

Статистический анализ анапластической трансформации пациентов с глиальными опухолями супратенториальной локализации

А.М. Исмаилов

**Федеральное государственное автономное образовательное учреждение
высшего образования «Национальный исследовательский университет
«Высшая школа экономики», Российская Федерация, 101000, г. Москва,
ул. Мясницкая, д. 20**

*** Контактная информация: Исмаилов Алы Мехтиевич, магистр ВШЭ
ФКН АДБМ**

Email: neuro.promotion@gmail.com

Введение:

Проведен подробный статистический анализ пациентов претерпевших анапластическую трансформацию глиальных опухолей супратенториальной локализации головного мозга. Глиальные опухоли - одни из самых частых первичных опухолей головного мозга. В соответствии с классификацией выделяют 4 степени злокачественности данных опухолей: Grade I-IV. Первые две степени считаются “доброкачественными”. Grade I - редкий случай, чаще представленный гистологической формой под названием пилоцитарная астроцитома. В нашем датасете таких опухолей нет. Grade II - частый нейроонкологический феномен, подразделяющийся по старой классификации на три основных типа в соответствии с гистологическим происхождением: астроцитомы - опухоли происходящие из астроцитарной глии, олигодендроглиомы - предшественники олигодендроглии и смешанные опухоли, имеющие как астро- так и олиго- компоненты - олигоастроцитомы. Grade III-IV - опухоли злокачественные, нередко происходящие от их “доброкачественных” предшественников, как правило глиом Grade II. Grade III глиома называется анапластической астроцитомой / анапластической олигоастроцитомой / анапластической олигодендроглиомой. Grade IV - представлена в виде единственной гистологической формы - глиобластомы. Процесс перехода от Grade II к Grade III-IV получил название “анапластическая трансформация”.

Цель исследования:

Выяснить влияние различных факторов на процесс анапластической трансформации, таких как: гистологическое происхождение опухоли, экспрессия

Ki-67, индекс Карновского (Karnofsky Performance Status), тип проводимой операции.

Материалы и методы:

Для проведения анализа был взят датасет, составленный лично мной в процессе обучения в университете, в институте нейрохирургии им. акад. Н.Н. Бурденко. Датасет состоит из данных 139 пациентов. Все пациентам было проведено как минимум две нейрохирургические операции: первая с патоморфологической верификацией Grade II глиомы, вторая с верификацией анапластической трансформации в Grade III или Grade IV опухоль. Период между первой и второй операцией мы будем называть “временем до трансформации”. Пациенты с временем до трансформации менее 4 месяцев были отфильтрованы, поскольку вероятно была допущена ошибка при первой гистологической верификации. После исключения таких пациентов в датасете осталось 134 наблюдения. У части пациентов есть информация об индексе пролиферативной активности [Ki-67](#). Он может быть высоким, средним и низким. Всего могло проводиться два типа операции: либо удаление опухоли, либо СТБ - стереотаксическая биопсия. Также у всех пациентов есть информация об [индексе Карновского](#) (KPS) до операции - он отражает степень функциональной дееспособности организма. В конце концов у каждого пациента есть информация о том в какую опухоль трансформировалась изначальная доброкачественная глиома: в Grade III либо сразу в глиобласту Grade IV.

	histology_1	ki67	surgery_type	kps	time_to_trans	histology_2
0	астроцитомы	*	удаление	70	22	глиобластома
1	астроцитомы	*	СТБ	90	56	глиобластома
2	астроцитомы	*	СТБ	70	38	анапластическая астроцитомы
3	астроцитомы	*	удаление	90	31	глиобластома
4	астроцитомы	*	СТБ	90	34	глиобластома
5	олигодендроглиомы	*	удаление	80	136	анапластическая олигодендроглиомы
6	астроцитомы	*	удаление	90	87	глиобластома
7	олигоастроцитомы	высокий	удаление	90	132	глиобластома
8	астроцитомы	*	удаление	80	17	анапластическая астроцитомы
9	олигодендроглиомы	*	СТБ	90	110	анапластическая олигодендроглиомы

Все статистические тесты и визуализация были проведены с помощью языка программирования [Python](#). Визуализация проводилась с подключением библиотек [Matplotlib](#) и [Seaborn](#). Для статистической обработки использовались библиотеки [Pandas](#), [Numpy](#), [SciPy](#). Проверка распределения на нормальность проводилась с помощью критерия согласия Колмогорова-Смирнова и теста Шапиро-Уилка: [kstest](#) и [shapiro](#) из библиотеки SciPy. Для проведения двухвыборочного теста Стьюдента использовался функция [ttest_ind](#) из модуля

SciPy. Для критерия Манна-Уитни использовалась функция [mannwhitneyu](#) из модуля SciPy. Однофакторный дисперсионный анализ был проведен с помощью функции [f_oneway](#) из модуля SciPy. Точный тест Фишера был использован с помощью функции [fisher_exact](#) из модуля SciPy.

