

# BREAST CANCER CLASSIFICATION

based on histopathological images

by:  
Robin Demesmaeker  
MA3, EE

# CONTENT

- Introduction
- Acquisition
- Exploration
- Exploitation and Evaluation
- Conclusion

# INTRODUCTION



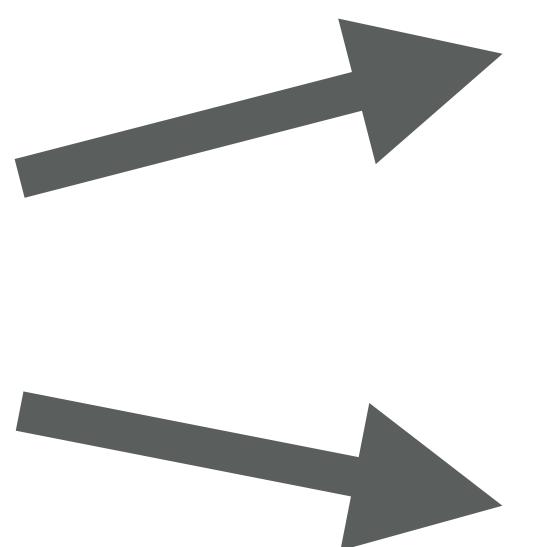
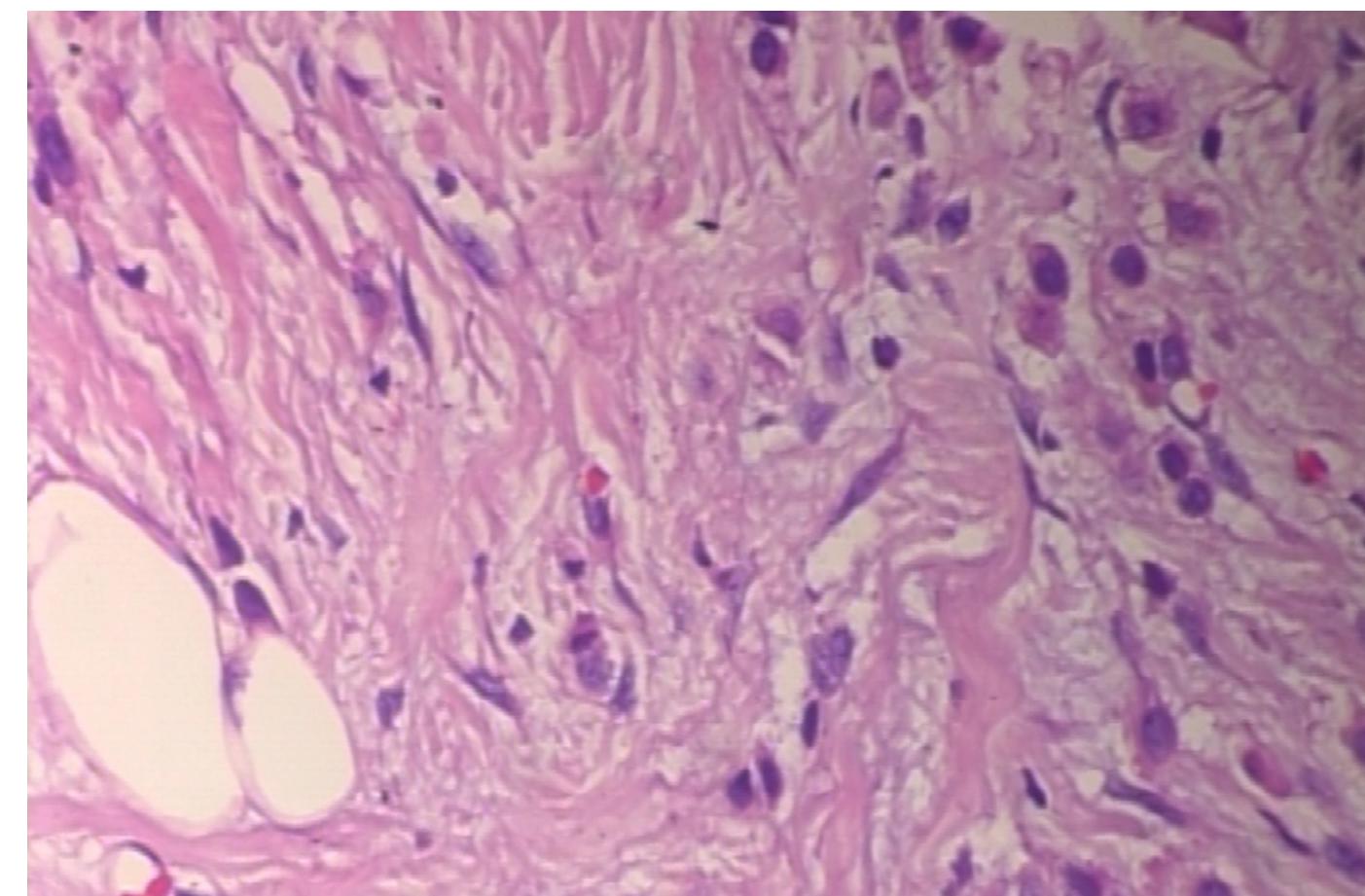
Example: disease classification

# ACQUISITION (I)

BreakHis dataset (2016) of histopathological images

Breast tumors: benign and malignant

460x700 images at 40x, 100x, 200x and 400x magnification  $\sim 1900/\text{magnification}$



Benign or malignant?

Which subtype? 4 benign, 4 malignant

# ACQUISITION (2)

Download: png images + metadata files

Preprocessing: resize images

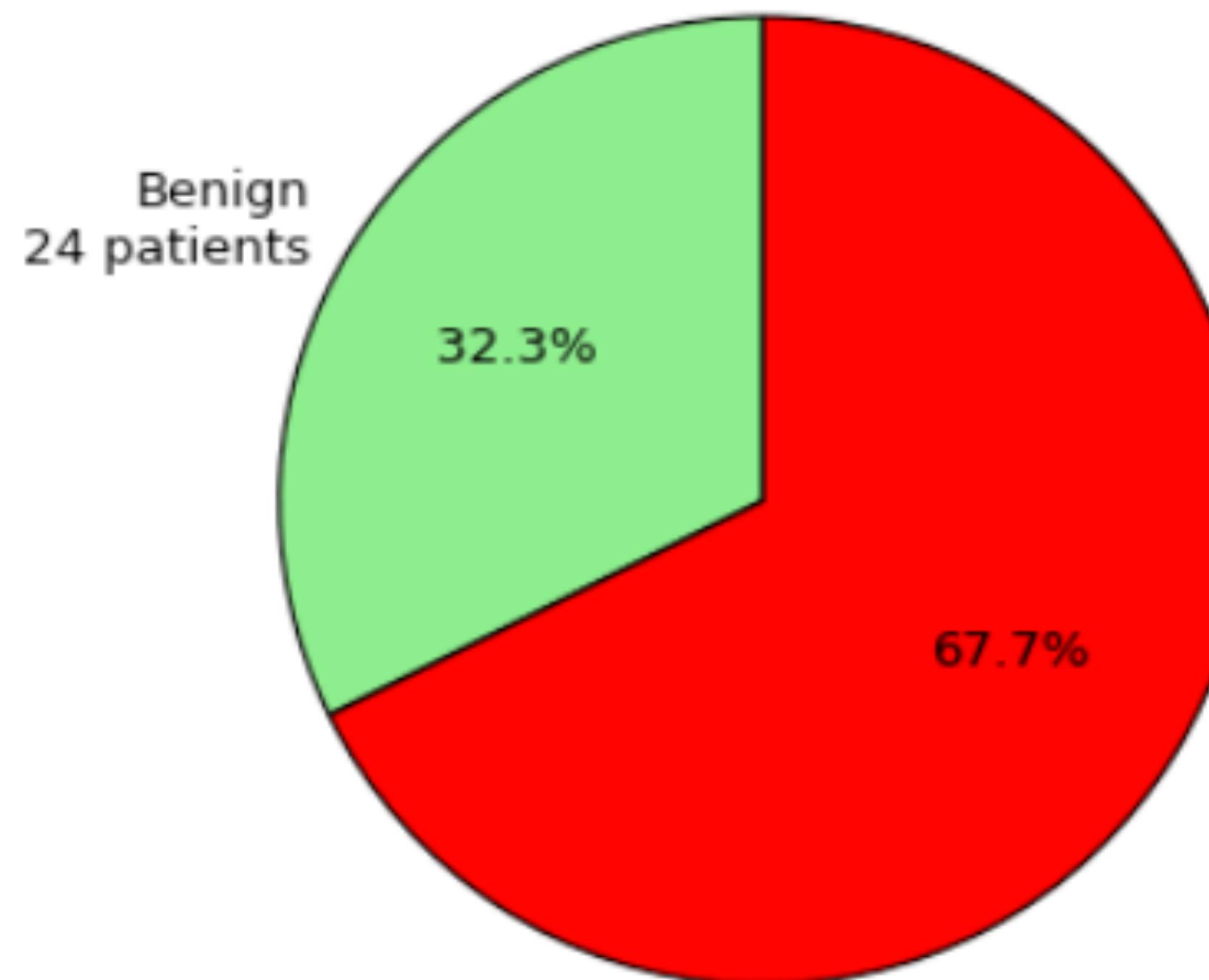
Preparing for use: make sqlite databases

# EXPLORATION

- Content of the dataset
- A sense for the data
- Simple features

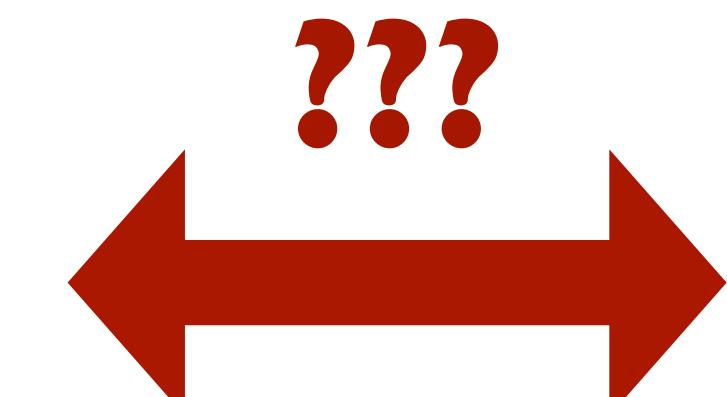
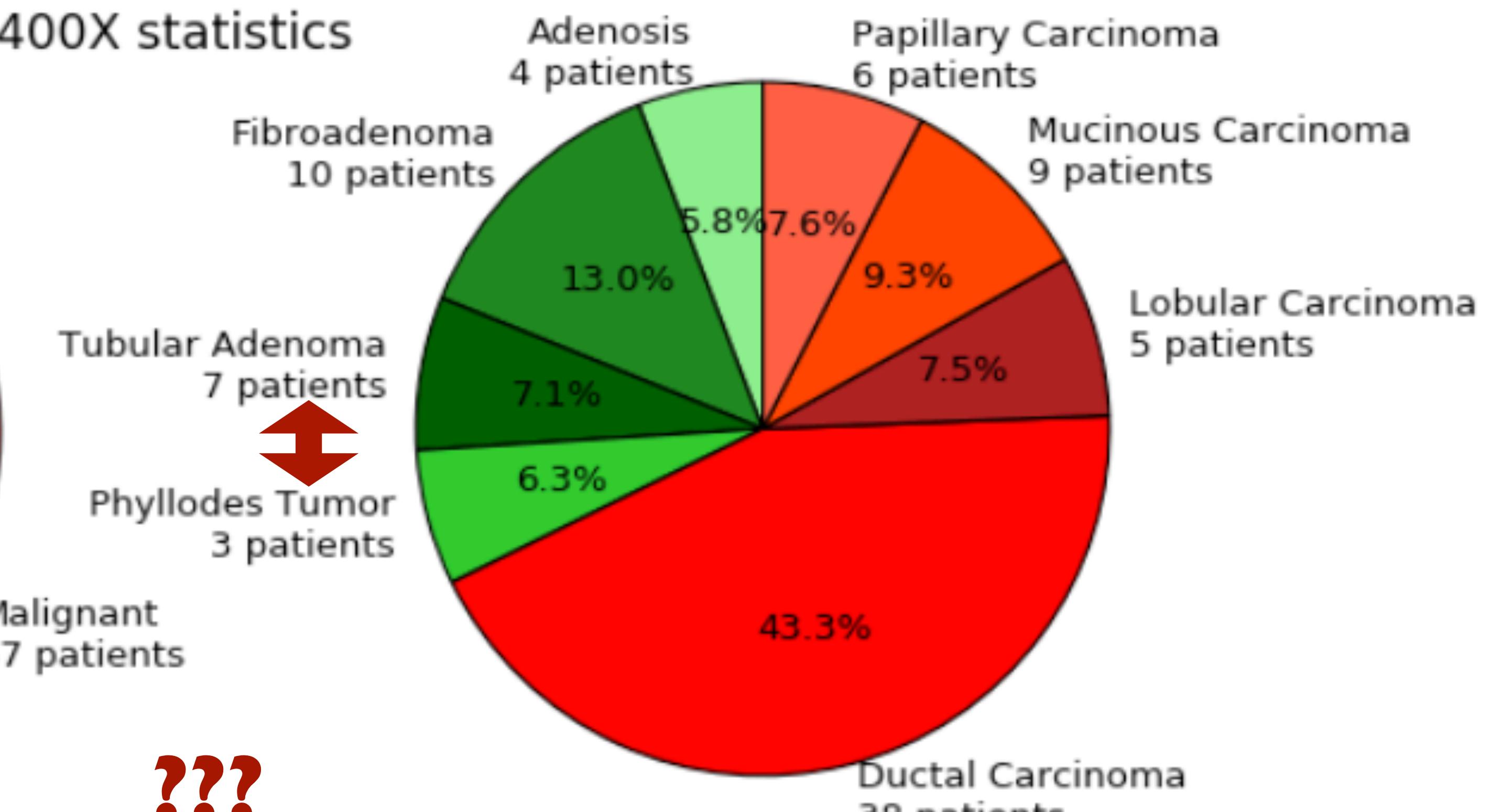


