



Институт профильного  
образования "Импульс"

Москва  
2024

# Применение теории нечетких множеств в анализе медицинских изображений

Филимонова Ольга  
Овсянников Алексей  
Бирюкова Наталья





$Y$  четкое множество

$$Y = \{x | x \in X\}$$

$\tilde{Y}$  нечеткое множество

$$\tilde{Y} = \{(x, \mu_{\tilde{Y}}(x)) | x \in X\}$$

$\mu_{\tilde{Y}}(x)$  функция принадлежности  $x$  к  $\tilde{Y}$

$$\mu_{\tilde{Y}} \in [0; 1]$$

Вероятность

$$p = 0,5$$



Вероятность того, что пиксель  
черный, составляет 50%

Функция принадлежности

$$\mu = 0,5$$



Пиксель на 50% принадлежит  
классу черных пикселей





# Метод нечеткой кластеризации C-средних (FCM)

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. In *Computers & Geosciences* (Vol. 10, Issues 2–3, pp. 191–203). Elsevier BV. [https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)

## Модификации FCM

Li, Y., & Li, G. (2010). Fast Fuzzy c-Means Clustering Algorithm with Spatial Constraints for Image Segmentation. In *Lecture Notes in Electrical Engineering* (pp. 431–438). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-12990-2\\_49](https://doi.org/10.1007/978-3-642-12990-2_49)

Chen, S., & Zhang, D. (2004). Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure. In *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* (Vol. 34, Issue 4, pp. 1907–1916). Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/tsmcb.2004.831165>

Szilágyi L, Benyó Z, Szilágyi SM, Adam HS (2003) MR brain image segmentation using an enhanced fuzzy c-means algorithm. *Proc 25th Ann Int Conf IEEE EMBS*, pp 724–726

Wu, J., Wang, X., Wei, T., & Fang, C. (2023). Full-parameter adaptive fuzzy clustering for noise image segmentation based on non-local and local spatial information. In *Computer Vision and Image Understanding* (Vol. 235, p. 103765). Elsevier BV. <https://doi.org/10.1016/j.cviu.2023.103765>

Krinidis, S., & Chatzis, V. (2010). A Robust Fuzzy Local Information C-Means Clustering Algorithm. In *IEEE Transactions on Image Processing* (Vol. 19, Issue 5, pp. 1328–1337). Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/tip.2010.2040763>

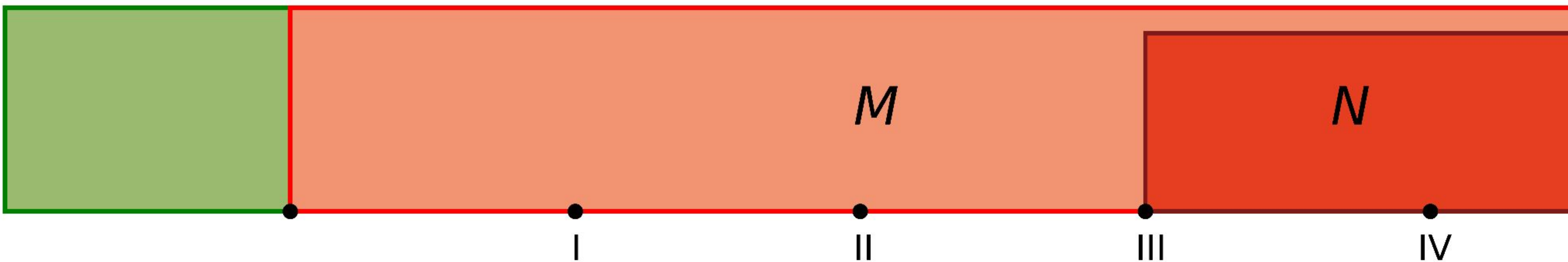
Wu, C., & Wang, Z. (2022). A modified fuzzy dual-local information c-mean clustering algorithm using quadratic surface as prototype for image segmentation. In *Expert Systems with Applications* (Vol. 201, p. 117019). Elsevier BV. <https://doi.org/10.1016/j.eswa.2022.117019>

Song, Q., Wu, C., Tian, X., Song, Y., & Guo, X. (2022). Kernel-based fuzzy local information clustering algorithm self-integrating non-local information. In *Digital Signal Processing* (Vol. 122, p. 103351). Elsevier BV. <https://doi.org/10.1016/j.dsp.2021.103351>

Wei, T., Wang, X., Li, X., & Zhu, S. (2022). Fuzzy subspace clustering noisy image segmentation algorithm with adaptive local variance & non-local information and mean membership linking. In *Engineering Applications of Artificial Intelligence* (Vol. 110, p. 104672). Elsevier BV. <https://doi.org/10.1016/j.engappai.2022.104672>