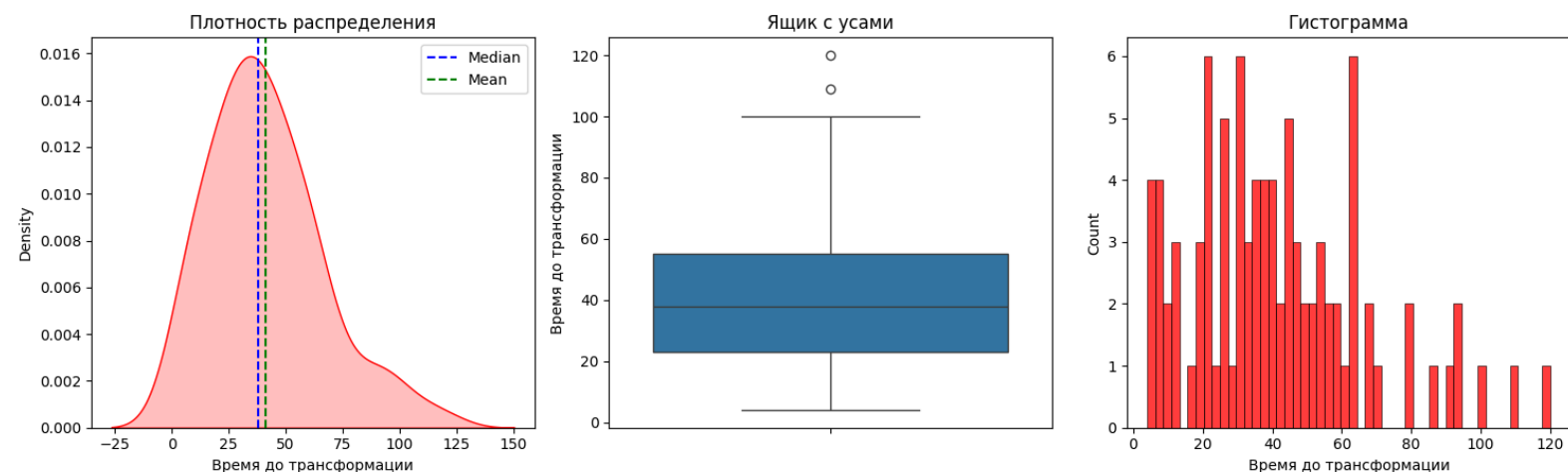
[Ссылка](#) на ноутбук со всеми калькуляциями и визуализациями.

Результаты.

• Описательная статистика

Гистологический тип астроцитомы - самый обширный в нашем датасете. Таких пациентов 91. Ниже представлены графики плотности распределения, ящик с усами и гистограмма.

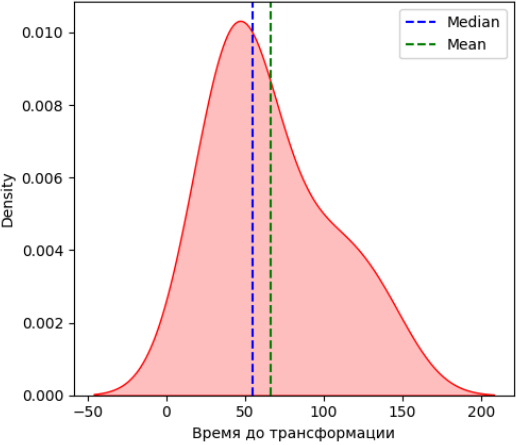
Описательная статистика гистологического типа "астроцитомы"



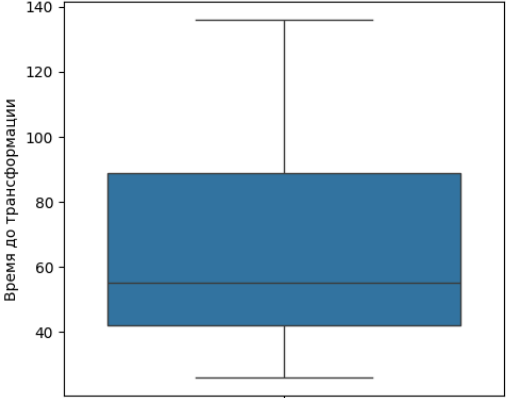
Олигодендроглиом меньше всего, их 9. Ниже представлены графики плотности распределения, ящик с усами и гистограмма.

Описательная статистика гистологического типа "олигодендроглиомы"

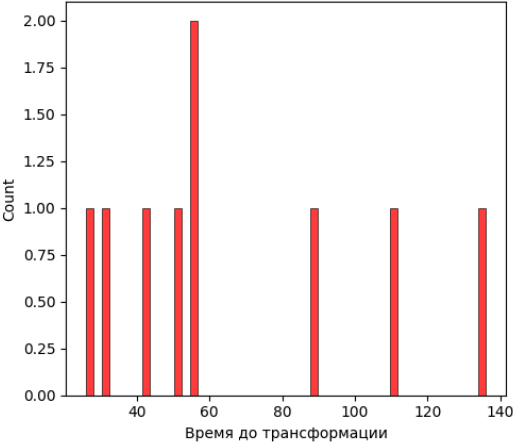
Плотность распределения



Ящик с усами



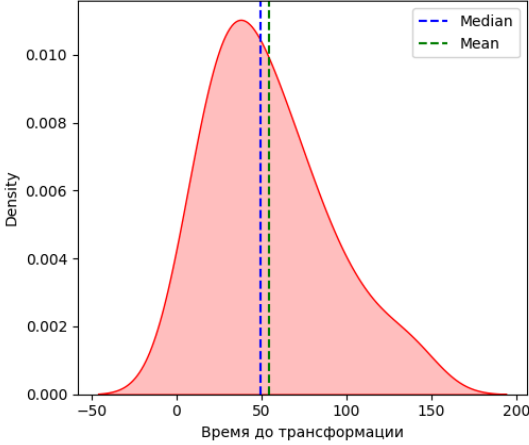
Гистограмма



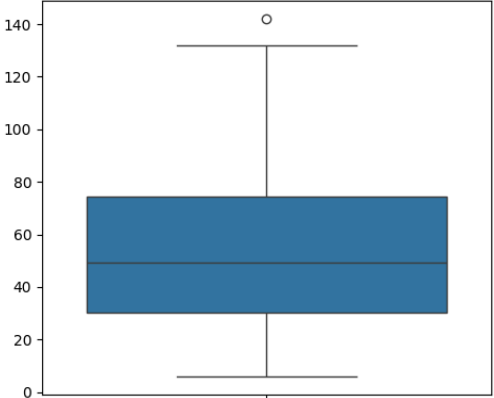
Олигоастроцитом 34. Ниже представлены графики плотности распределения, ящик с усами и гистограмма.