# CONTENT OF THE DATASET



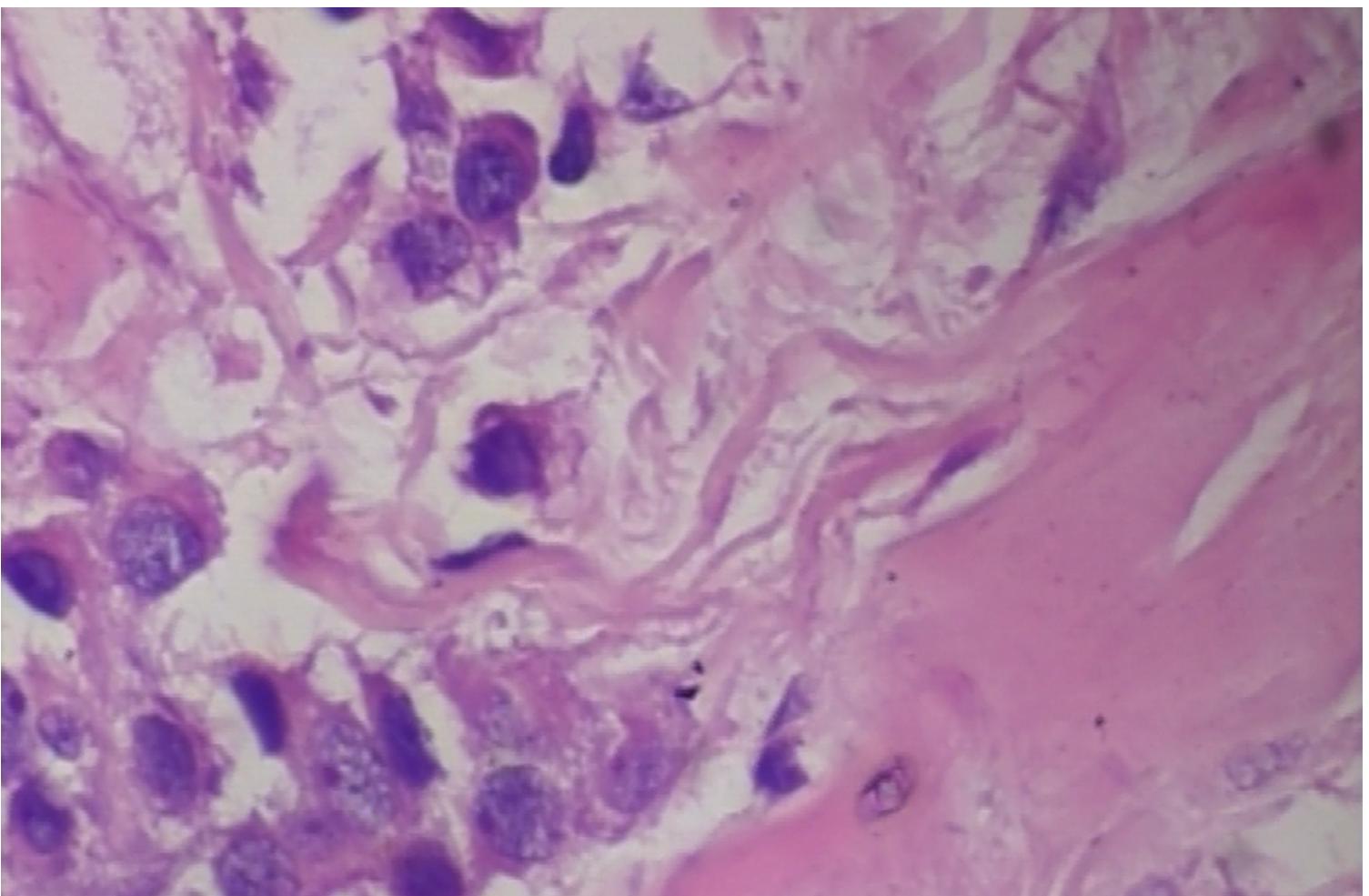
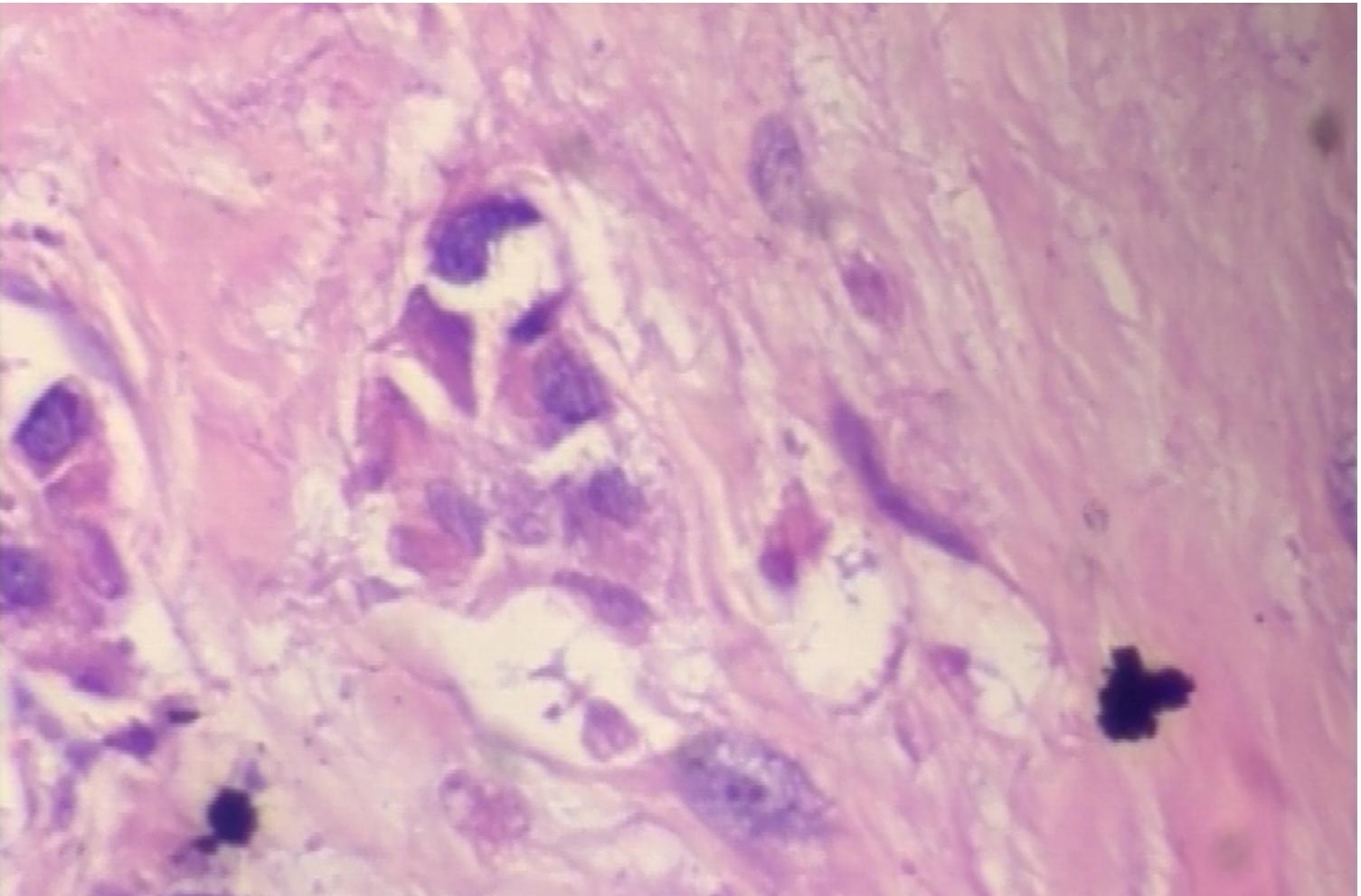
81 patients

400X statistics

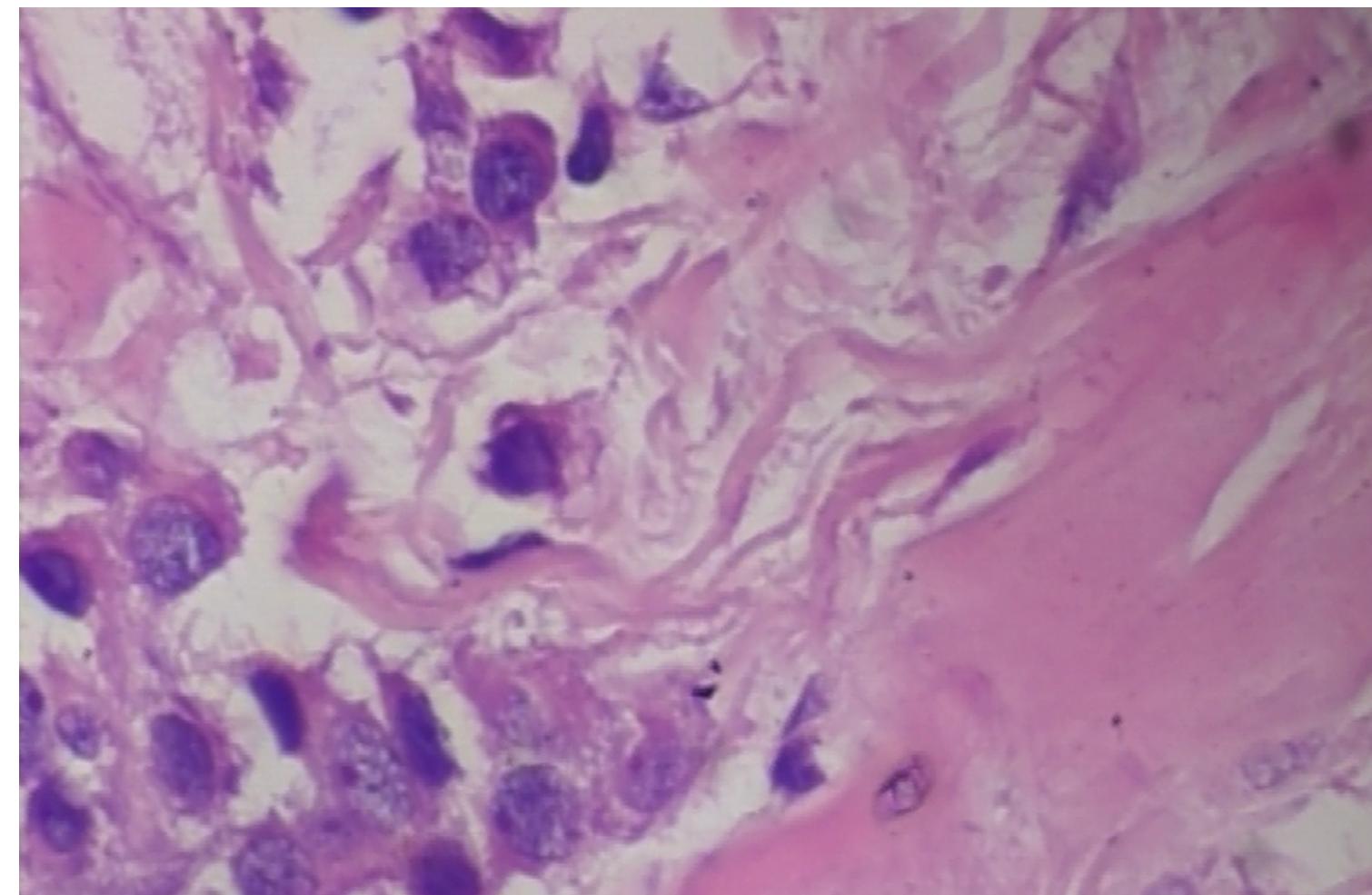
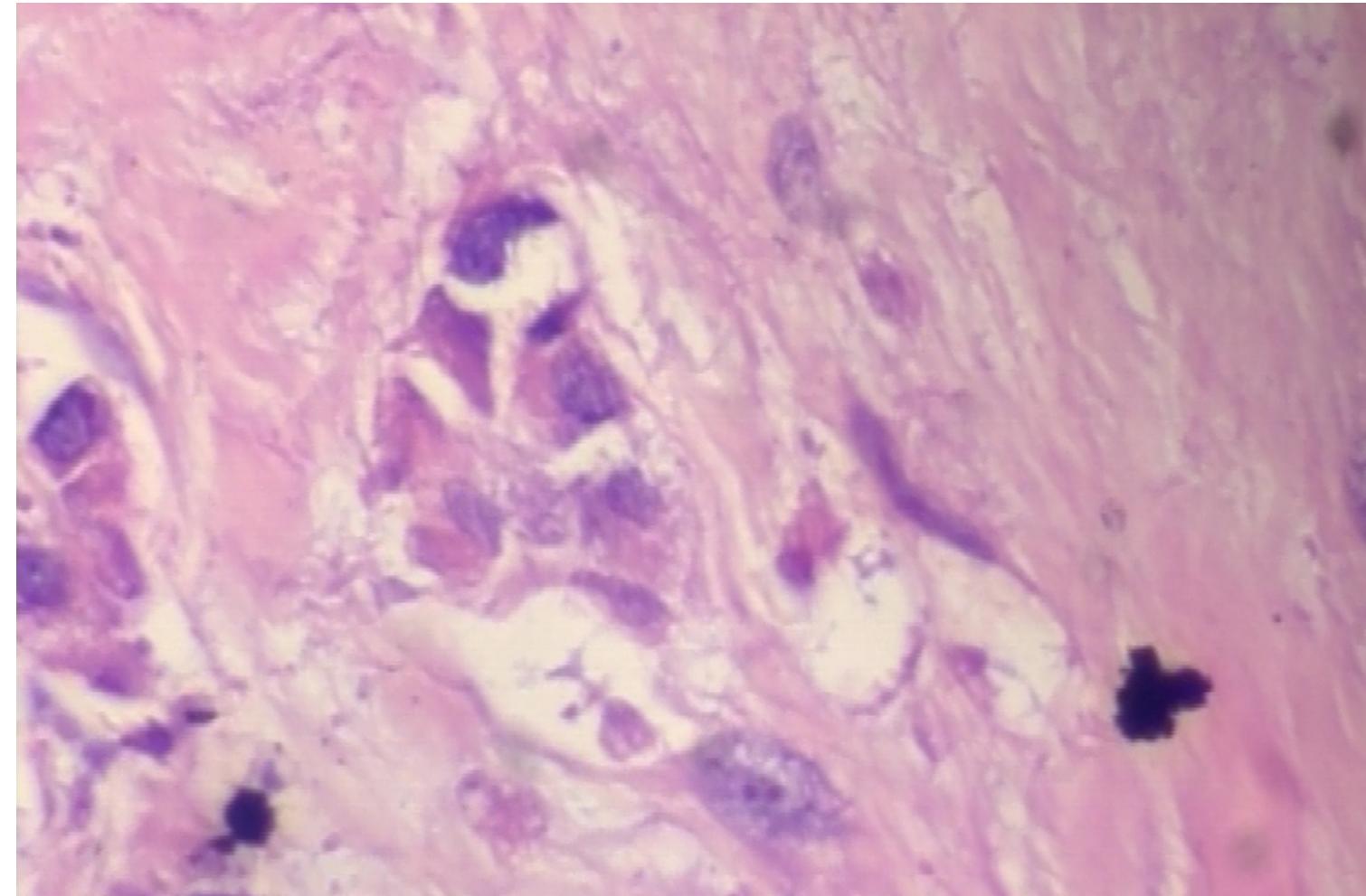


