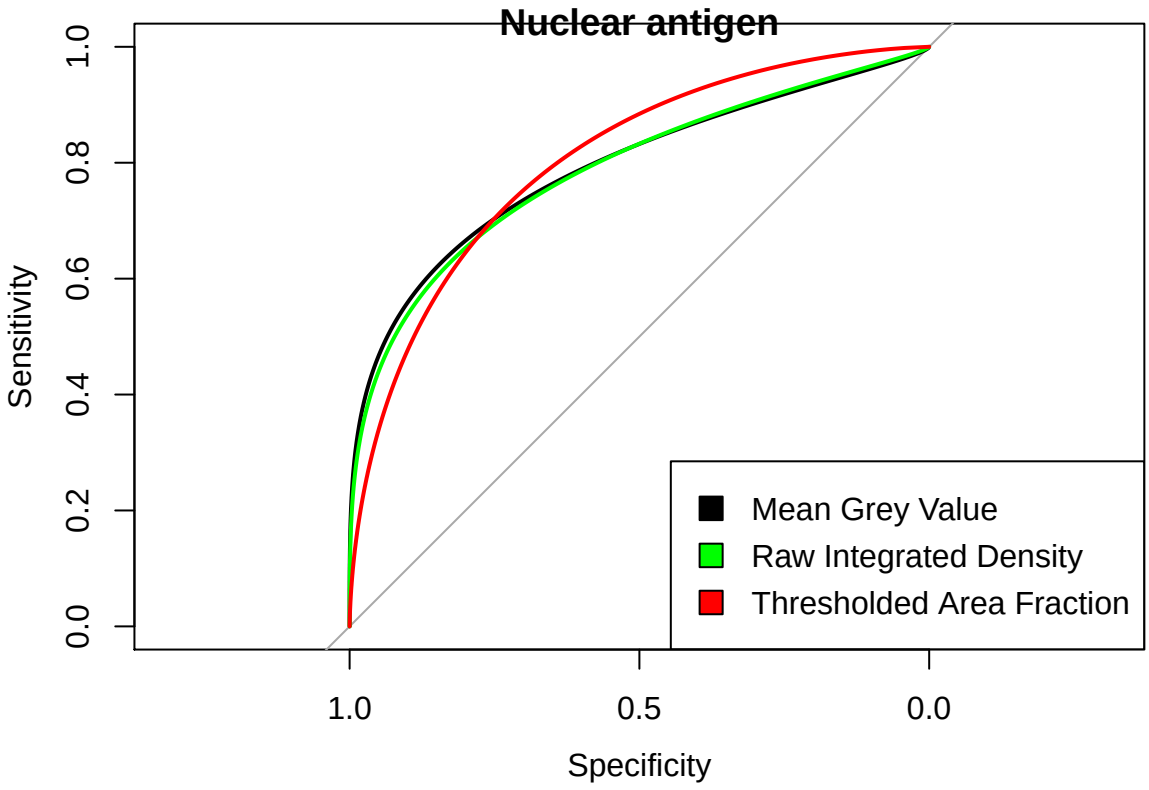
```
plot(smooth(roc_Zscore_fish_jun2018$Mean))
lines.roc(smooth(roc_Zscore_fish_jun2018$RawIntDen), col = "green")
lines.roc(smooth(roc_Zscore_fish_jun2018$li_thr_frArea), col = "red")
legend(legend = c("Mean Grey Value", "Raw Integrated Density", "Thresholded Area Fraction"),fill = c("bl
title(main = "mRNA FISH")
```




```
plot(smooth(roc_Zscore_ihc_jun2018$Mean))
lines.roc(smooth(roc_Zscore_ihc_jun2018$RawIntDen), col = "green")
lines.roc(smooth(roc_Zscore_ihc_jun2018$li_thr_frArea), col = "red")
legend(legend = c("Mean Grey Value", "Raw Integrated Density", "Thresholded Area Fraction"),fill = c("bl
title(main = "Cytoplasmic antigen")
```




```
plot(smooth(roc_Zscore_brdu_jun2018$Mean))
lines.roc(smooth(roc_Zscore_brdu_jun2018$RawIntDen), col = "green")
lines.roc(smooth(roc_Zscore_brdu_jun2018$li_thr_frArea), col = "red")
legend(legend = c("Mean Grey Value", "Raw Integrated Density", "Thresholded Area Fraction"),fill = c("bl
title(main = "Nuclear antigen")
```



The value of the selected features as a binary predictor was analysed separately for the different channel classes (mRNA FISH channels, Cytoplasmic antigen channels and the Nuclear antigen channels). Each of those features were found to be good enough predictors of the "ground truth" (Table 1). For simplicity of the implementation we chose the fraction of the thresholded area as a predictor. Next the area under the nonsmoothed curve for Mean grey value as main classification test is calculated and confidence intervals are constructed via bootstrapping with 10^5 resamples. The same bootstrapping strategy is used to find the threshold which maximizes the sum of Sensitivity and Specificity of the classificator

```
# mRNA FISH channel
auc(roc_Zscore_fish_jun2018$li_thr_frArea)
```

```
## Area under the curve: 0.816
```

```
ci(roc_Zscore_fish_jun2018$li_thr_frArea,boot.n = 100000)
```

```
## 95% CI: 0.7736-0.8583 (DeLong)
```

```
ci.thresholds(roc_Zscore_fish_jun2018$li_thr_frArea, thresholds = "best",boot.n = 100000 )
```

```
## 95% CI (1e+05 stratified bootstrap replicates):
##  thresholds sp.low sp.median sp.high se.low se.median se.high
##        5.62 0.5301    0.6011  0.6721 0.8254     0.873  0.9206
```

```
# Cellular antigen channel
auc(roc_Zscore_ihc_jun2018$li_thr_frArea)
```

```
## Area under the curve: 0.9454
```

```
ci(roc_Zscore_ihc_jun2018$li_thr_frArea,boot.n = 100000)
```

```
## 95% CI: 0.9182-0.9725 (DeLong)
```

```
ci.thresholds(roc_Zscore_ihc_jun2018$li_thr_frArea, thresholds = "best",boot.n = 100000 )
```

```
## 95% CI (1e+05 stratified bootstrap replicates):
##  thresholds sp.low sp.median sp.high se.low se.median se.high
##         3.1 0.8037    0.8589   0.908 0.8824    0.9314  0.9804
```

```
# Nuclear antigen channel
auc(roc_Zscore_brdu_jun2018$li_thr_frArea)
```

```
## Area under the curve: 0.7372
```

```
ci(roc_Zscore_brdu_jun2018$li_thr_frArea,boot.n = 100000)
```

```
## 95% CI: 0.3302-1 (DeLong)
```

```
ci.thresholds(roc_Zscore_brdu_jun2018$li_thr_frArea, thresholds = "best",boot.n = 100000 )
```

```
## 95% CI (1e+05 stratified bootstrap replicates):
##  thresholds sp.low sp.median sp.high se.low se.median se.high
##      11.015 0.8372    0.8837  0.9256      0    0.6667       1
```

The "nuclear antigen channel"contains images marked for either BrdU or Ki67. The number of positive cells is extremely low and whence the wide confidence intervals for the area under the curve for those images. Due to the small number of positives in the training data for the nuclear antigen channel the ROC curve is not interpretable. Which is why we chose a reasonably high threshold for positivity at 30% coverage of the masks.