Описательная статистика смешанного гистологического типа "олигоастроцитомы"

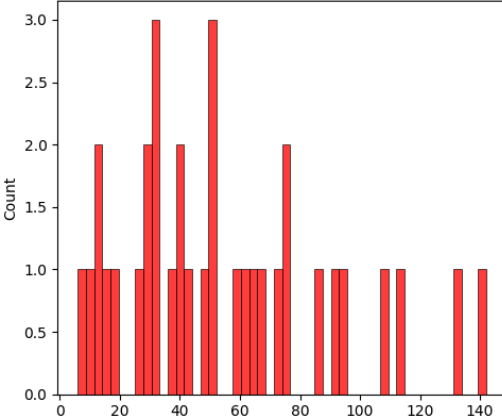
Плотность распределения



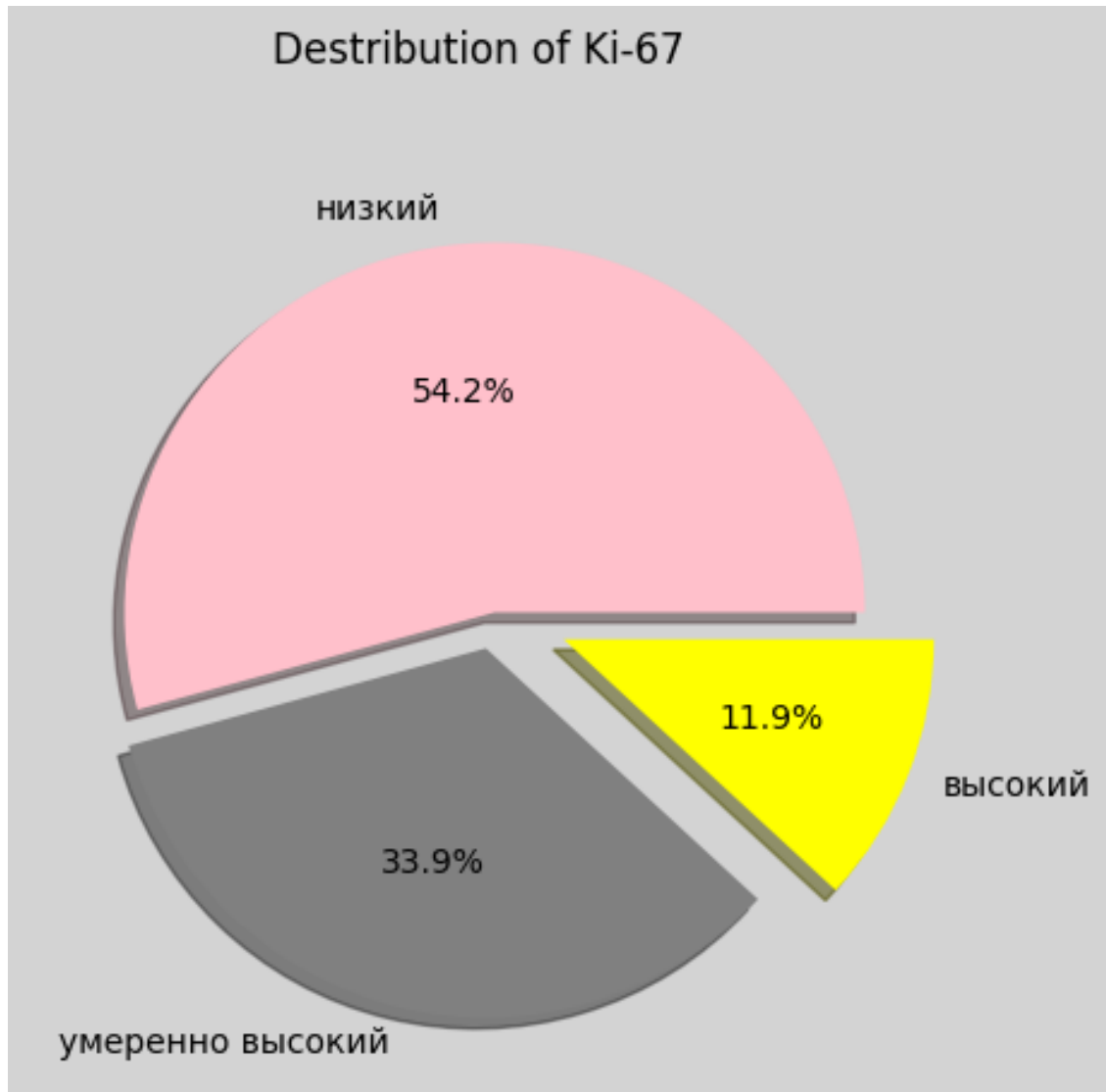
Ящик с усами



Гистограмма

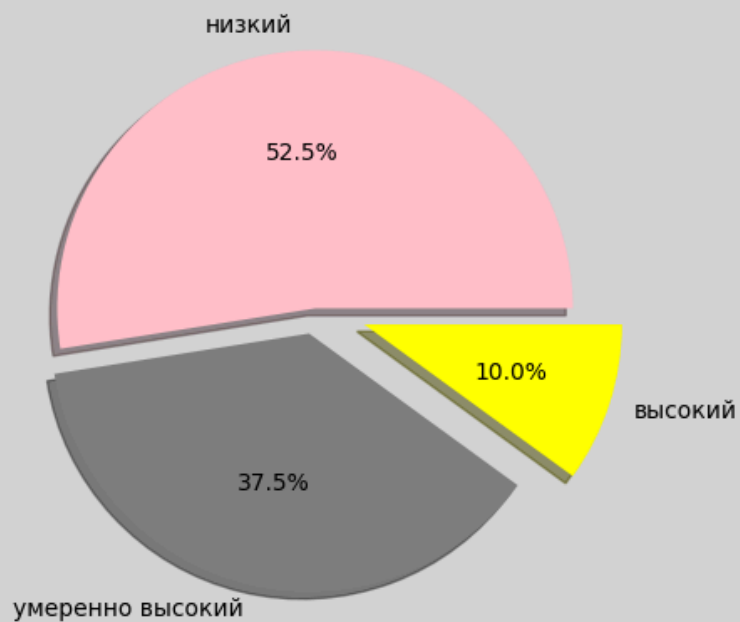


Данные об индексе Ki-67 были у 59 пациентов. Посмотрим на круговую диаграмму:

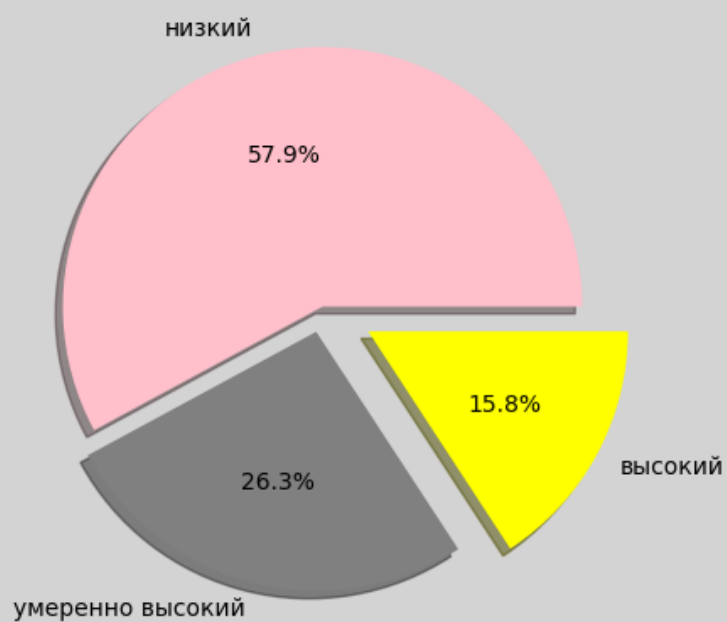


Далее я отсортировал пациентов по гистологическому типу и визуализировал распределение пациентов по индексу Ki-67 но в каждом гистологическом типе по отдельности. Поскольку олигодендроглиом слишком мало, я решил объединить олигодендроглиомы и олигоастроцитомы в единую группу пациентов с олигокомпонентом в гистологии.