Нет рака

Стадии рака



$P_M$  вероятность наличия рака любой стадии

Например,

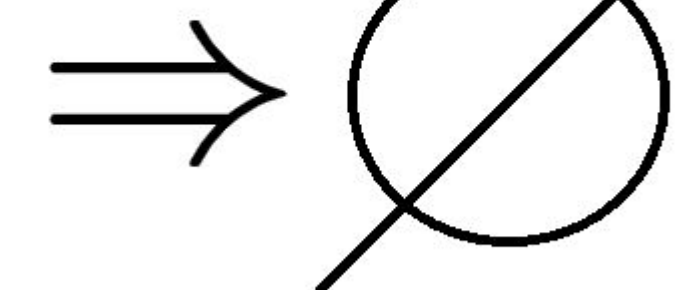
$$P_M = 0,2$$

$P_N$  вероятность наличия рака третьей стадии

$$P_N = 0,6$$

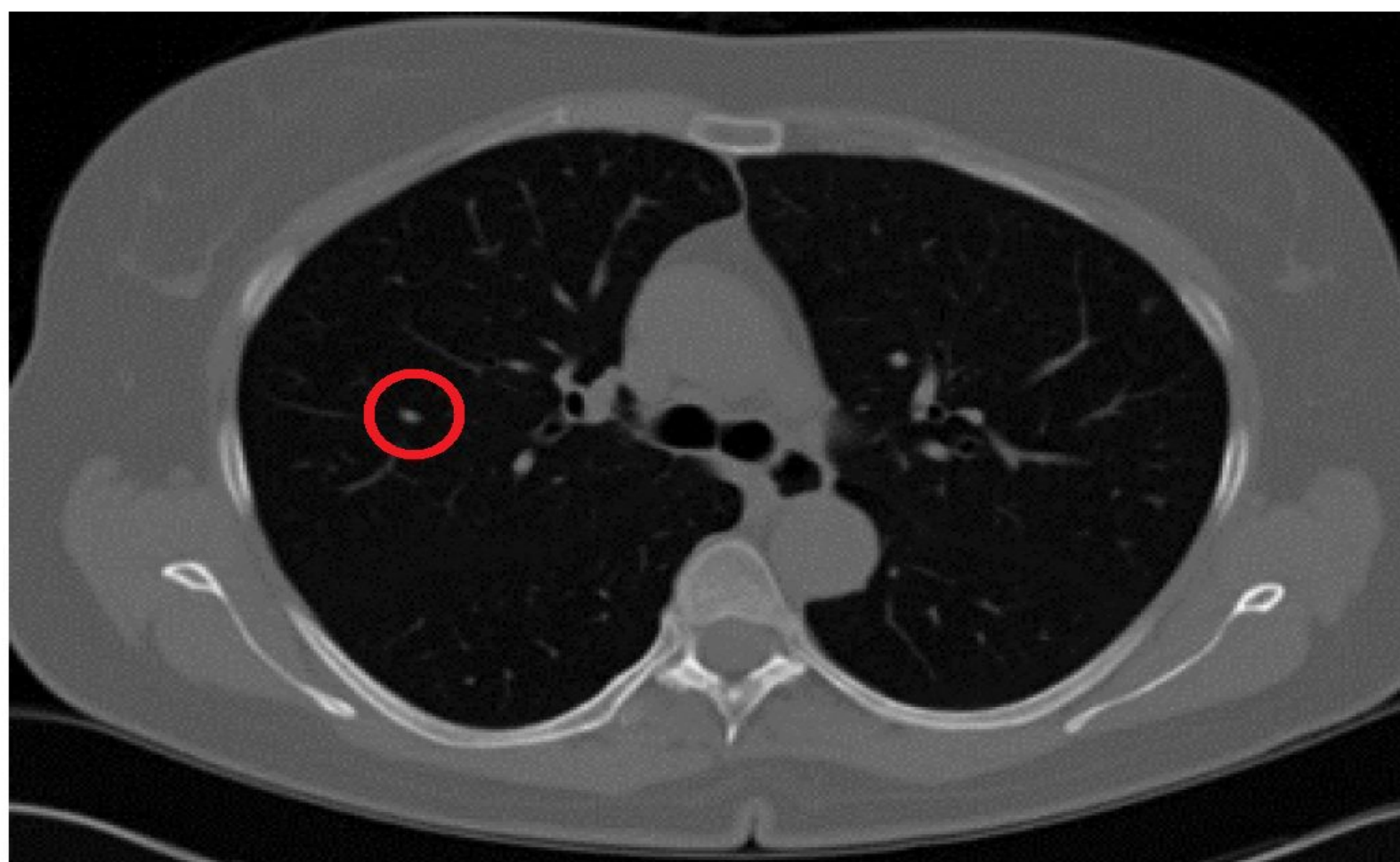
Но

$$N \subseteq M$$