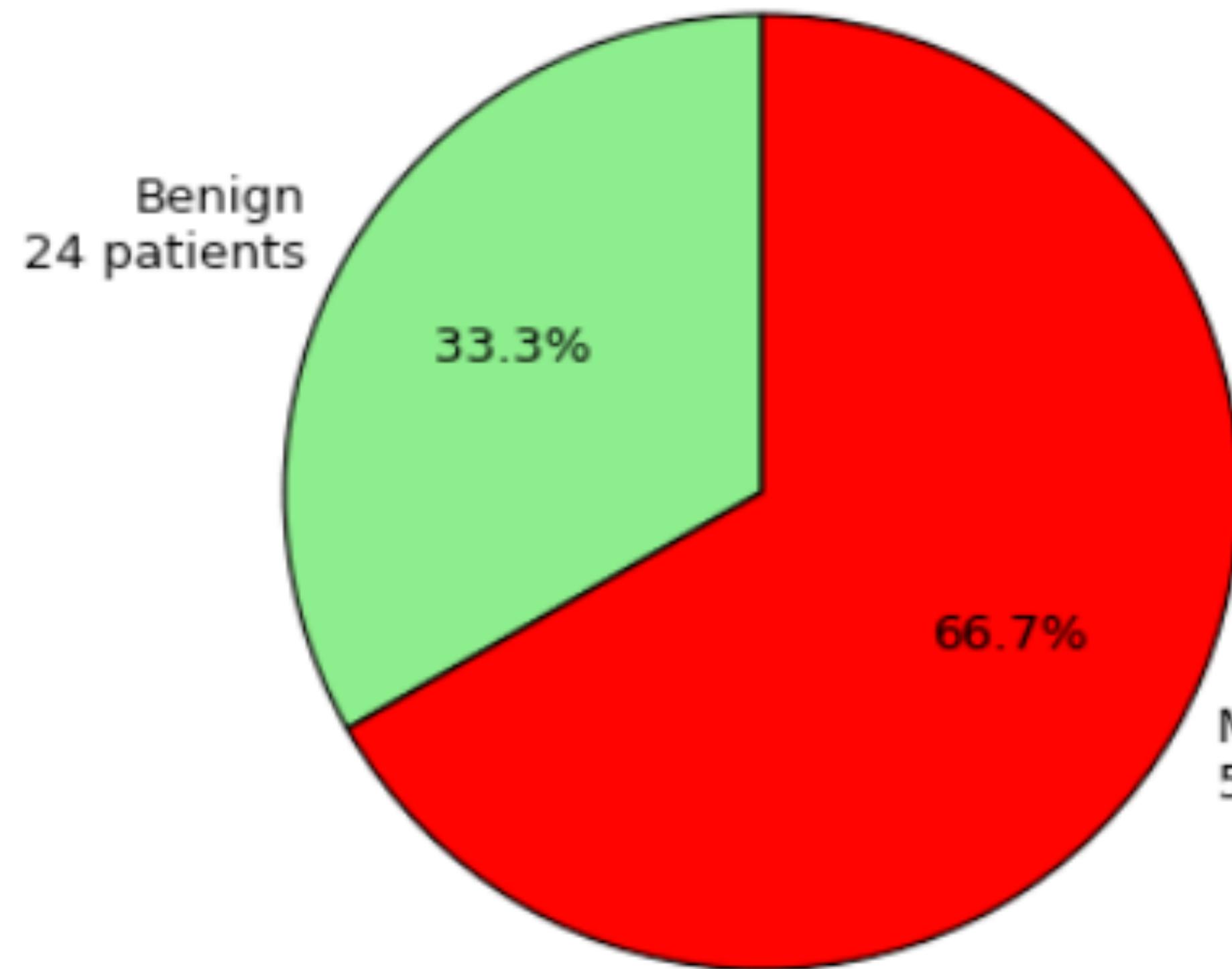
82 patients

patient I3412 **DC**

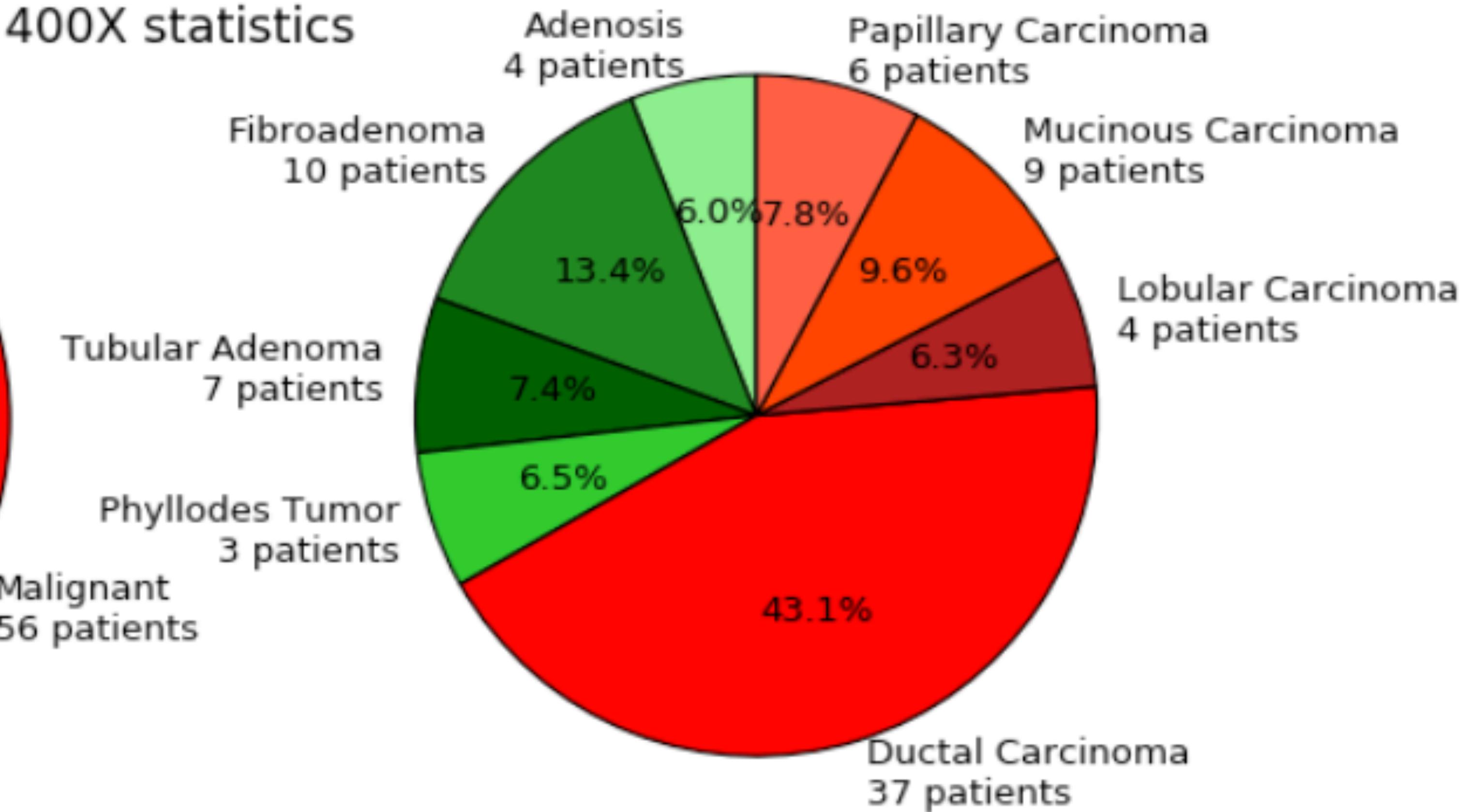


patient I3412 **LC**





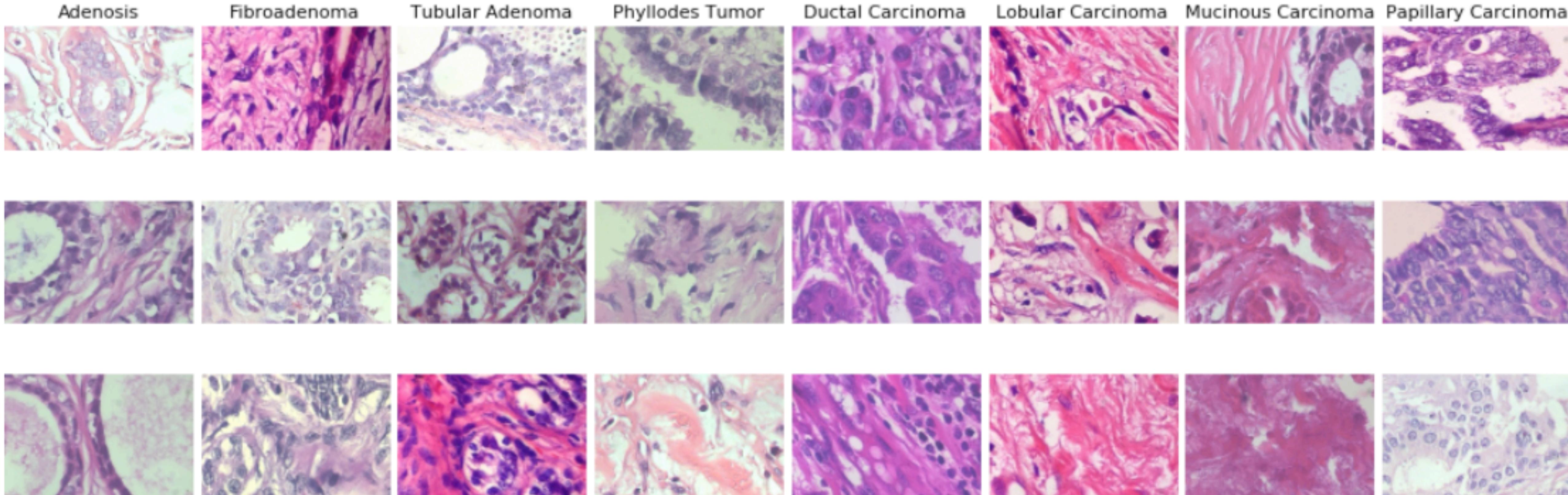
## 400X statistics



Highly biased data!