Destribution of Ki-67 in astrocytomas group

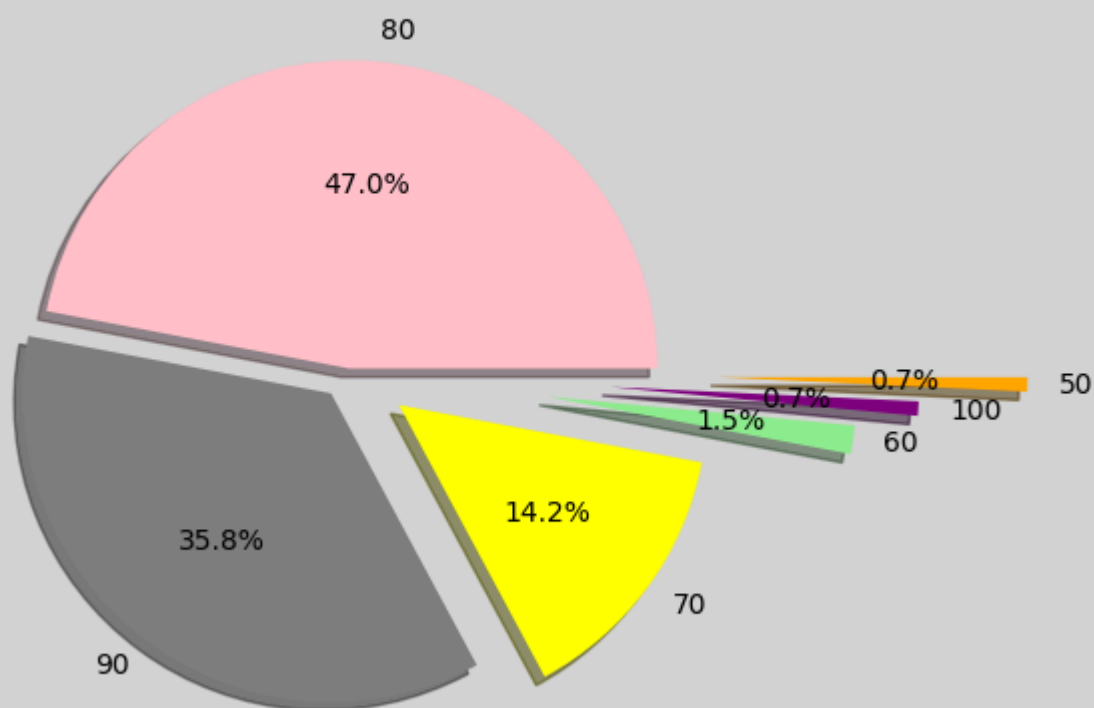


Destribution of Ki-67 in oligo-tumor group