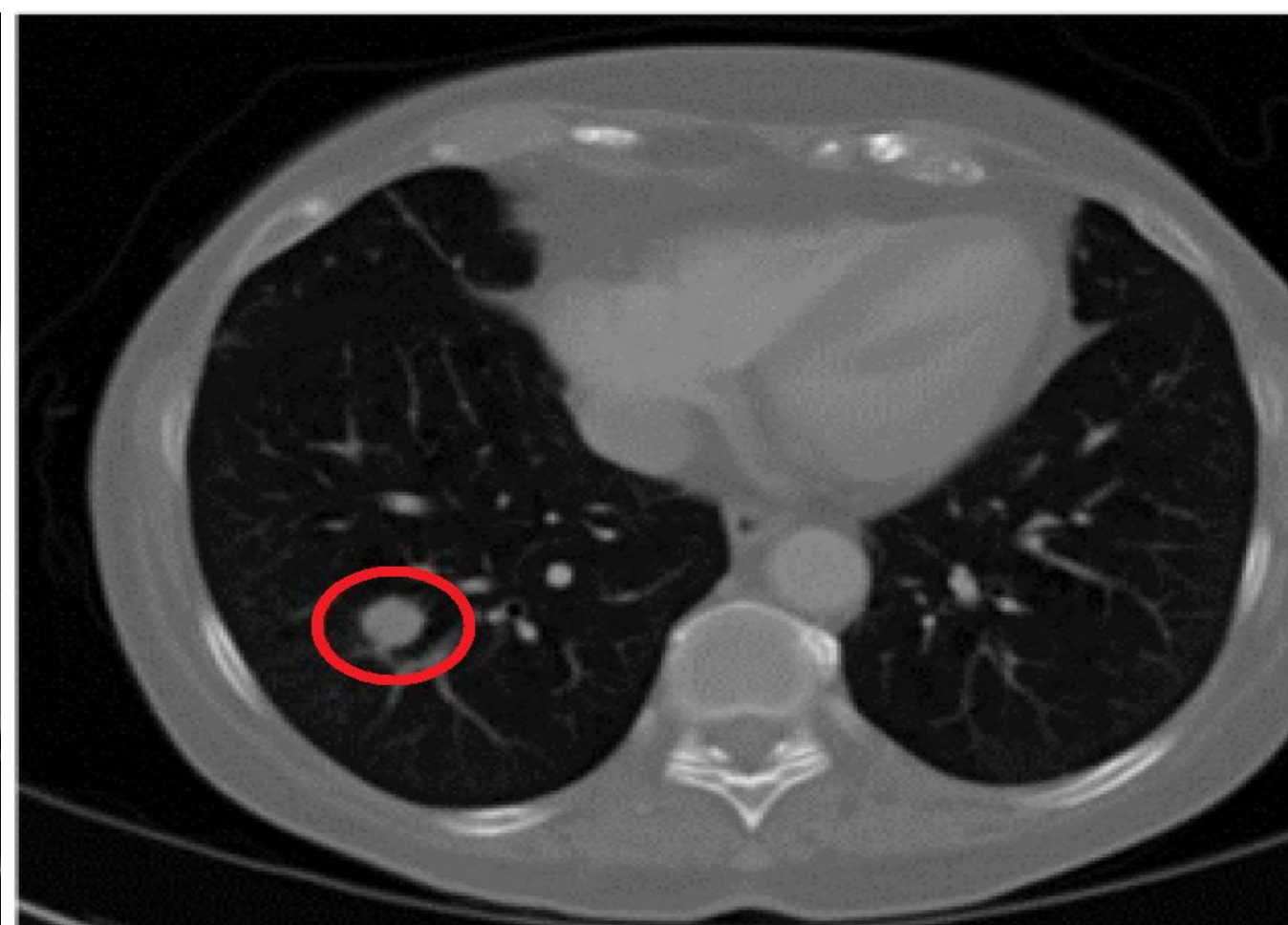


$$P_N \leq P_M$$

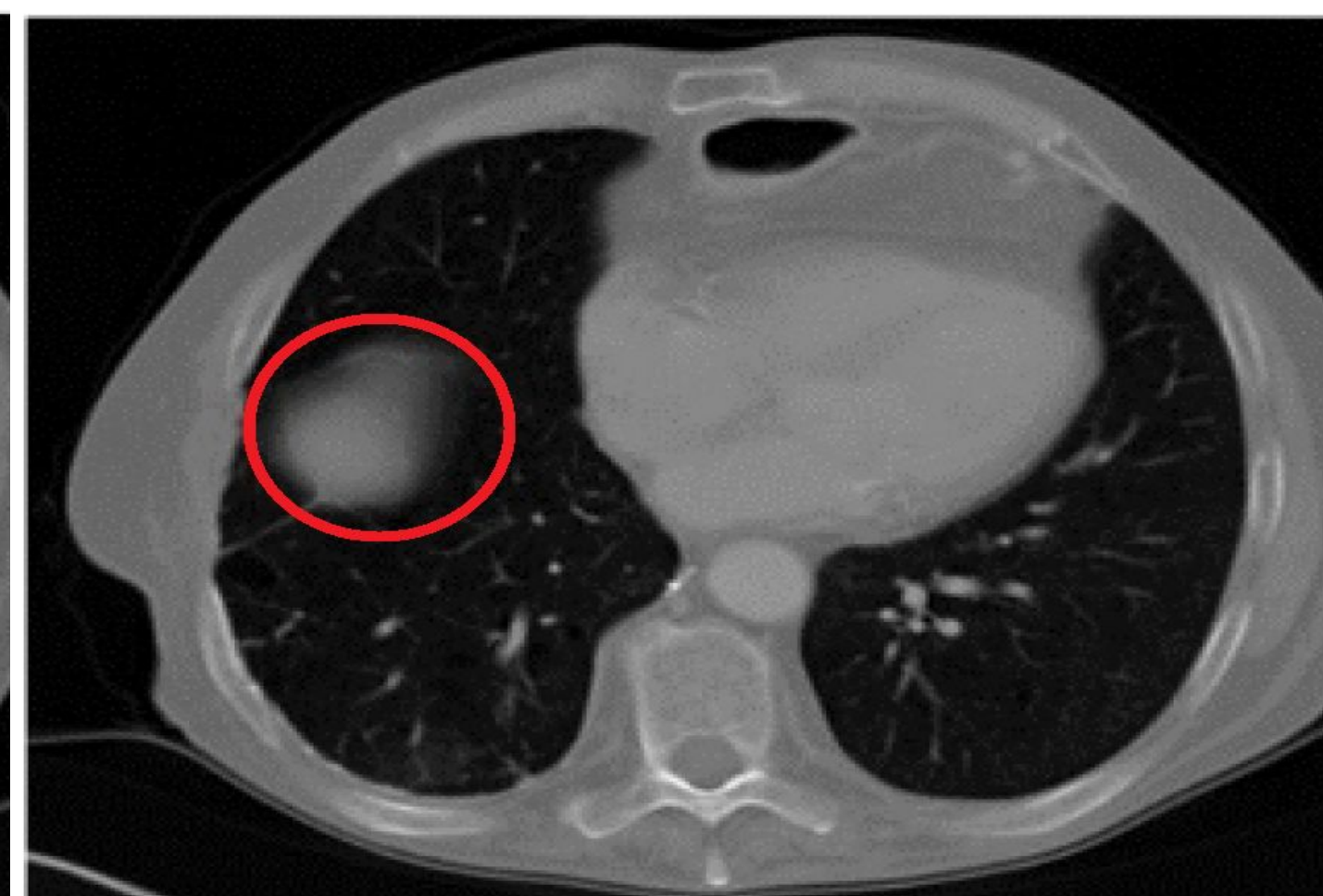
## Стадия II



## Стадия III



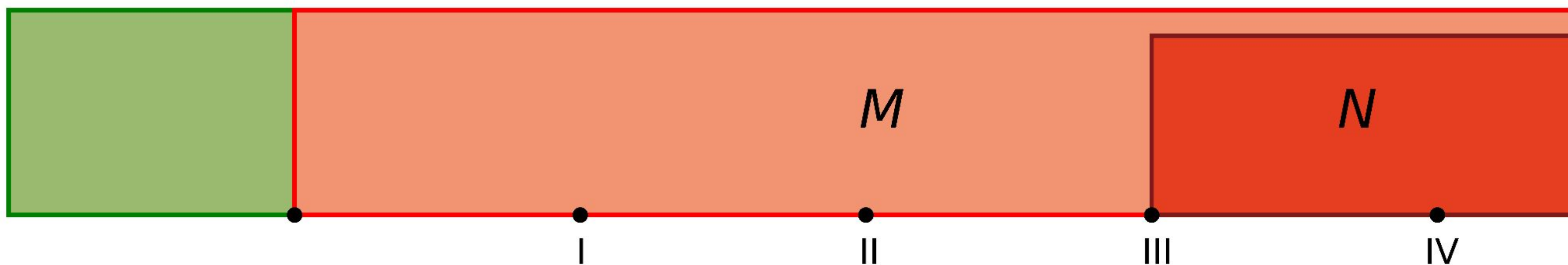
## Стадия IV



Снимки КТ, демонстрирующие разные стадии рака легкого<sup>1</sup>.