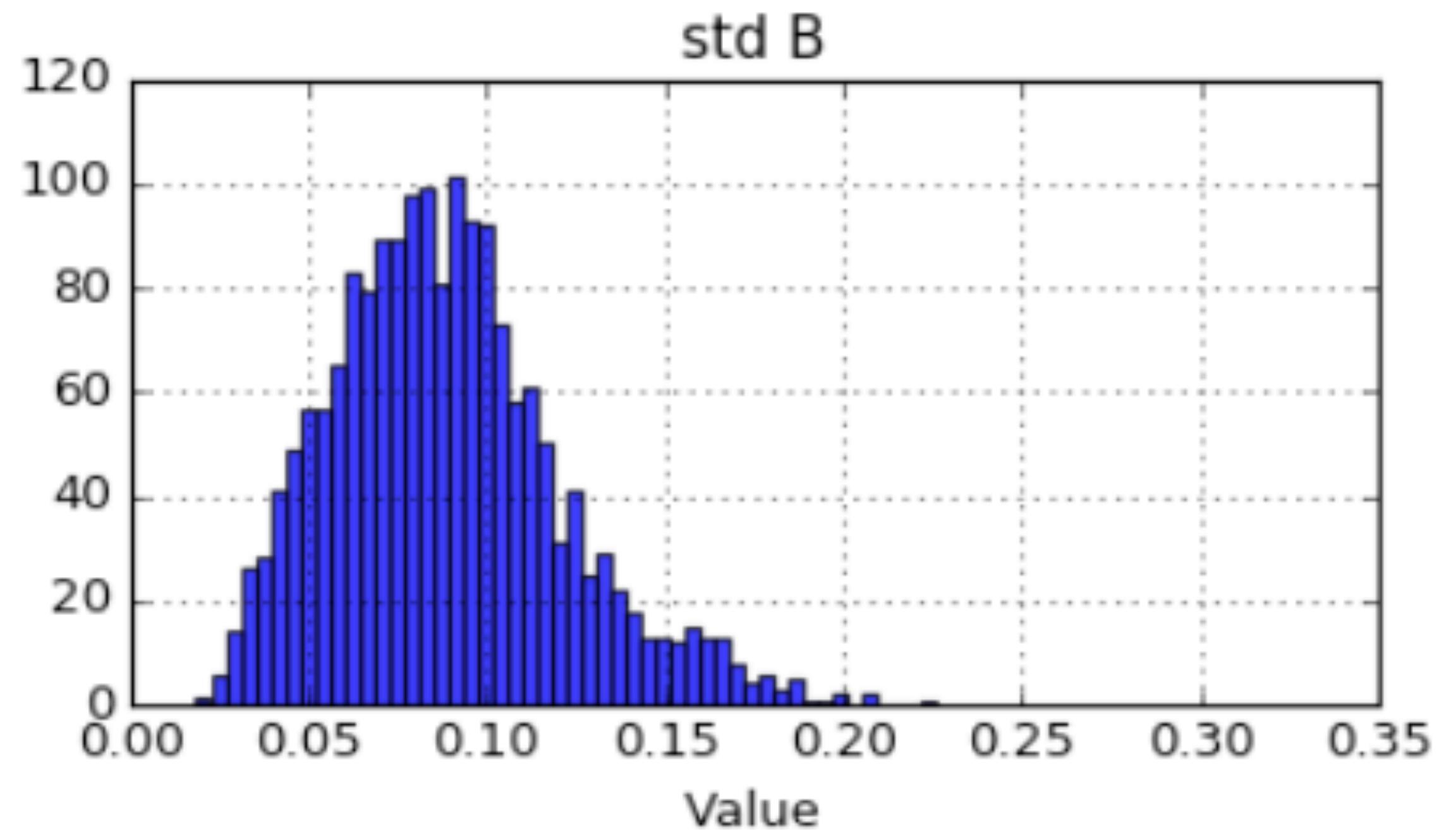
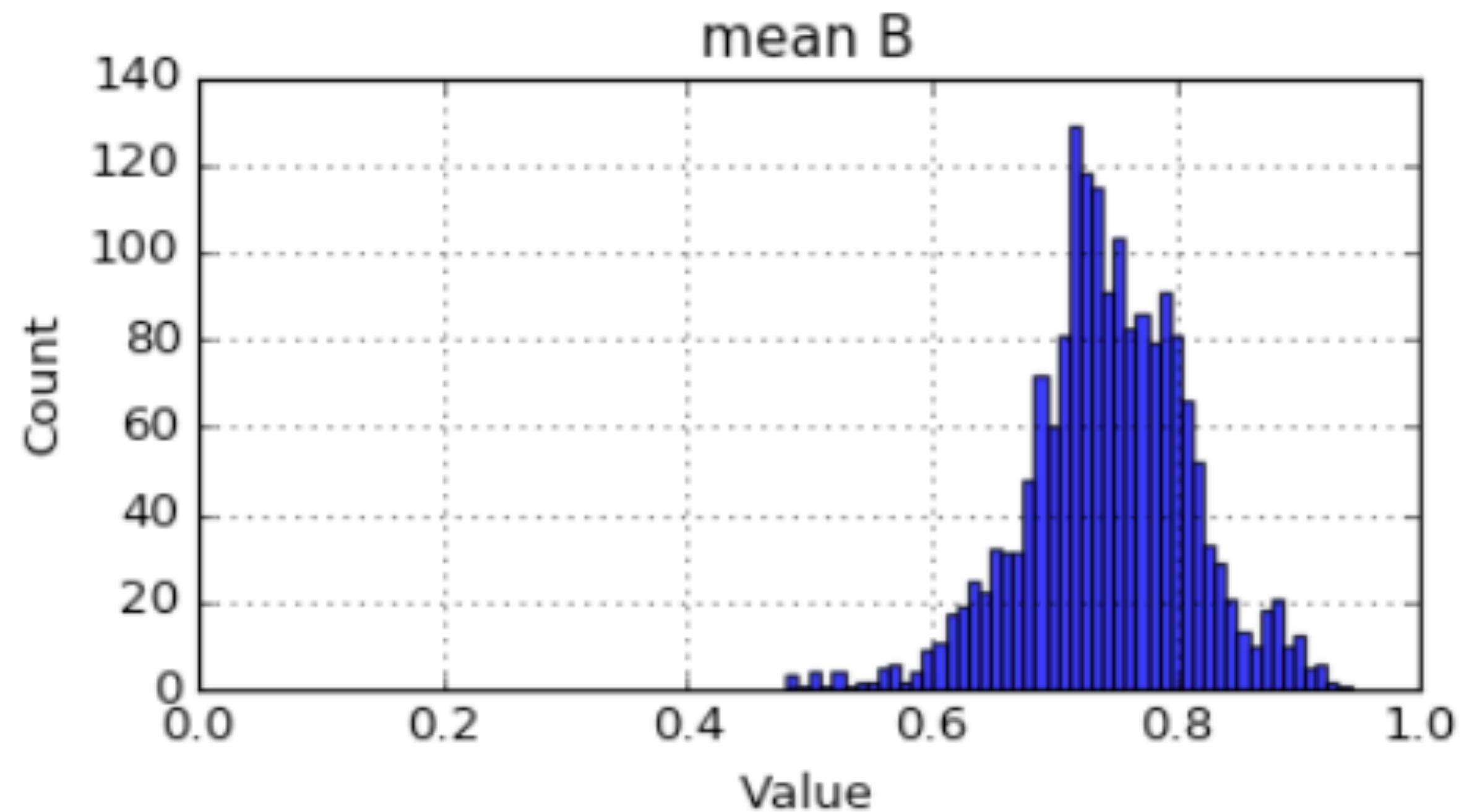
# A SENSE FOR THE DATA

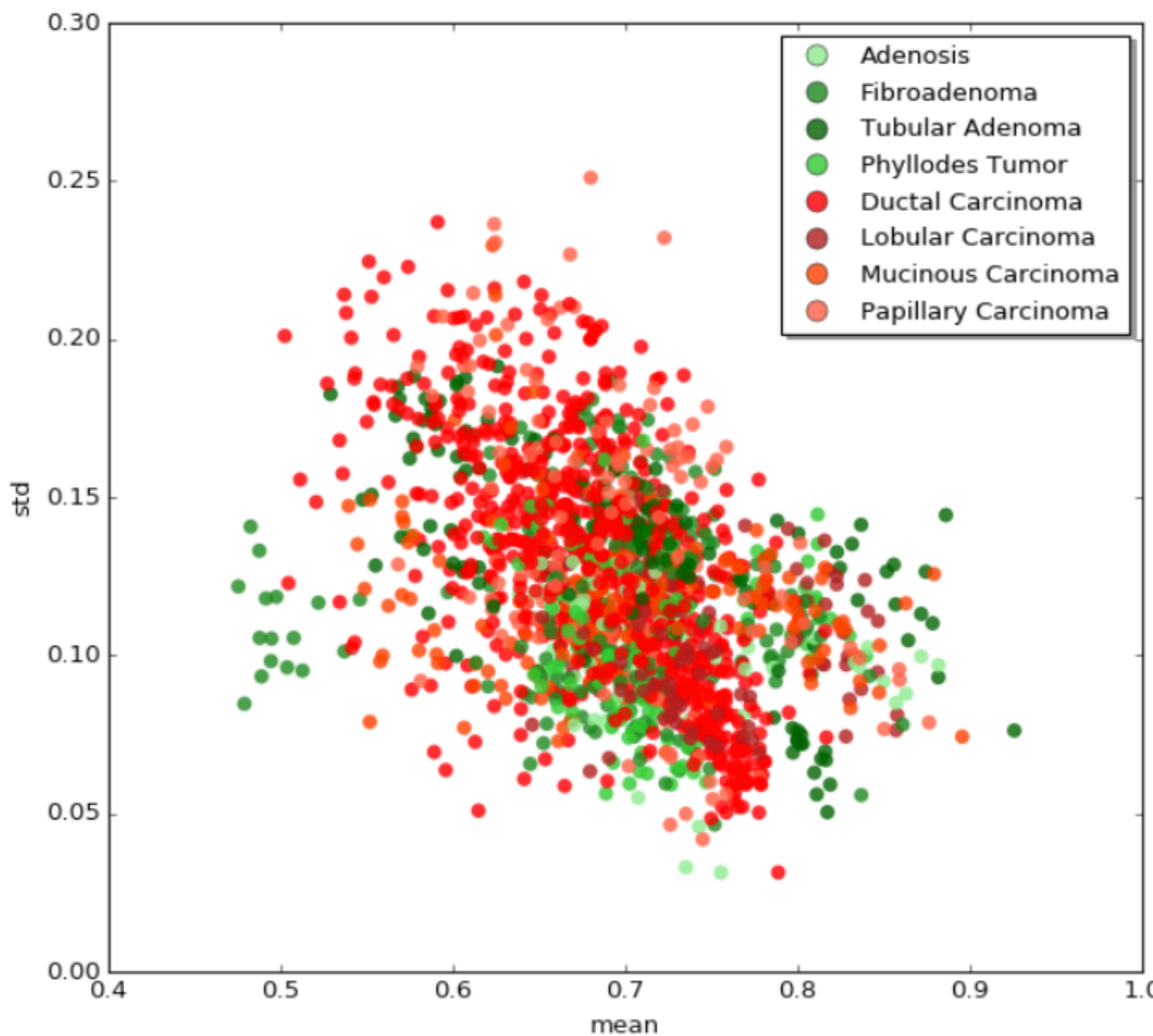
400X



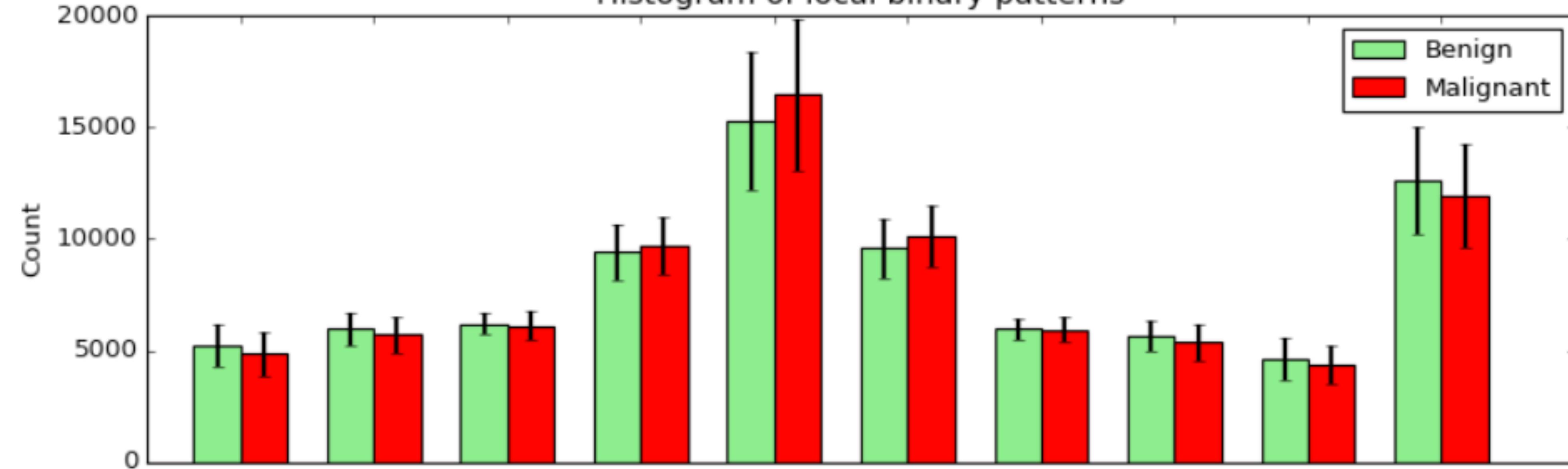
Without biological expertise: good luck!