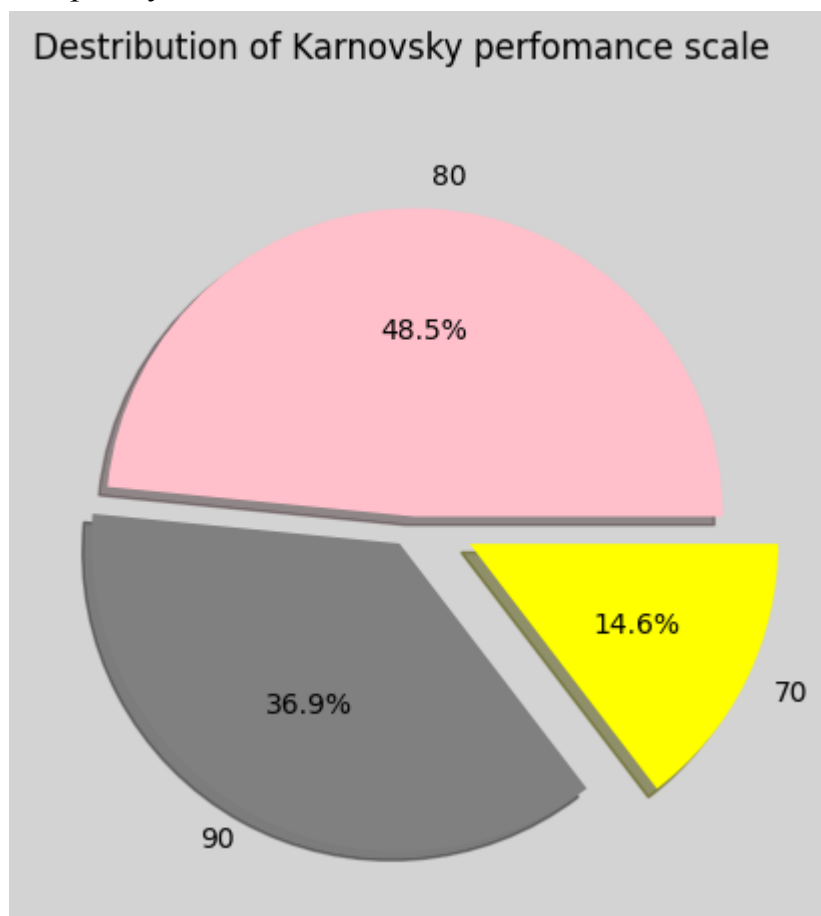


Далее посмотрим на распределение KPS.

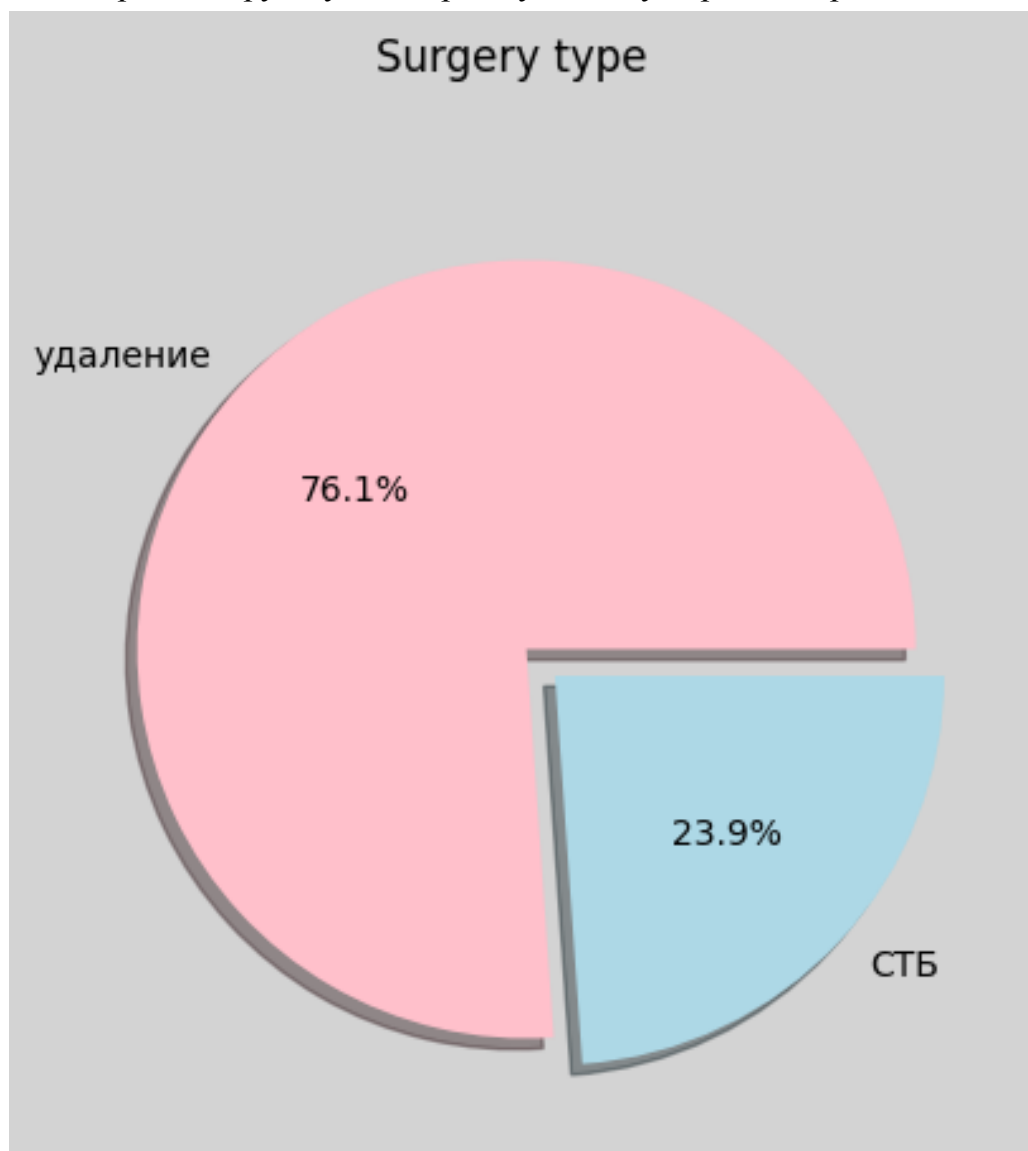
Destribution of Karnovsky perfomance scale



