



Skoltech
Skolkovo Institute of Science and Technology



Методы реконструкции индивидуальных данных выживаемости, оценка пропорциональности рисков и информативности цензурирования

К.В. Угловская, К.В. Жуденков,
К.В. Песков, Е.Е. Храмеева

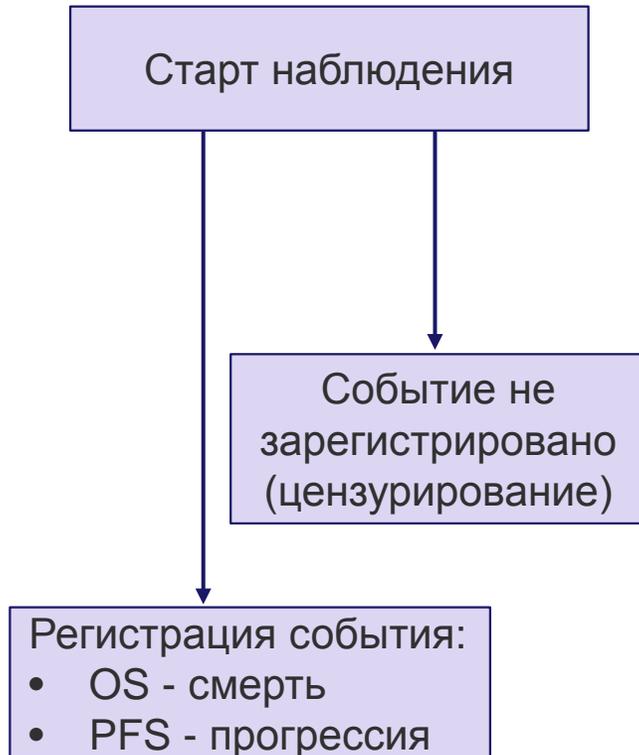
16.10.2025

План

1. Введение: анализ выживаемости в клинических исследованиях
2. Цель и задачи
3. Выбор данных выживаемости
4. Результаты:
 - 2 метода реконструкции
 - Проверка пропорциональности рисков
 - Проверка информативности цензурирования
5. Выводы



Введение: Анализ выживаемости, оценочная функция КМ



№	Time	Status	Treatment
1	0.462	1	1
2	0.707	1	1
3	0.707	1	1
4	0.707	1	1
5	0.707	1	1
6	0.707	1	1
7	0.707	1	1
8	0.757	0	1
9	1.046	1	1
10	1.180	1	1
11	1.180	1	1
12	1.180	1	1
13	1.180	1	1
14	1.360	0	1
15	1.389	0	1
16	1.462	1	1
17	1.462	1	1