# SIMPLE FEATURES

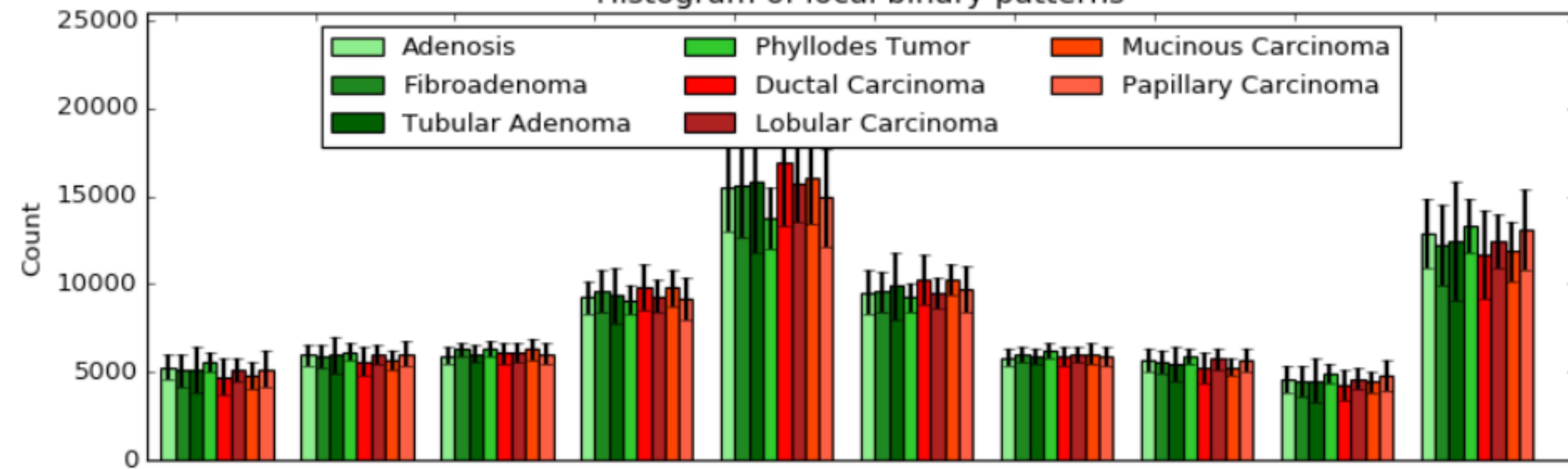




### Histogram of local binary patterns



### Histogram of local binary patterns



# EXPLOITATION AND EVALUATION

- Convolutional neural networks: why?
- Patch-based classification
- Models

# CONVOLUTIONAL NEURAL NETWORKS: WHY?

Conventional solutions:

kNN, SVM, random  
forest, ...

- Need explicit feature  
calculation

Neural networks:

simple, recurrent,  
**convolutional**, ...

+ Learn implicit features  
representative for  
problem

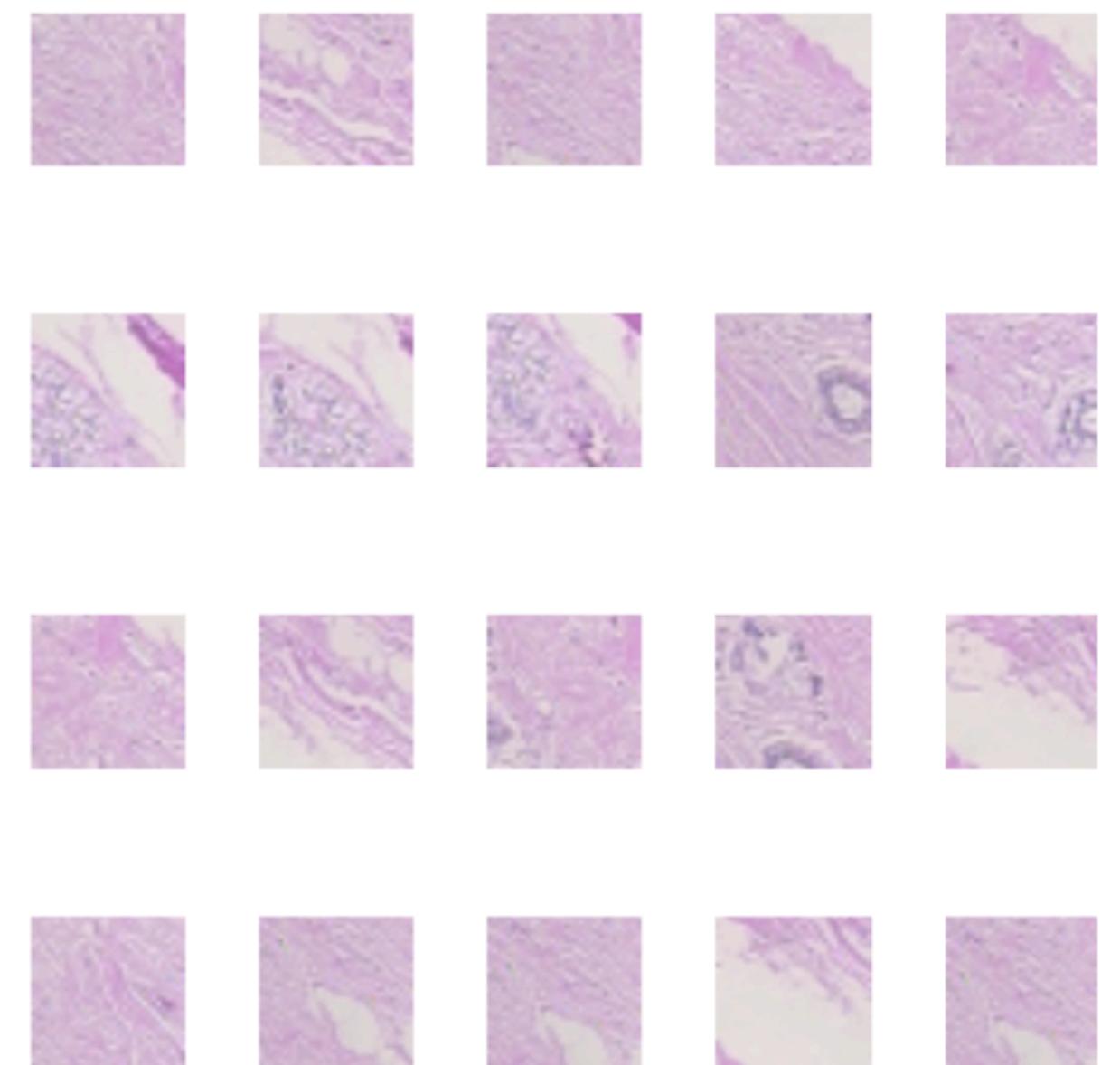
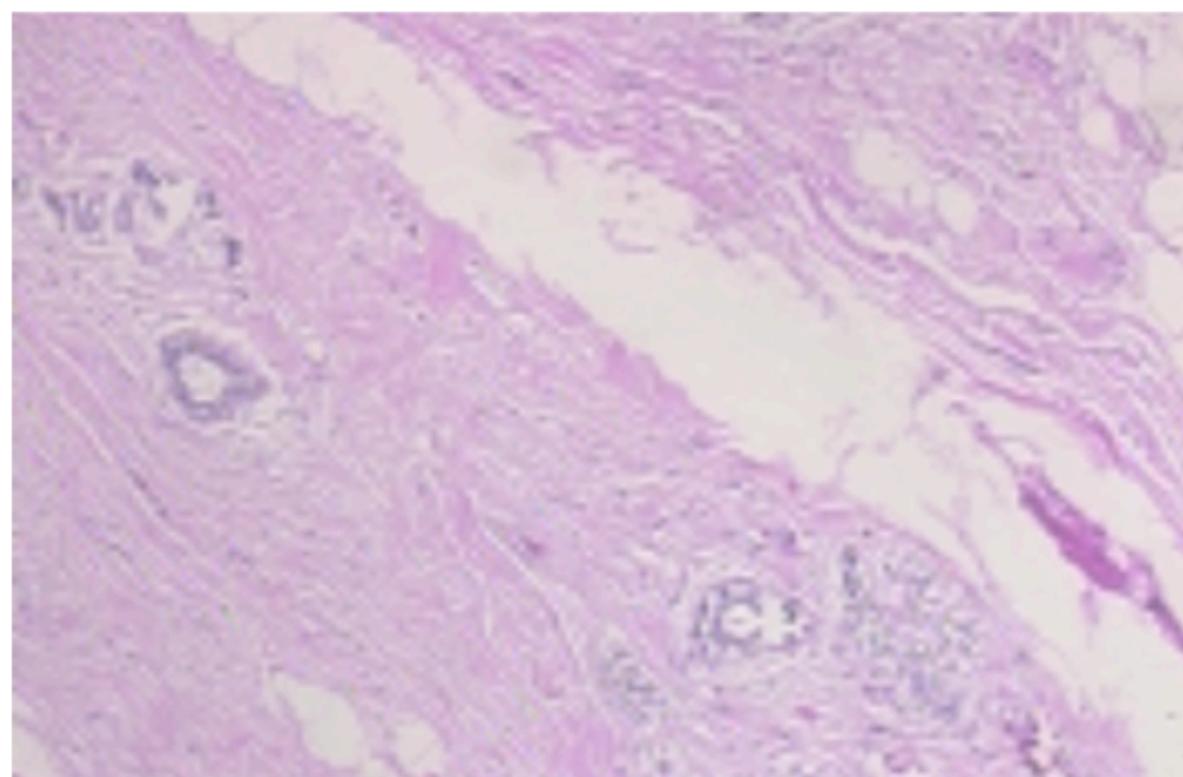
# PATCH-BASED CLASSIFICATION

- 32x32x3 patches
- Combining results:

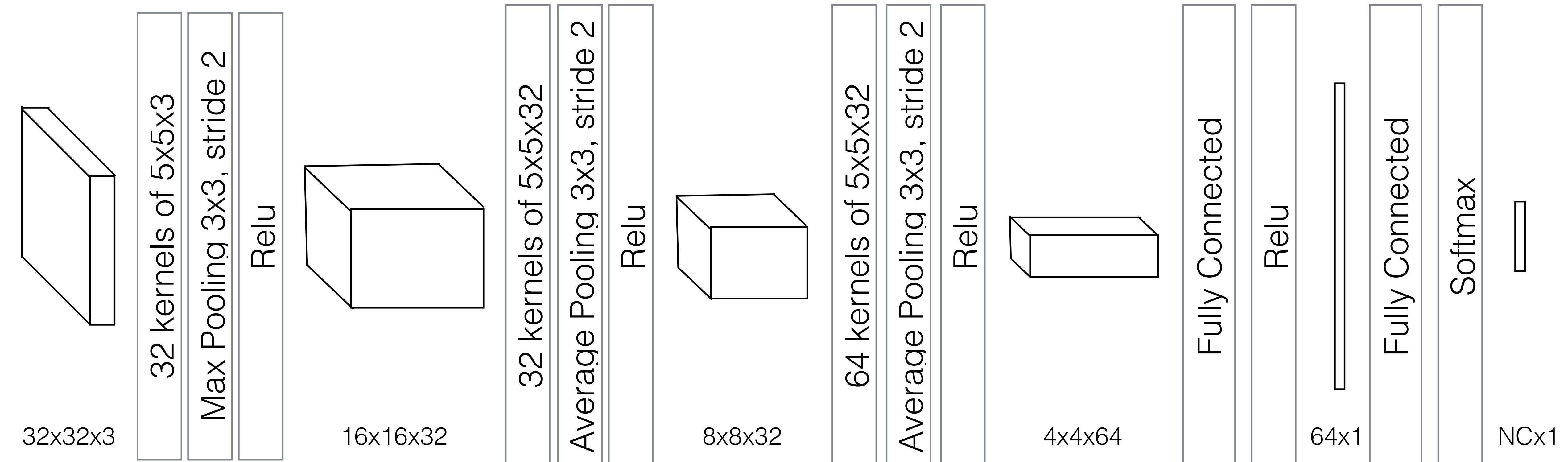
patch prediction =  $\text{argmax}(y_{patch})$

image prediction =  $\text{argmax} \left( \sum_{i=1}^{N_{patch}} y_{patch,i} \right)$

patient prediction =  $\text{argmax} \left( \sum_{i=1}^{N_{images}} \sum_{j=1}^{N_{patch}} y_{patch,i,j} \right)$