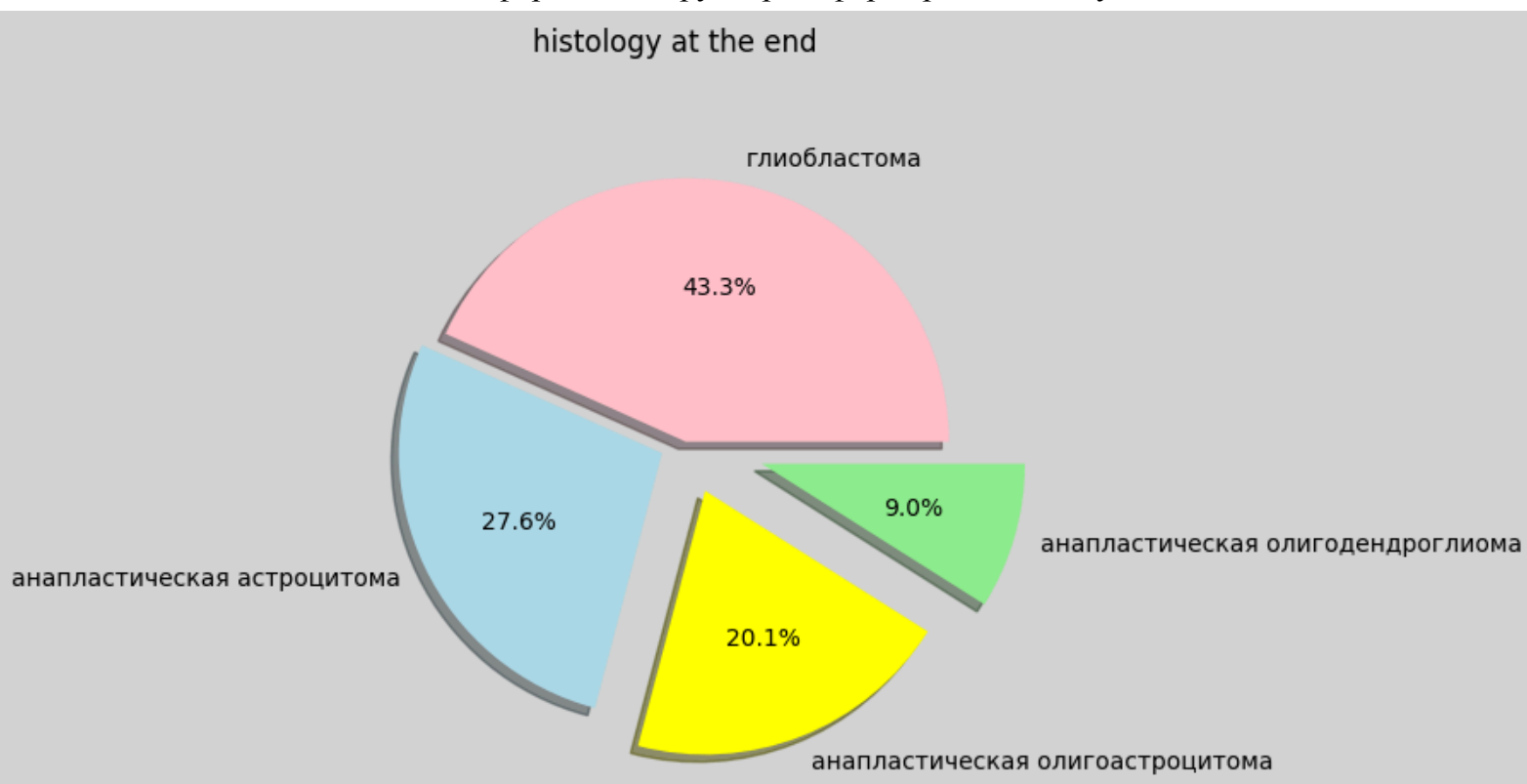
Больше всего пациентов с индексом 70, 80 и 90. С KPS 60 и 100 по одному наблюдению, с KPS 60 два наблюдения. Убрав их мы получаем такую диаграмму:



Посмотрим на круговую диаграмму по типу первой операции:



И итоговая гистологическая форма в которую трансформировались опухоли:



Статистическое планирование.