Нет рака

Стадии рака



Калибровка  $P_M$   
с помощью  $P_N$



# Классические алгоритмы калибровки вероятностей

## Модель Платта

Platt, John. (2000). Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. *Adv. Large Margin Classif.* 10.

## Изотоническая регрессия

Chakravarti, N. (1989). Isotonic Median Regression: A Linear Programming Approach. *Mathematics of Operations Research*, 14(2), 303–308. <http://www.jstor.org/stable/3689709>

## Гистограммная калибровка

Bianca Zadrozny and Charles Elkan. 2001. Obtaining calibrated probability estimates from decision trees and naive bayesian classifiers. In *Proceedings of the Eighteenth International Conference on Machine Learning*. 609--616.

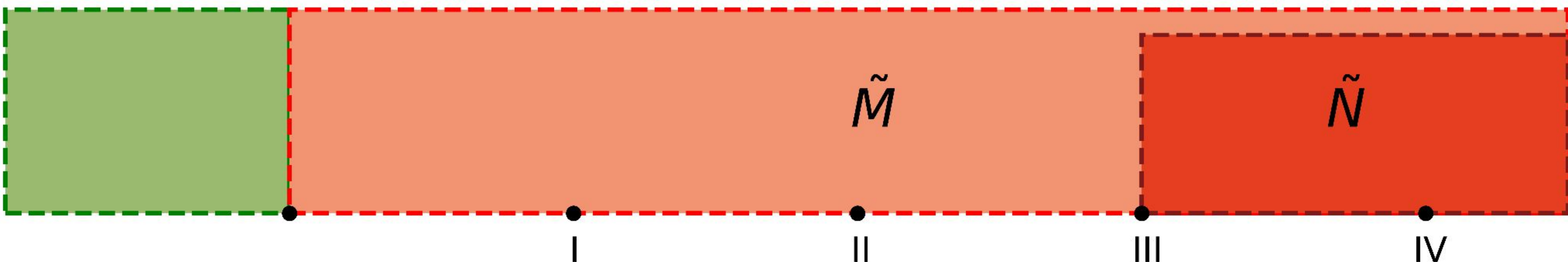
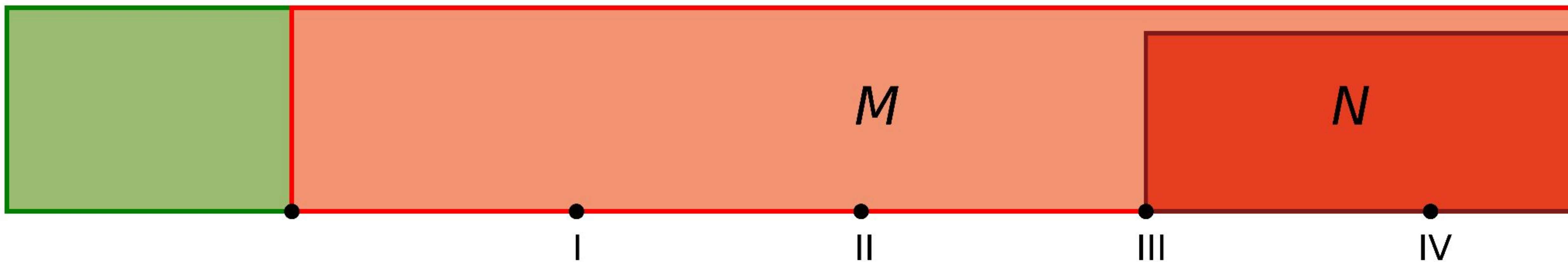
## Метода ВВQ (Binning into Quantiles)

Naeini MP, Cooper GF, Hauskrecht M. Obtaining Well Calibrated Probabilities Using Bayesian Binning. *Proc AAAI Conf Artif Intell*. 2015 Jan;2015:2901-2907. PMID: 25927013; PMCID: PMC4410090.