# MODELS



CONVOLUTIONAL

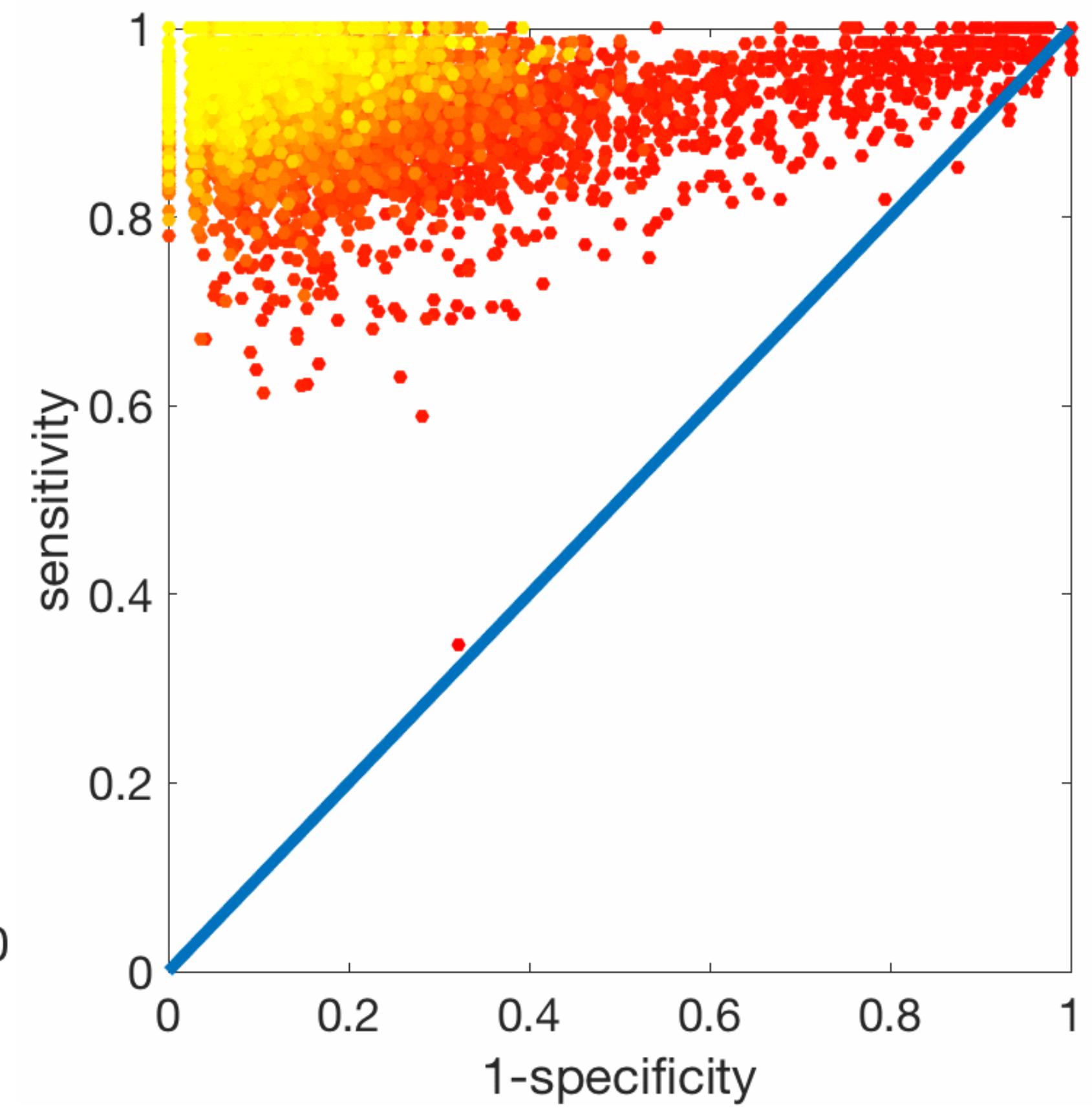
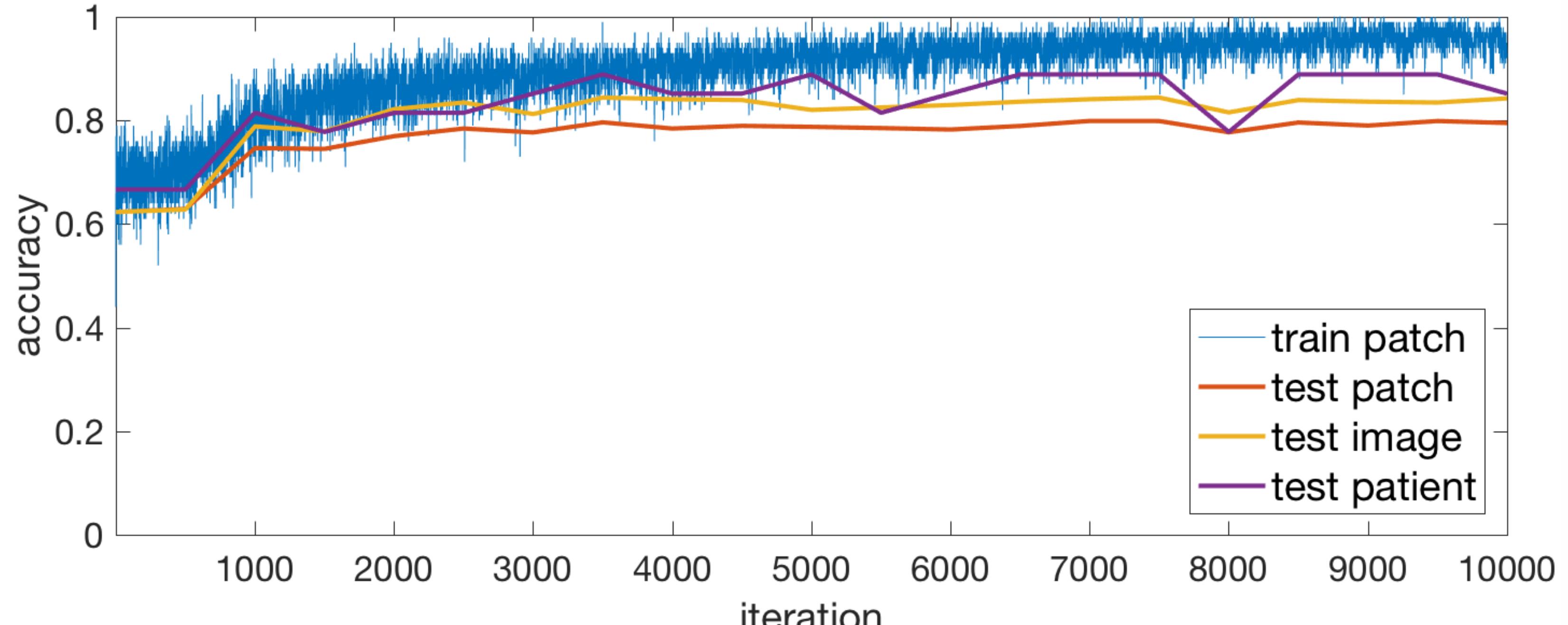
CONVOLUTIONAL

CONVOLUTIONAL

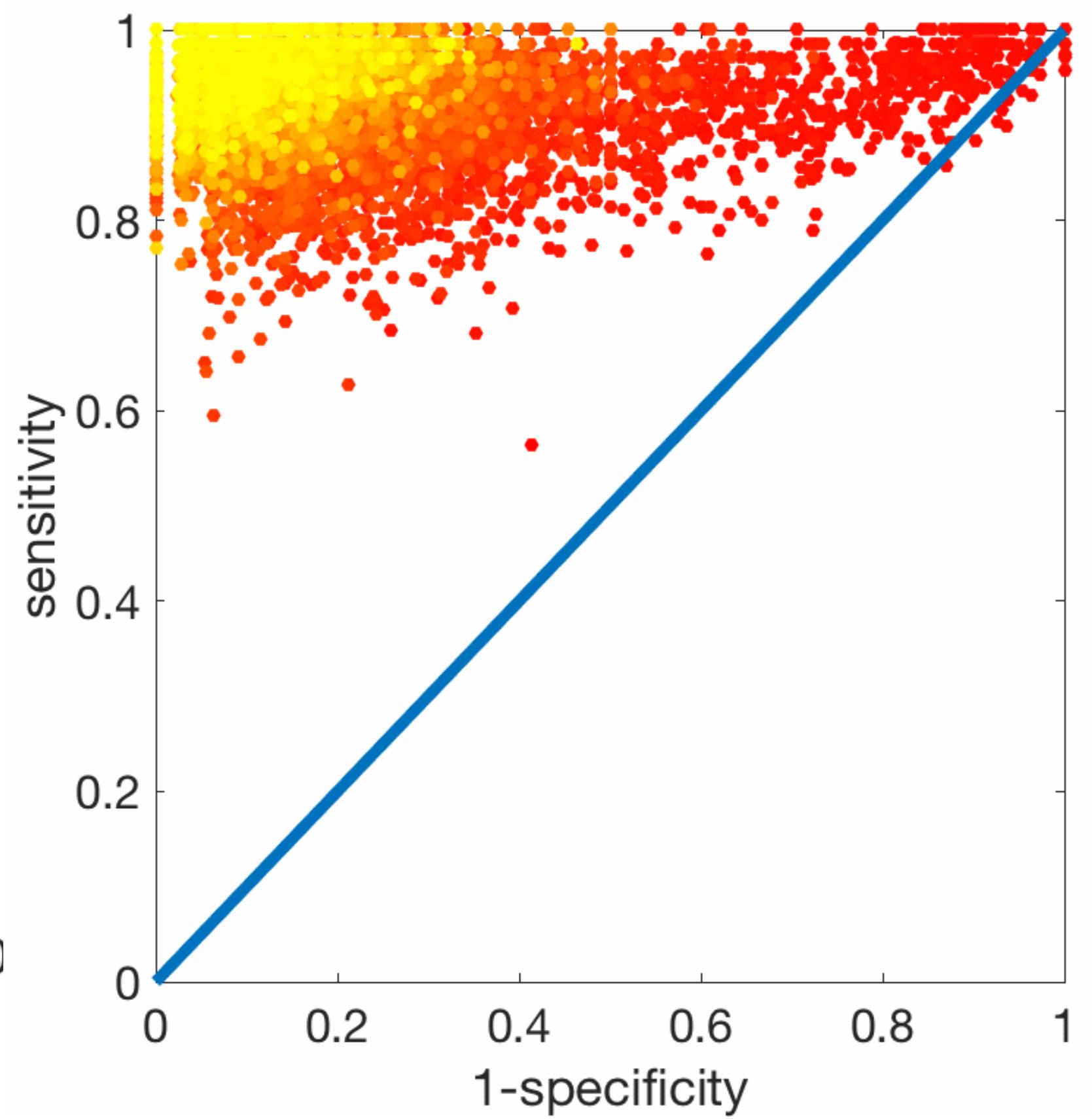
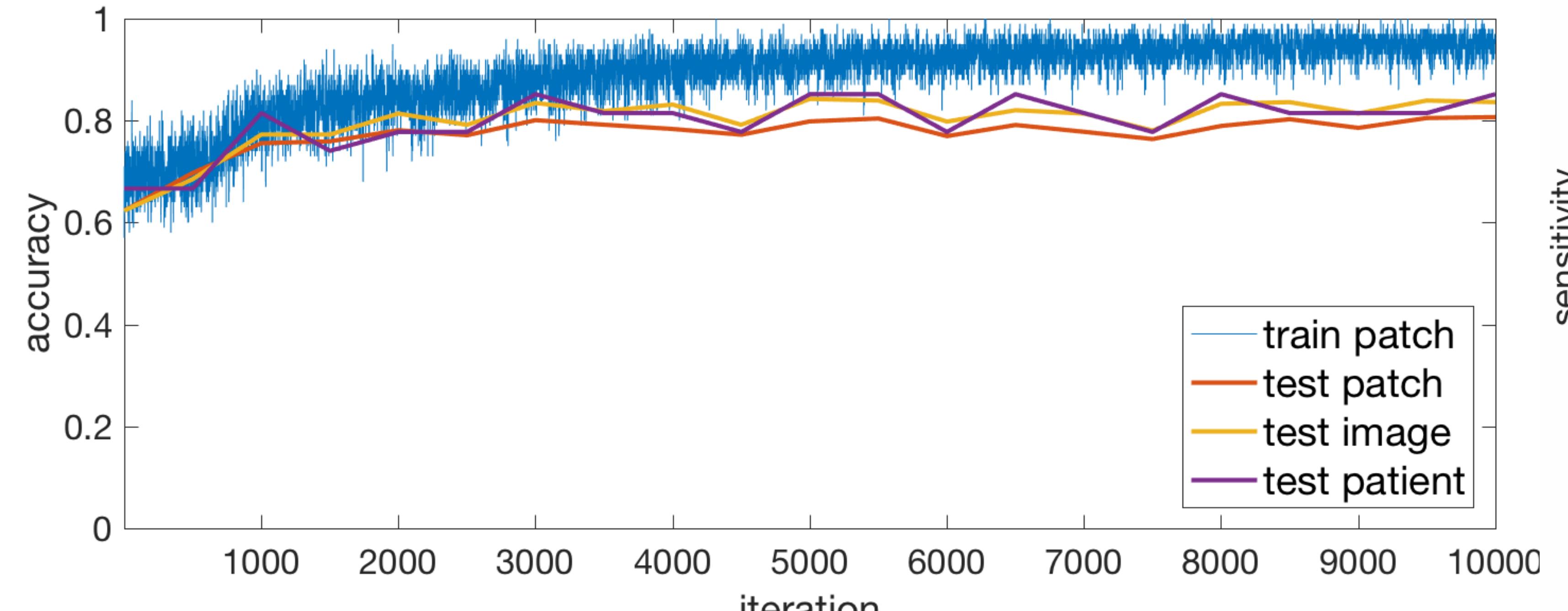
FC