Были выдвинуты следующие гипотезы:

1. Распределение времени до трансформации:
 - H_0 : время до трансформации распределено нормально
 - H_1 : время до трансформации распределено не нормально
2. Влияние гистологического типа опухоли на скорость трансформации:
 - H_0 : гистологический тип опухоли не имеет влияния на скорость трансформации
 - H_1 : гистологический тип опухоли имеет влияние на скорость трансформации
3. Влияние типа первой операции на скорость трансформации:
 - H_0 : тип операции не имеет влияние на скорость трансформации
 - H_1 : тип операции имеет влияние на скорость трансформации
4. Влияние предоперационного и индекса KPS на скорость трансформации:
 - H_0 : предоперационный индекс KPS не имеет влияние на скорость трансформации
 - H_1 : предоперационный индекс KPS имеет влияние на скорость трансформации

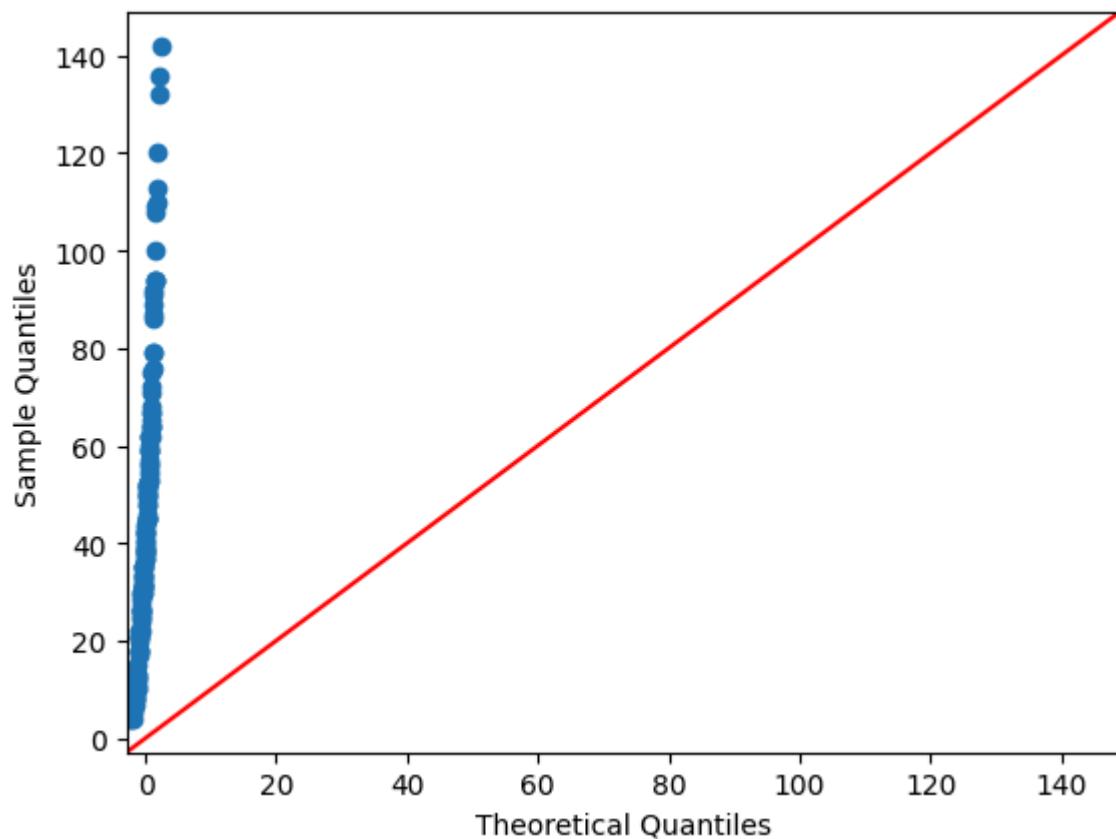
5. Влияние индекса Ki-67 на скорость трансформации:
 - H_0 : индекс Ki-67 не имеет влияние на скорость трансформации
 - H_1 : индекс Ki-67 имеет влияние на скорость трансформации
6. Влияние первичной гистологии на результат трансформации (Grade III или сразу Grade IV):
 - H_0 : первичная гистология не имеет влияние на результат итоговой трансформации
 - H_1 : первичная гистология имеет влияние на результат трансформации
7. Влияние первичной гистологии на индекс Ki-67:
 - H_0 : первичная гистология не имеет влияния на индекс Ki-67
 - H_1 : первичная гистология имеет влияние на индекс Ki-67

Уровень значимости 0,05.

Результаты.

Распределение времени до трансформации:

- H_0 : время до трансформации распределено нормально
- H_1 : время до трансформации распределено не нормально



По данным QQ-plot распределение отличается от нормального значительно.

Применим **критерий согласия Колмогорова**:

$p\text{value}=0.0$

Критерий Шапиро-Уилка:

$p\text{value}=0.0000037659$

Мы отклоняем нулевую гипотезу. Данные распределены не нормально.