## Нет рака

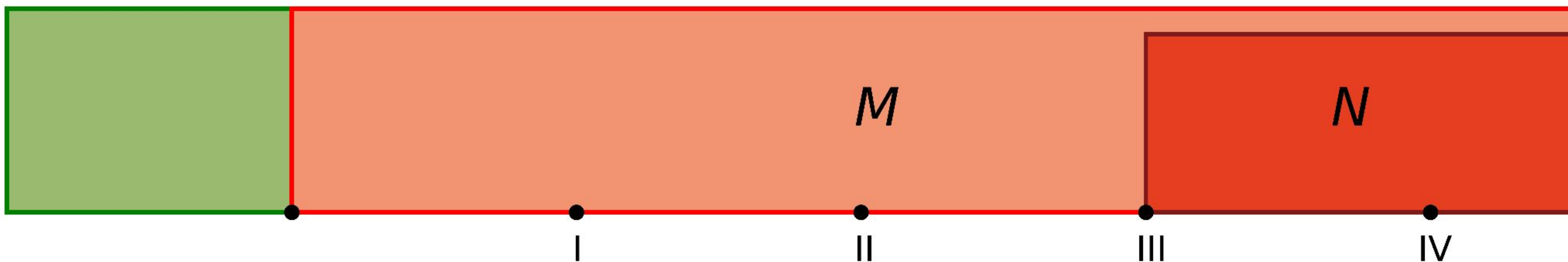
## Стадии рака





Нет рака

Стадии рака



Вероятность наличия рака любой стадии для  $i$ -ого пациента<sup>2</sup>

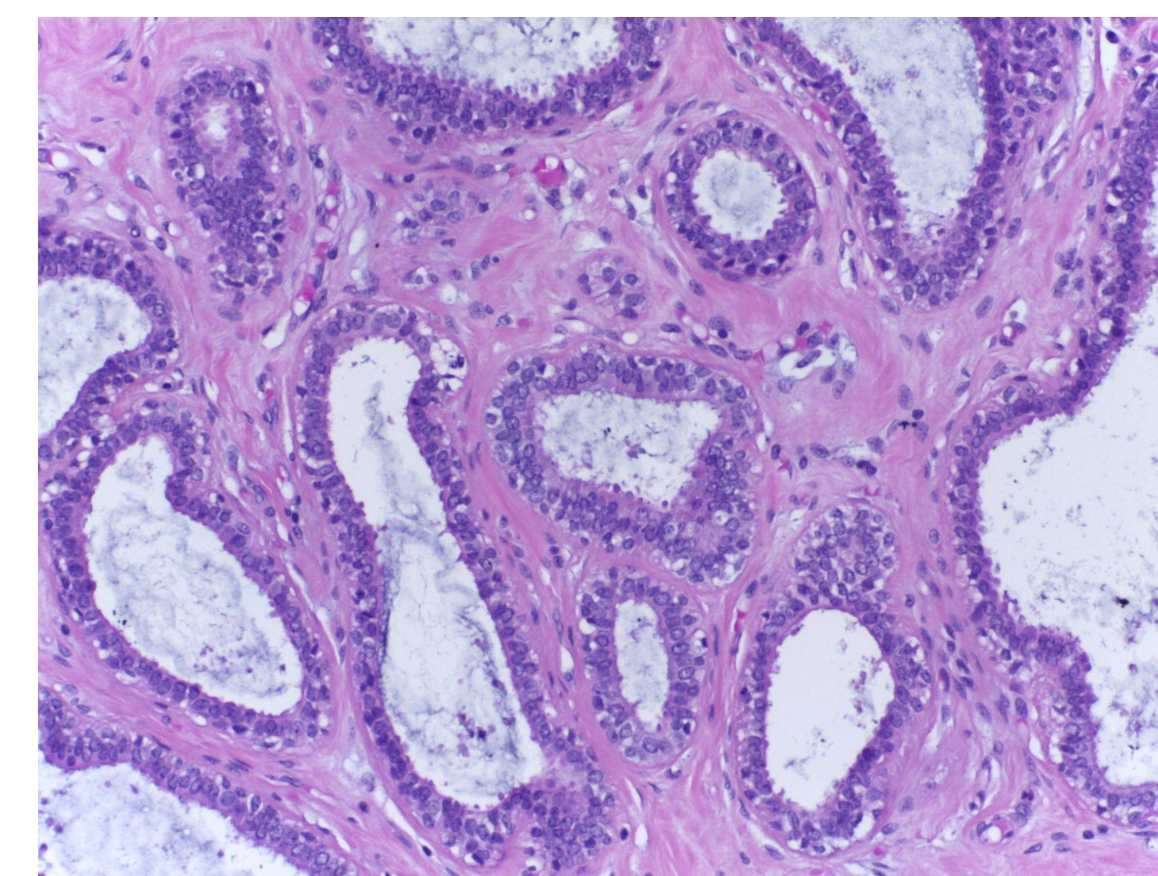
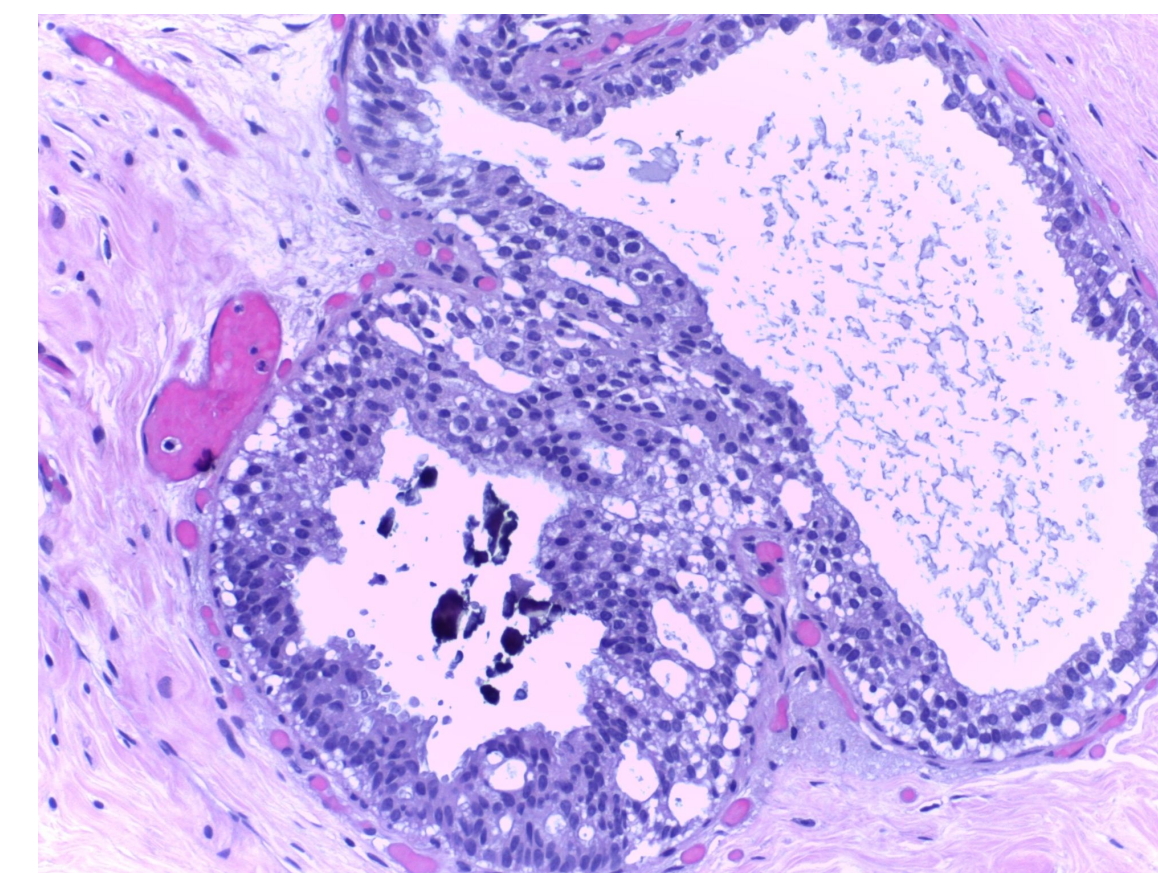
$$\theta_i = \frac{1}{2} \left( \frac{\max\{P_N(\hat{x}_i), P_M(\hat{x}_i)\}}{1 + \exp(A \cdot \min\{P_N(\hat{x}_i), P_M(\hat{x}_i)\} + B)} - \frac{\min\{1 - P_N(\hat{x}_i), 1 - P_M(\hat{x}_i)\}}{1 + \exp(C \cdot \max\{1 - P_N(\hat{x}_i), 1 - P_M(\hat{x}_i)\} + D)} + 1 \right)$$

$$A, B, C, D \in \mathbb{R}$$

## Рак молочной железы

400 гистологических изображений молочной железы<sup>3</sup>:

- **множество здоровых пациентов**  
200 изображений или без патологий, или с выявленными доброкачественными новообразованиями
- **множество пациентов с ранней стадией рака**  
100 изображений с выявленным преинвазивным раком (Carcinoma in situ)
- **множество пациентов с поздней стадией рака**  
100 изображений с выявленной инвазивной карциномой (Invasive carcinoma)

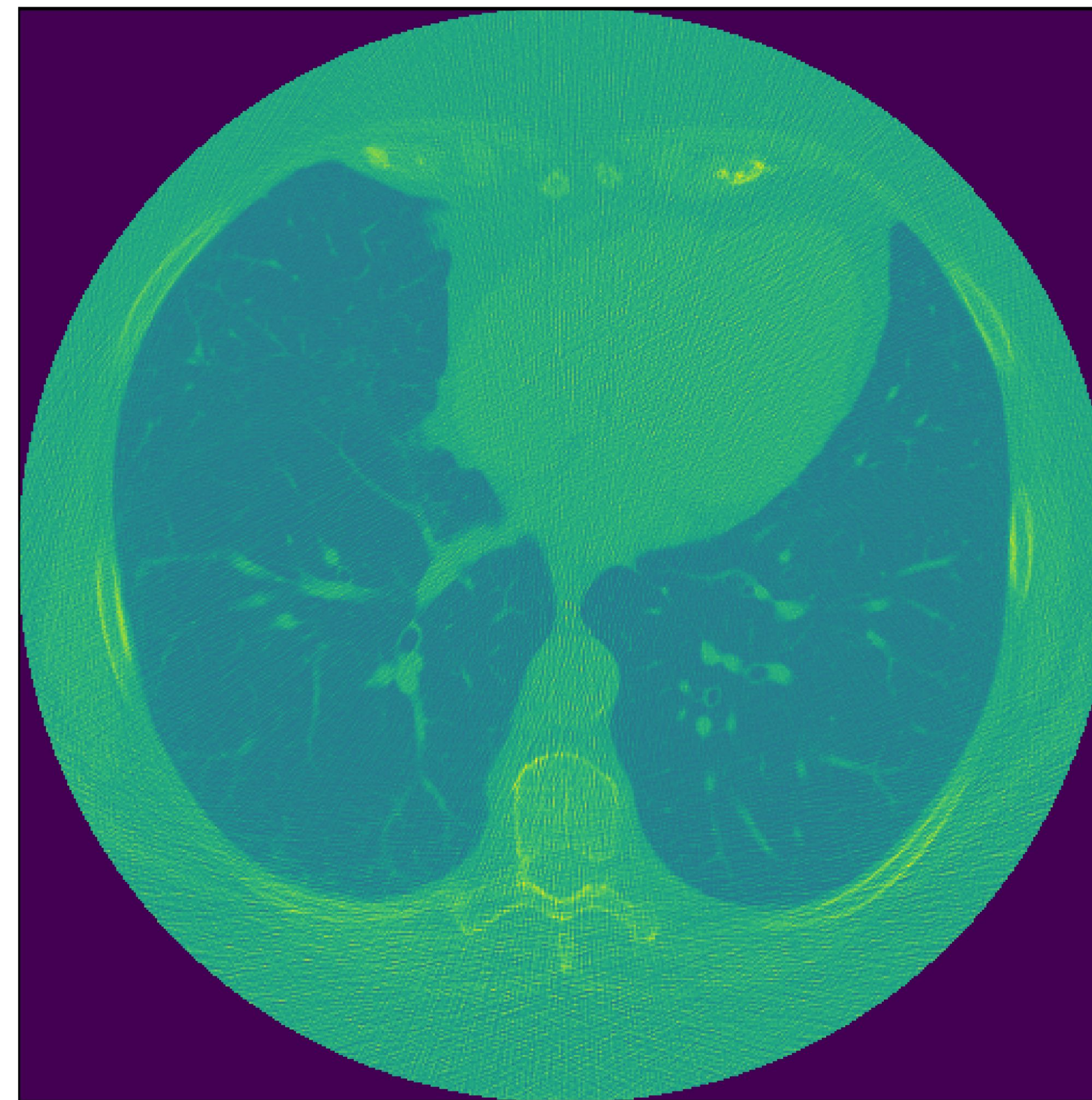




## Рак легкого

2159 снимков КТ из архива изображений рака<sup>4,5</sup>:

- **множество здоровых пациентов**  
2113 изображений
- **множество пациентов с первой стадией рака**  
20 изображений
- **множество пациентов со второй, третьей или четвертой стадией рака**  
26 изображений



4. Armato III, Samuel G., McLennan et al (2015). Data From LIDC-IDRI (Version 4) [dataset]. The Cancer Imaging Archive. <https://doi.org/10.7937/K9/TCIA.2015.LO9QL9SX>