FC

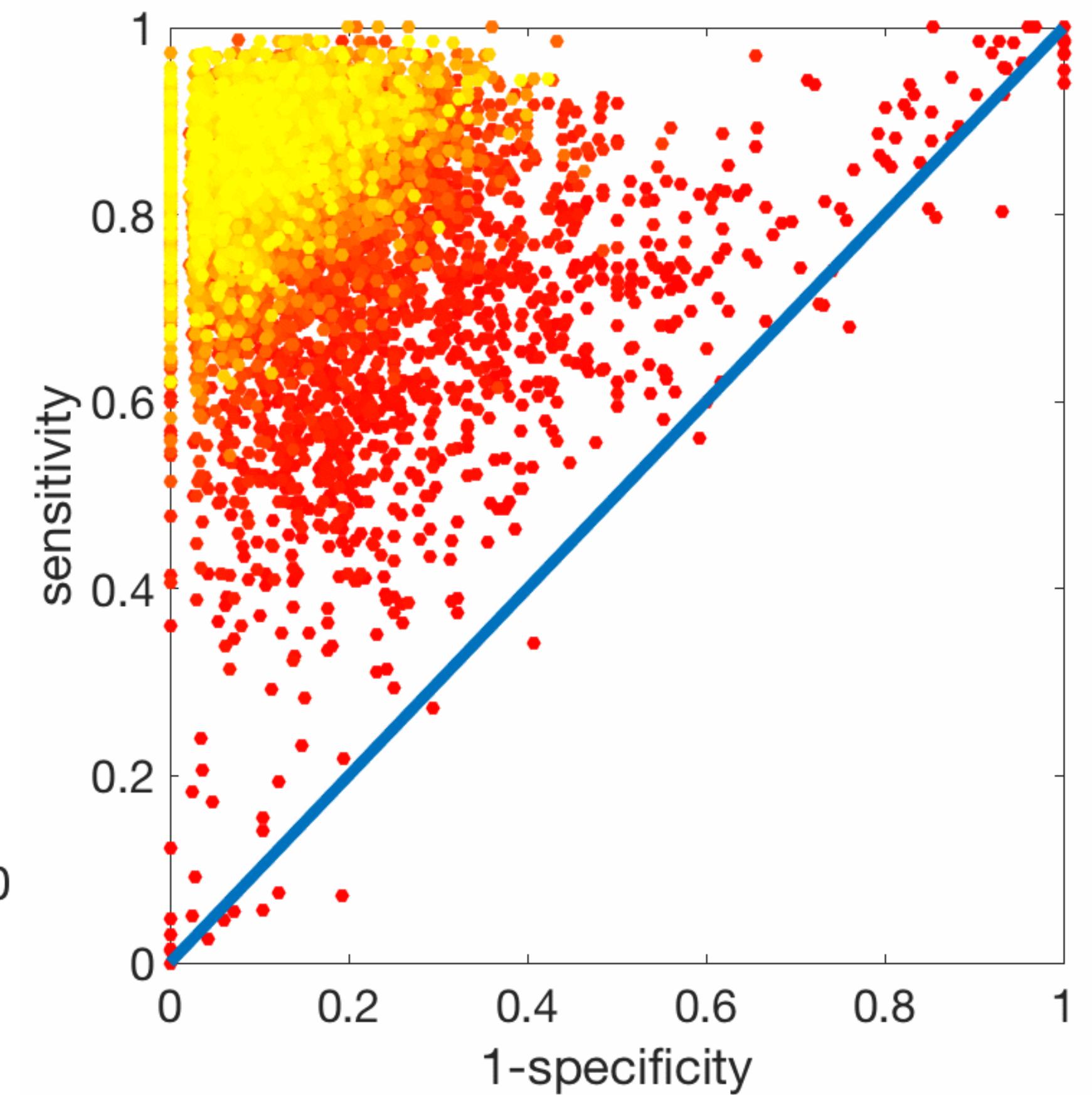
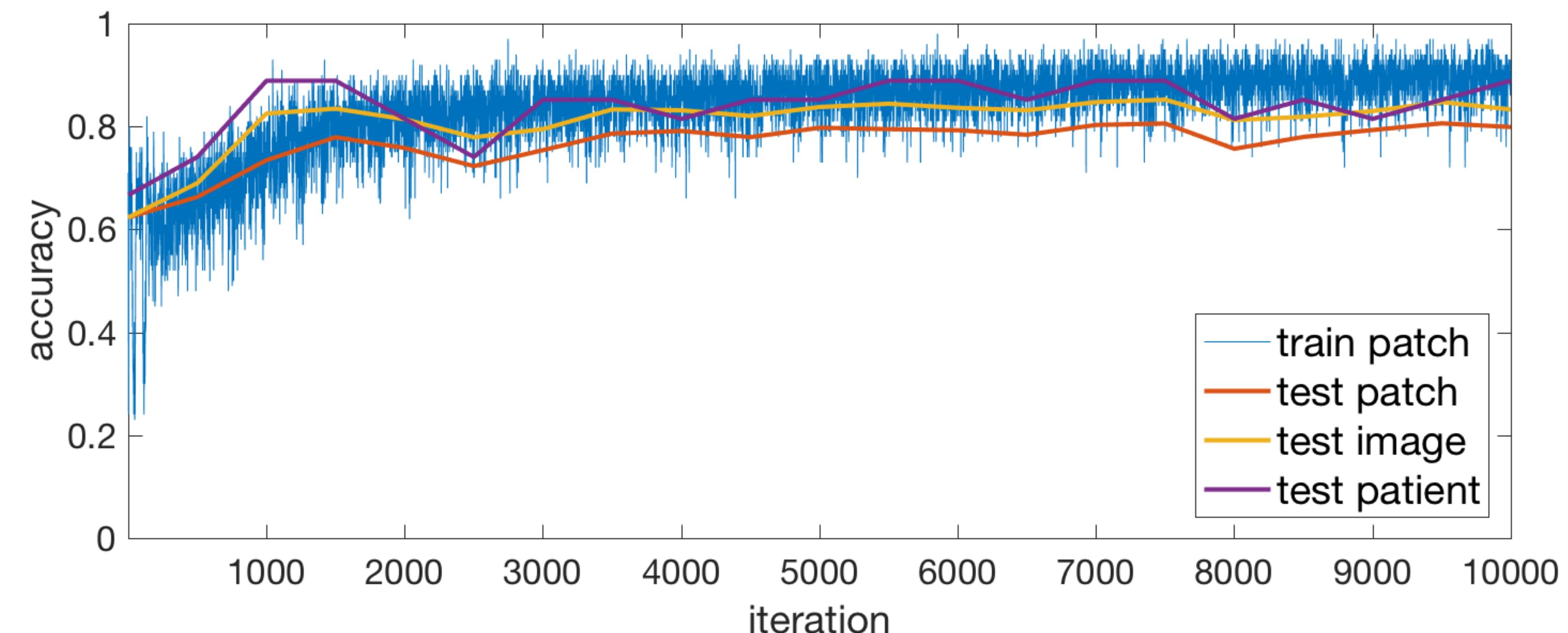
# FIRST RESULTS



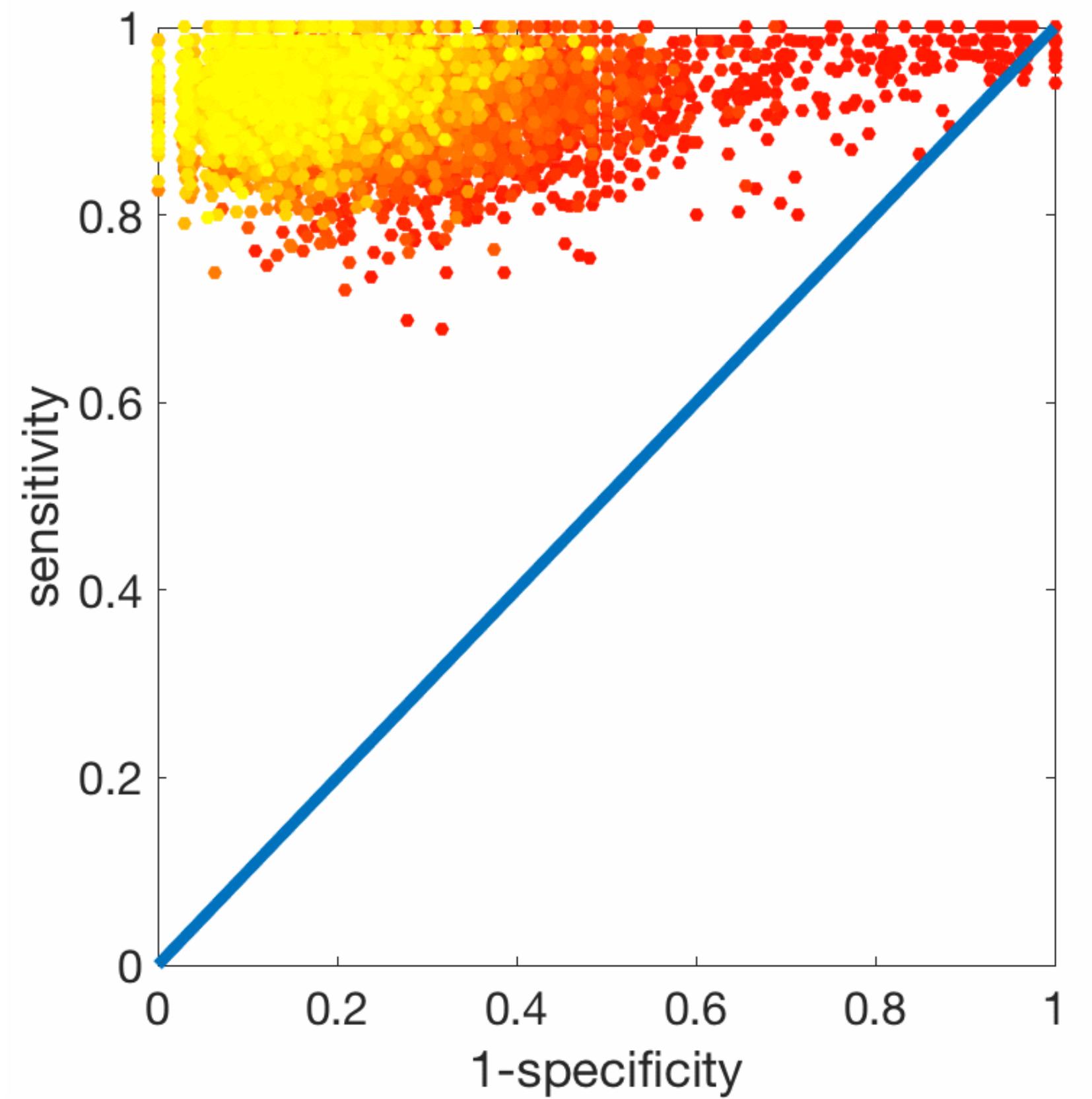
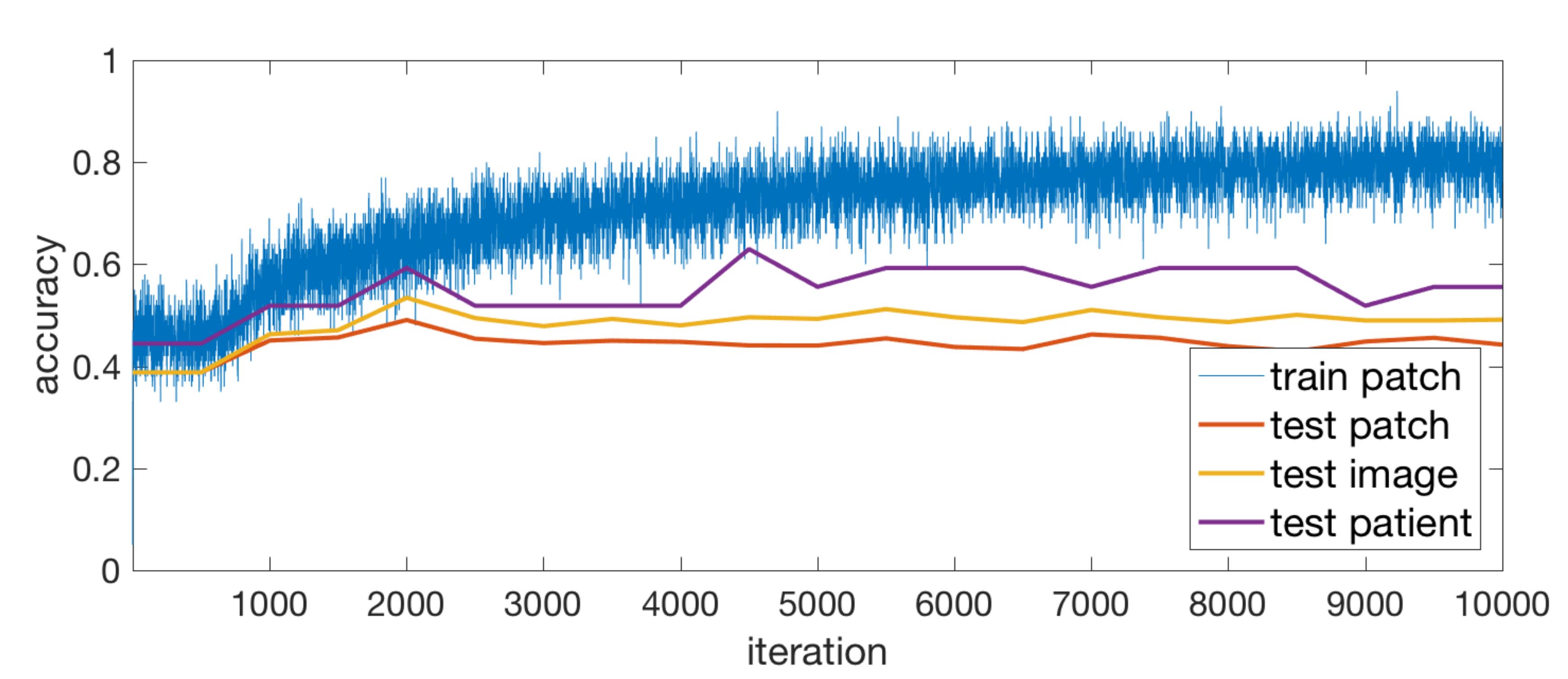
# SUBSAMPLING IN CONVOLUTION