Влияние гистологического типа опухоли на скорость трансформации:

- H_0 : гистологический тип опухоли не имеет влияния на скорость трансформации
- H_1 : гистологический тип опухоли имеет влияние на скорость трансформации

Поскольку имеются 3 гистологических типа опухоли (астроцитомы, олигоастроцитомы и олигодендроглиомы) на первом этапе был проведен **однофакторный дисперсионный анализ ANOVA** для определения разницы между группами:

$p\text{ value}=0.00836$ - уверенно отклоняет H_0 гипотезу. Теперь проведем попарные сравнения двухвыборочным t-критерием (в виду ненормальности распределения фокус будем держать на непараметрическом аналоге t-критерия) и применим поправку на множественное сравнение:

астроцитомы vs. олигодендроглиомы: $p\text{ value}=0.007725$

астроцитомы vs. олигоастроцитомы: $p\text{ value}=0.020186$

олигоастроцитомы vs. олигодендроглиомы: $p\text{ value}=0.38576$

Видим что скорость трансформации между олигоастроцитомами и олигодендроглиомами не отличается. Применим **поправку Холма — Бонферрони** и **метод Бенджамини — Хохберга**:

1) поправка Холма-Бонферрони:

$0.007725 * 3 = 0.02317$

$0.020186 * 2 = 0.04037$

С уверенностью можно отклонить H_0 в парах *астроцитомы vs олигоастроцитомы*, *астроцитомы vs олигодендроглиомы*.

2) метод Бенджамини — Хохберга применять нет смысла, критерий Холма-Бенферрони достаточно

Поскольку данные распределены не нормально, то предпочтительно использовать непараметрический аналог критерия Стьюдента:

Теперь применим **критерий Манна-Уитни**:

астроцитомы vs. олигодендроглиомы: $p\text{ value}=0.04055$

астроцитомы vs. олигоастроцитомы: $p\text{ value}=0.07219$

олигоастроцитомы vs. олигодендроглиомы: $p\text{ value}=0.37837$

В данном случае H_0 отклонить не получается уже для пар *астроцитомы vs. олигоастроцитомы* и *олигоастроцитомы vs. олигодендроглиомы*.

1) Посмотрим что даст поправка Бонферрони:

$$0.04055 * 3 = 0.12167$$

2) метод Бенджамини — Хохберга даст тот же результат

отклонить H_0 не удастся.

Также объединим опухоли с олигокомпонентом в одну группу и применим **двухвыборочный t-критерий**:

$$p\text{value}=0.003689$$

отклоняем H_0

Теперь применим **критерий Манна-Уитни (предпочтительнее)**:

$$p\text{value}=0.01841$$

отклоняем H_0

Влияние типа первой операции на скорость трансформации:

- H_0 : тип операции не имеет влияние на скорость трансформации
- H_1 : тип операции имеет влияние на скорость трансформации

Применим **двухвыборочный t-критерий Стьюдента**:

$$p\text{-value} = 0.34829$$

Применим **критерий Манна-Уитни**:

$$p\text{-value} = 0.43215$$

Мы не можем отклонить H_0

Влияние предоперационного и индекса KPS на скорость трансформации:

- H_0 : предоперационный индекс KPS не имеет влияние на скорость трансформации
- H_1 : предоперационный индекс KPS имеет влияние на скорость трансформации

Для проведения тестирования были взяты только KPS 70, 80 или 90.

Применим **однофакторный дисперсионный анализ ANOVA** для определения разницы между группами:

$$p\text{value}=0.55361$$

Разницы между группами не выявлено, нет оснований отклонять H_0 .

Влияние индекса Ki-67 на скорость трансформации:

- H_0 : индекс Ki-67 не имеет влияние на скорость трансформации
- H_1 : индекс Ki-67 имеет влияние на скорость трансформации

Применим **однофакторный дисперсионный анализ ANOVA** для определения разницы между группами:

$pvalue=0.6604$

Разницы между группами не выявлено, нет оснований отклонять H_0 .

Влияние первичной гистологии на результат трансформации (Grade III или сразу Grade IV):

- H_0 : первичная гистология не имеет влияние на результат итоговой трансформации
- H_1 : первичная гистология имеет влияние на результат трансформации

Для исследования данной гипотезы опухоли с олигокомпонентом были объединены в одну группу. В итоге получена таблица сопряженности 2x2.

Для исследования был использован **точный тест Фишера**.

	Astrocytomas	Oligogliomas
Grade III	40	36
Grade IV	51	7

$pvalue=0.00001284$

Уверенно отклоняем H_0 .

Влияние первичной гистологии на индекс Ki-67:

- H_0 : первичная гистология не имеет влияния на индекс Ki-67
- H_1 : первичная гистология имеет влияние на индекс Ki-67

В данном случае также были объединены опухоли с олигокомпонентом в одну группу. Индекс Ki-67 был представлен в 3 вариантах: низкий, умеренный и высокий. В итоге получилось 3 таблицы 2x2, в случае статистической значимости на первом этапе будут применены поправки на множественное сравнение. Для исследования был использован **точный тест Фишера**.

Таблицы:

	Astrocytomas	Oligogliomas
High Ki-67	4	3
Moderate Ki-67	15	5

pvalue=0.633363

	Astrocytomas	Oligogliomas
High Ki-67	4	3
Low Ki-67	21	11

pvalue=0.68566

	Astrocytomas	Oligogliomas
Moderate Ki-67	15	5
Low Ki-67	21	1

pvalue=0.08663

Нет оснований отклонять нулевую гипотезу.

Выводы.

В результате исследования была выявлена более агрессивная биологическая природа глиом астроцитарной гистологической формы, проявляющаяся как в более раннем наступлении анапластической трансформации по сравнению с опухолями имеющими олигодендроглиальный гистологический компонент в своем составе, так и в злокачественности итоговой гистологической формы: астроцитомы чаще трансформируются напрямую в Grade IV. Такие факторы как индекс Карновского, Ki-67, тип выполненной операции не повлияли на скорость

трансформации. Распределение Ki-67 среди гистологических типов так же не было значительно.