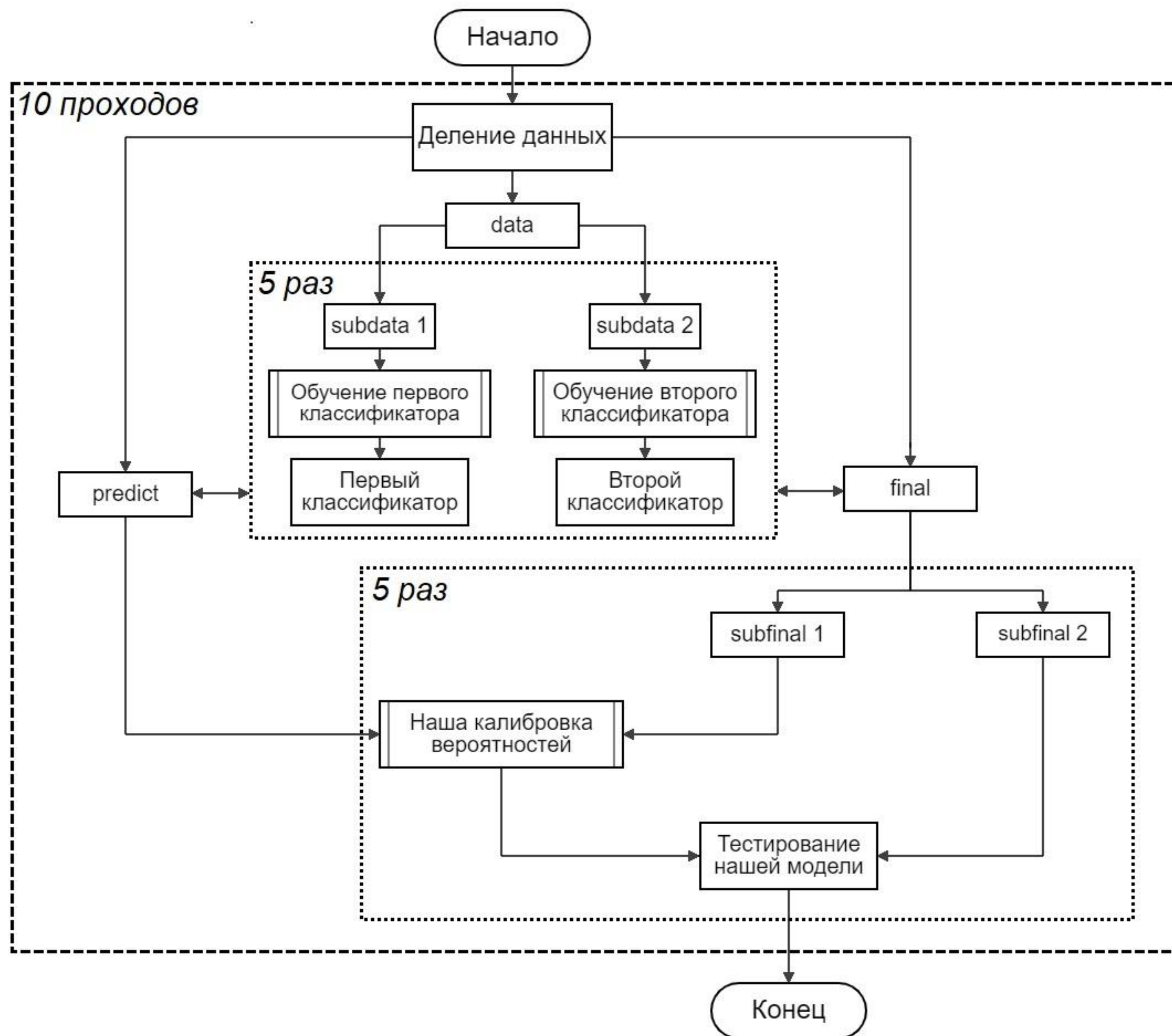
5. National Lung Screening Trial Research Team. (2013). Data from the National Lung Screening Trial (NLST) (Version 3) [dataset]. The Cancer Imaging Archive. <https://doi.org/10.7937/TCIA.HMQ8-J677>

## Объемы выборок

Форма рака	<i>subdata 1</i>	<i>subdata 2</i>	<i>predict</i>	<i>subfinal 1</i>	<i>subfinal 2</i>
РМЖ	120	160	160	40	40
РЛ	1203	1214	405	270	270
	1042	1052	567	270	270

Одно тестирование для рака молочной железы

Два тестирования с разным объемом выборок для рака легкого



## Логарифмическая потеря

$$\text{Log Loss} = -\frac{1}{l} \cdot \sum_{i=1}^l (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i))$$

$y_i$  истинный класс  $i$ -ого  
пациента

## Оценка Брайера

$$\text{BS} = \frac{1}{l} \sum_{i=1}^l (\hat{y}_i - y_i)^2$$

$\hat{y}_i$  предсказанная  
вероятность для  $i$ -  
ого пациента

$l$  размер выборки

## Площадь под кривой Precision-Recall (AUC-PR)



- Логарифмическая потеря
- Оценка Брайера
- Площадь под кривой Precision-Recall

## Относительное изменение метрики

$$\Delta = \frac{M_c - M_o}{M_o} \cdot 100\%$$

$M_o$  метрика, рассчитанная для исходных вероятностей

$M_c$  метрика, рассчитанная для откалиброванных вероятностей

## Классические методы калибровки

- Все модели, за исключением модели Платта, привели к переобучению
- Модель Платта оставила метрики без существенных изменений

## Наша модель: рак молочной железы

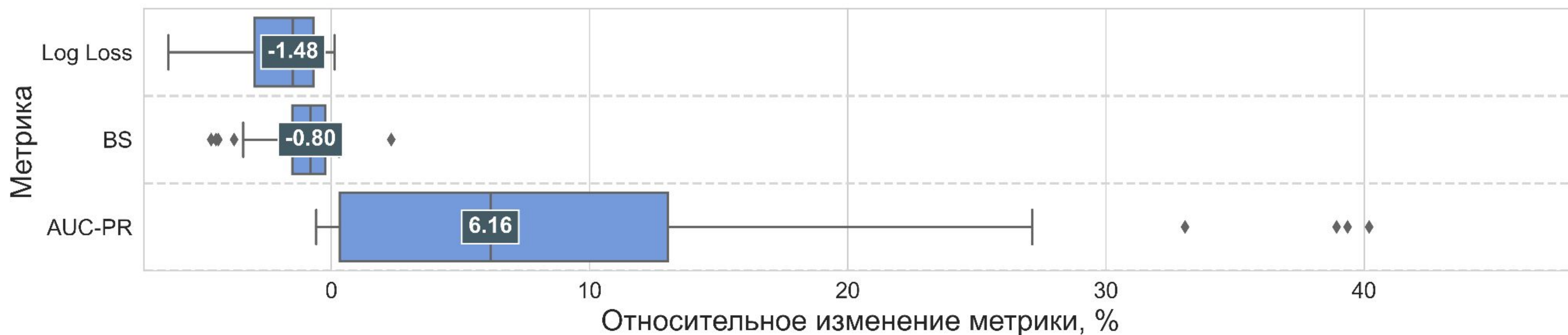


Диаграмма размаха относительного изменения метрики, полученная после работы нашего метода в случае рака молочной железы.