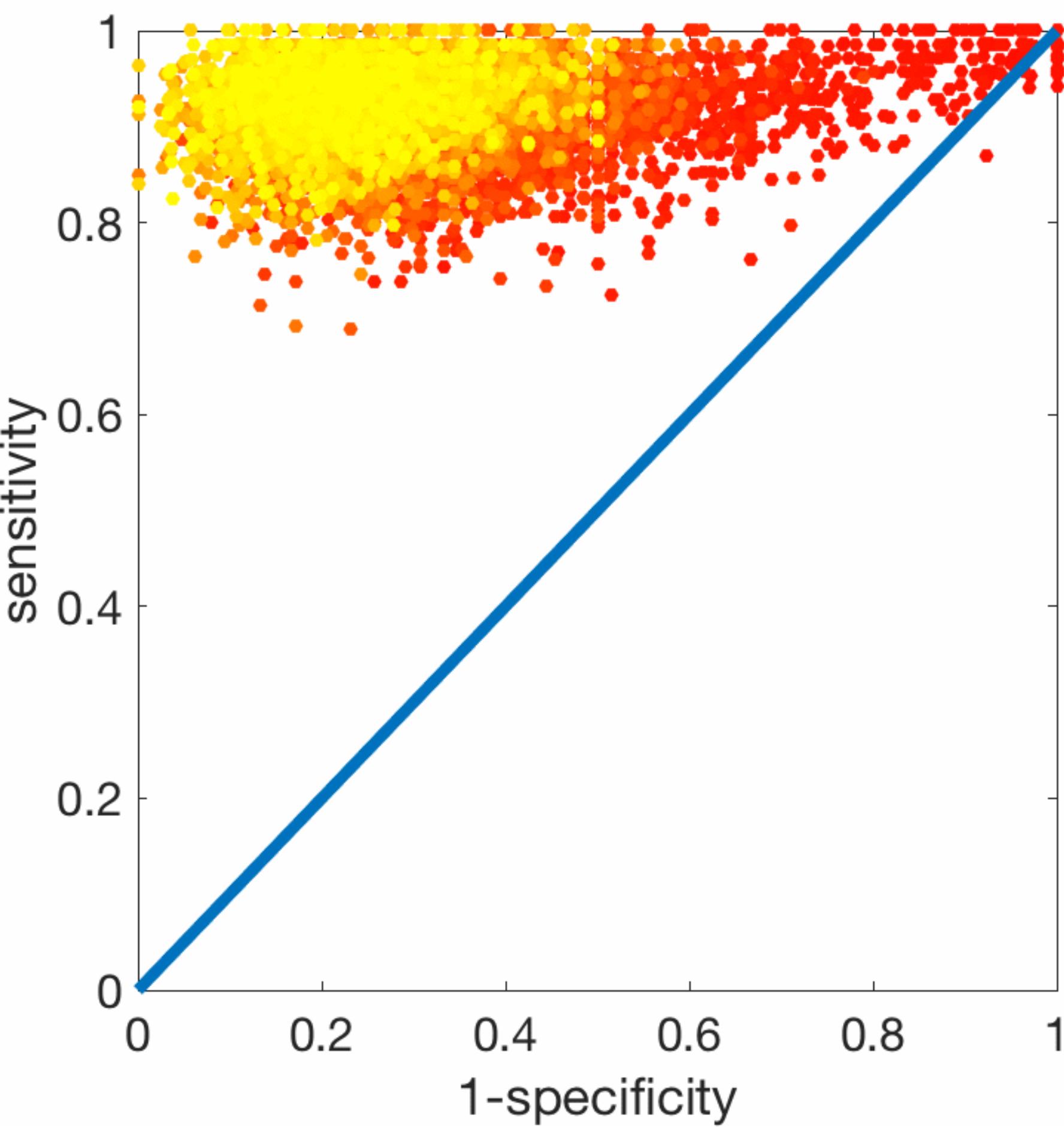
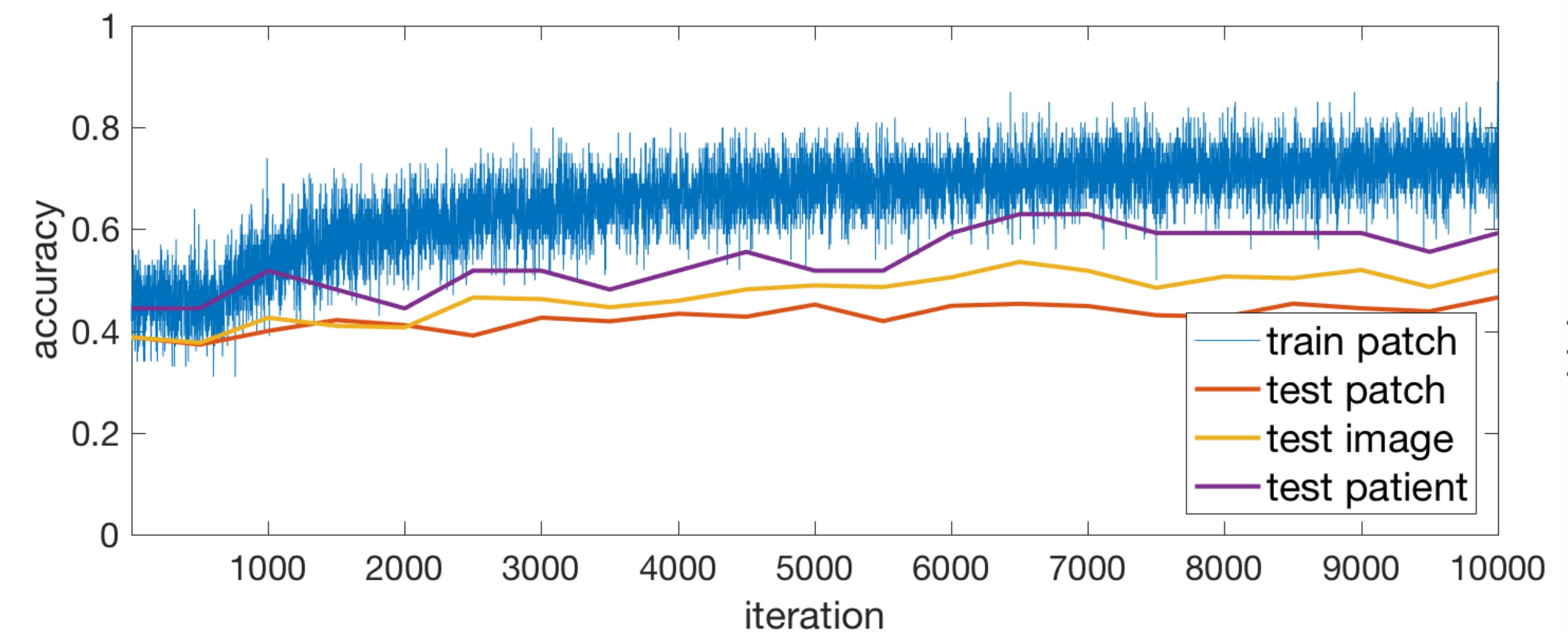
# CORRECTING THE BIAS?



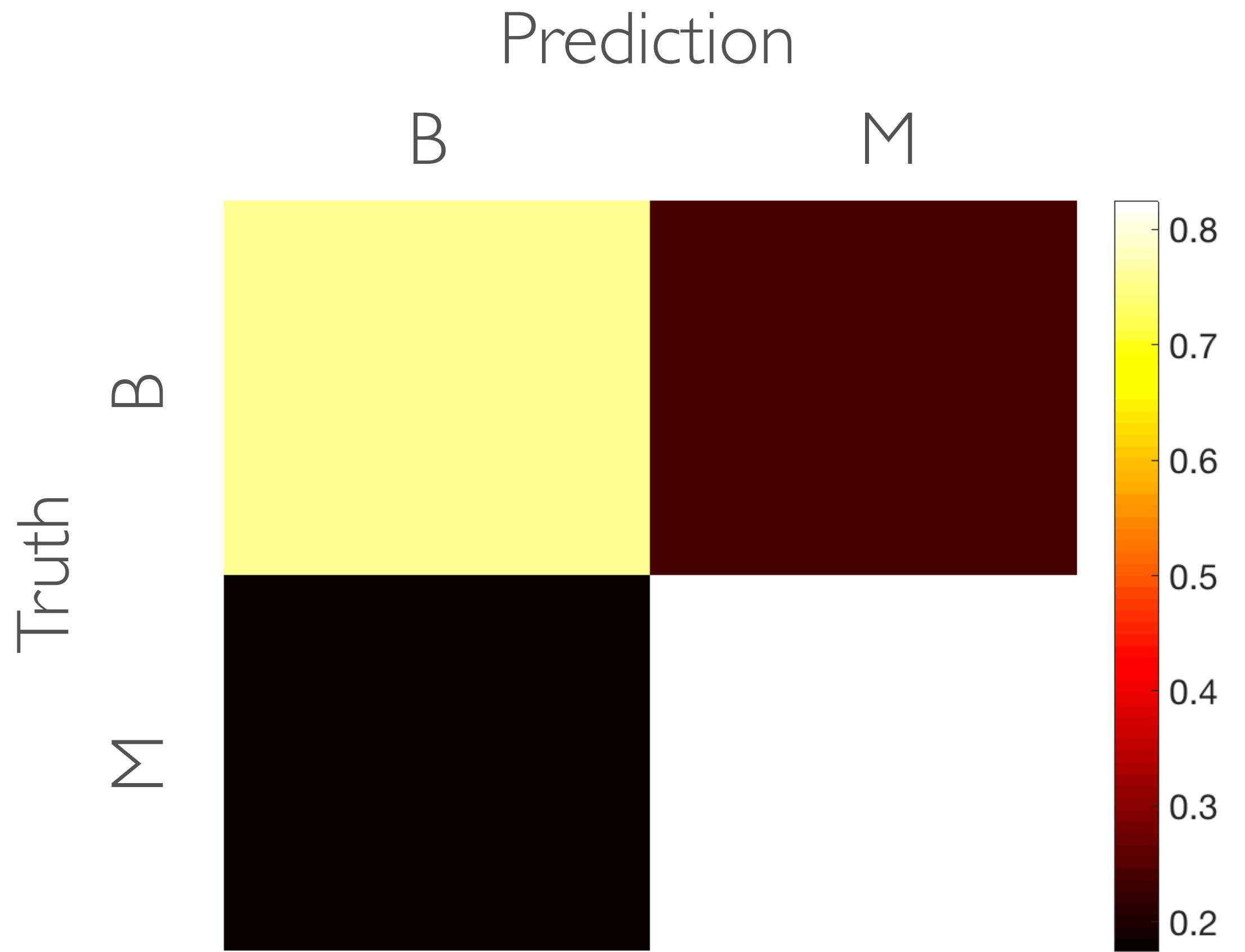
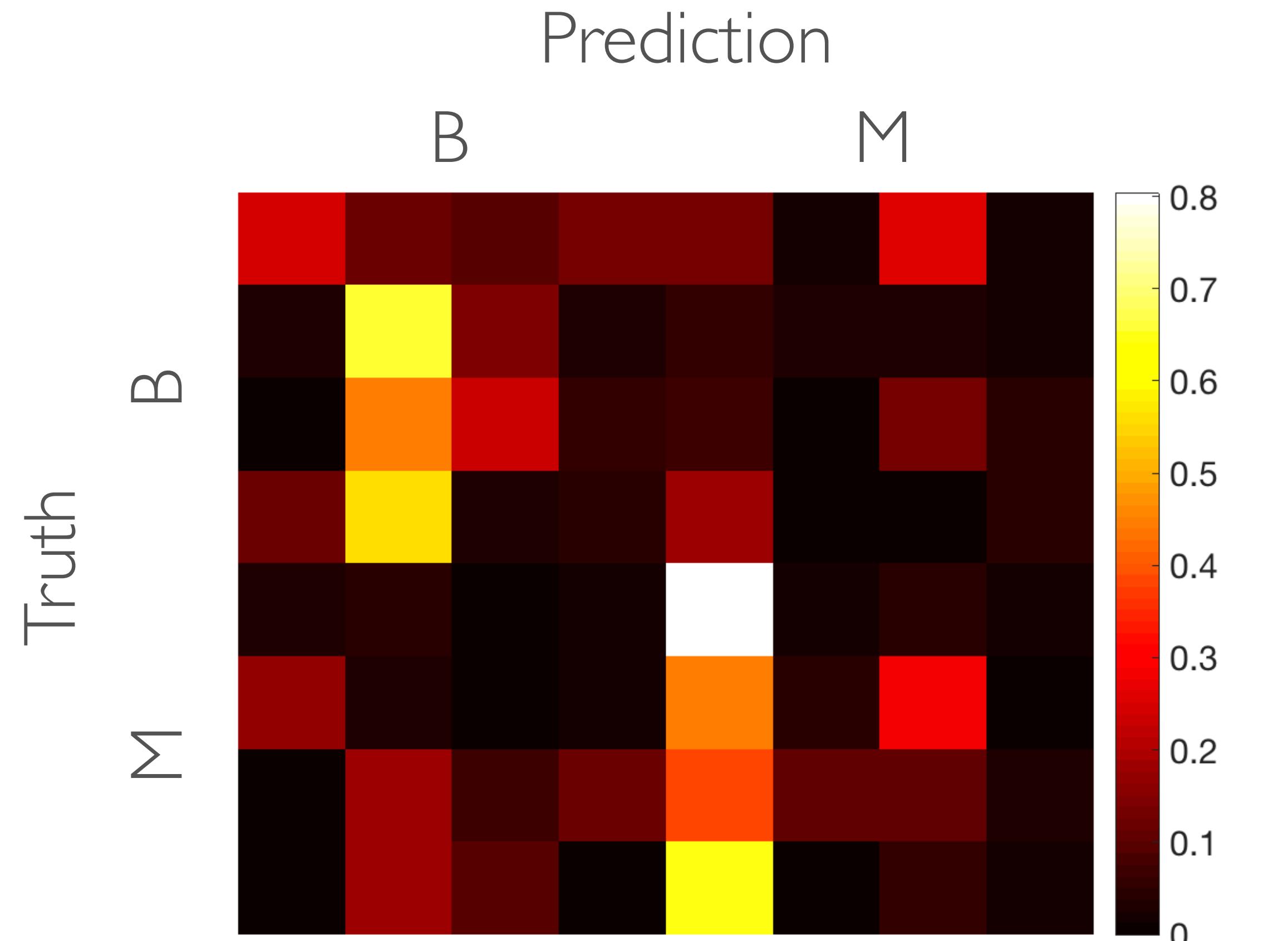
# MULTI-TYPE CLASSIFICATION



# HIGHER RESOLUTION?



# CONFUSION MATRIX



# CONCLUSION

- Model gives good results for binary classification
- Multi-class differentiation more challenging
- Project setting vs. Real life application
- For full evaluation: all folds + all magnifications



# THANKS!

# Questions?