## Наша модель: рак легкого

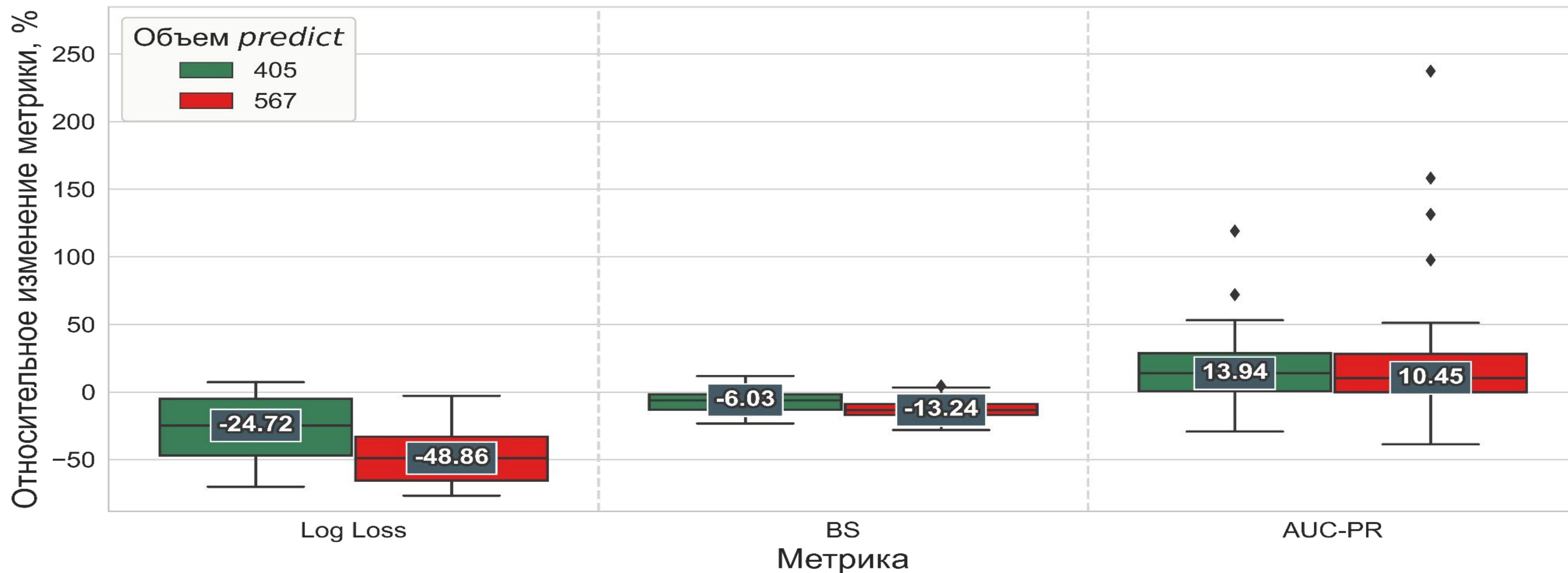
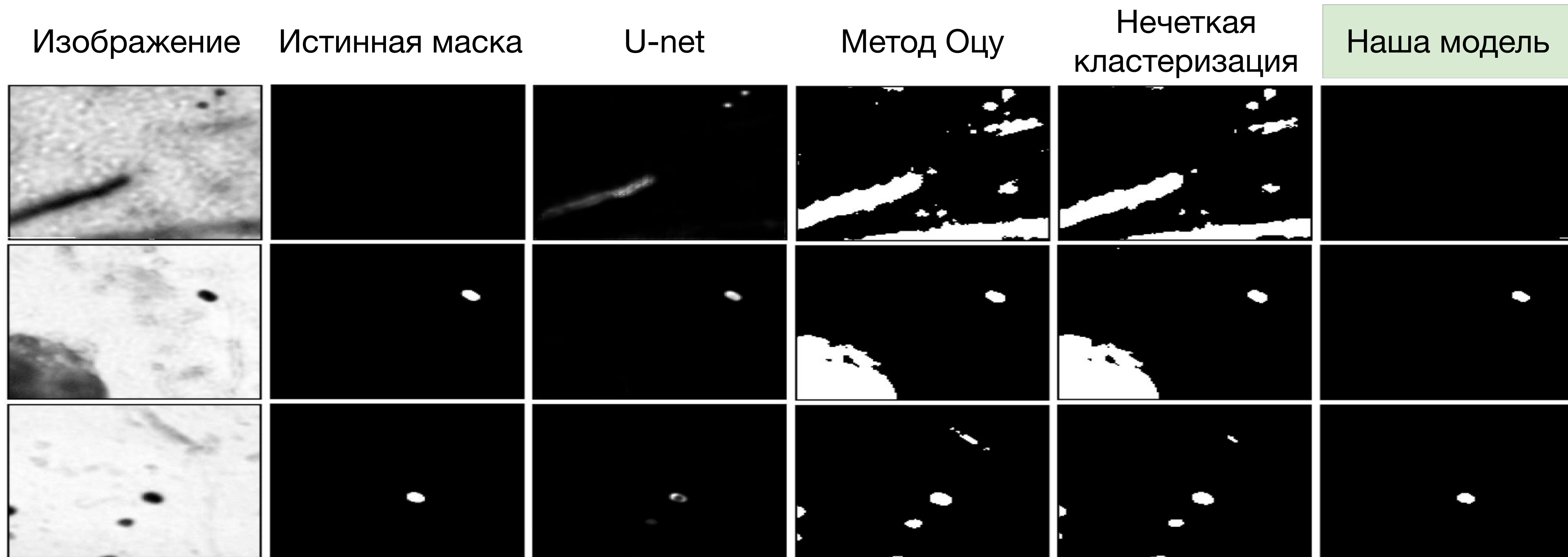


Диаграмма размаха относительного изменения метрики, полученная после работы нашего метода в случае рака легкого.





## Набор клинических бактериальных данных<sup>6</sup>



Сравнение разных алгоритмов сегментации

6. Wang, X., Shi, Y., Guo, S., Qu, X., Xie, F., Duan, Z., Hu, Y., Fu, H., Shi, X., Quan, T., Wang, K., & Xie, L. (2024). A Clinical Bacterial Dataset for Deep Learning in Microbiological Rapid On-Site Evaluation. In Scientific Data (Vol. 11, Issue 1). Springer Science and Business Media LLC. <https://doi.org/10.1038/s41597-024-03370-5>



Институт профильного  
образования “Импульс”

Применение теории нечетких  
множеств в анализе медицинских  
изображений

Заключение

Ольга Филимонова  
filimonova\_o\_a@staff.sechenov.ru

Спасибо за внимание!