Модели выживаемости

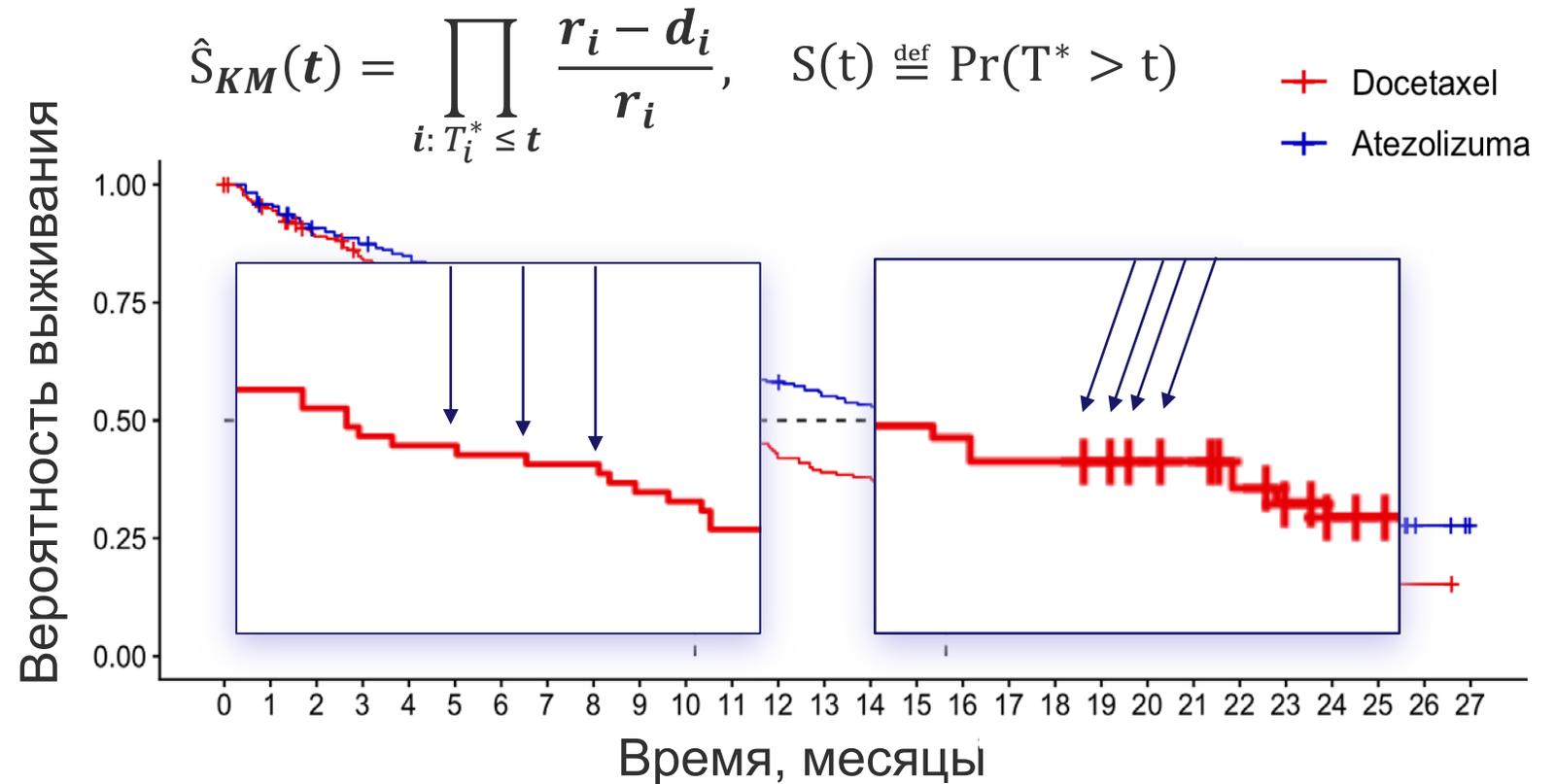
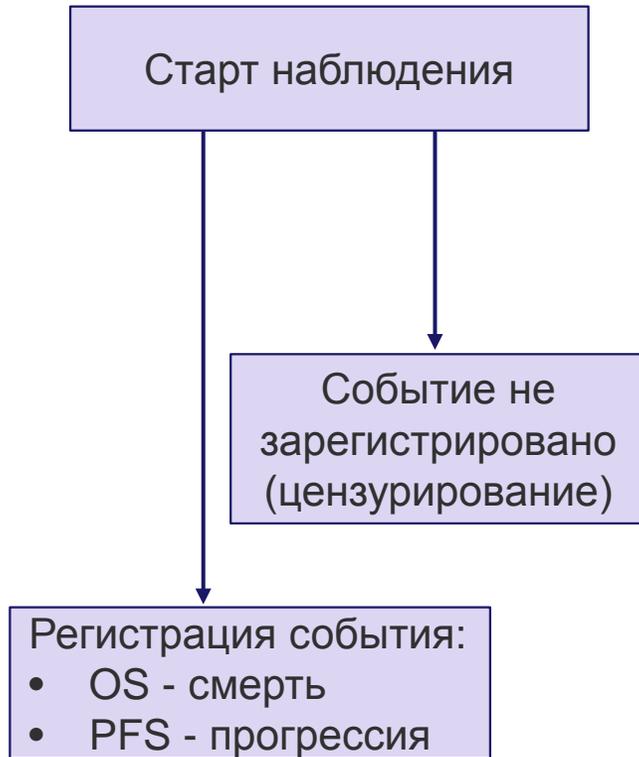
- Доля выживших
- Многофакторный анализ
- Оценка величины эффекта
- Предсказания для новых субъектов

↓

Мета-анализ – вершина доказательной медицины



Введение: Анализ выживаемости, оценочная функция КМ



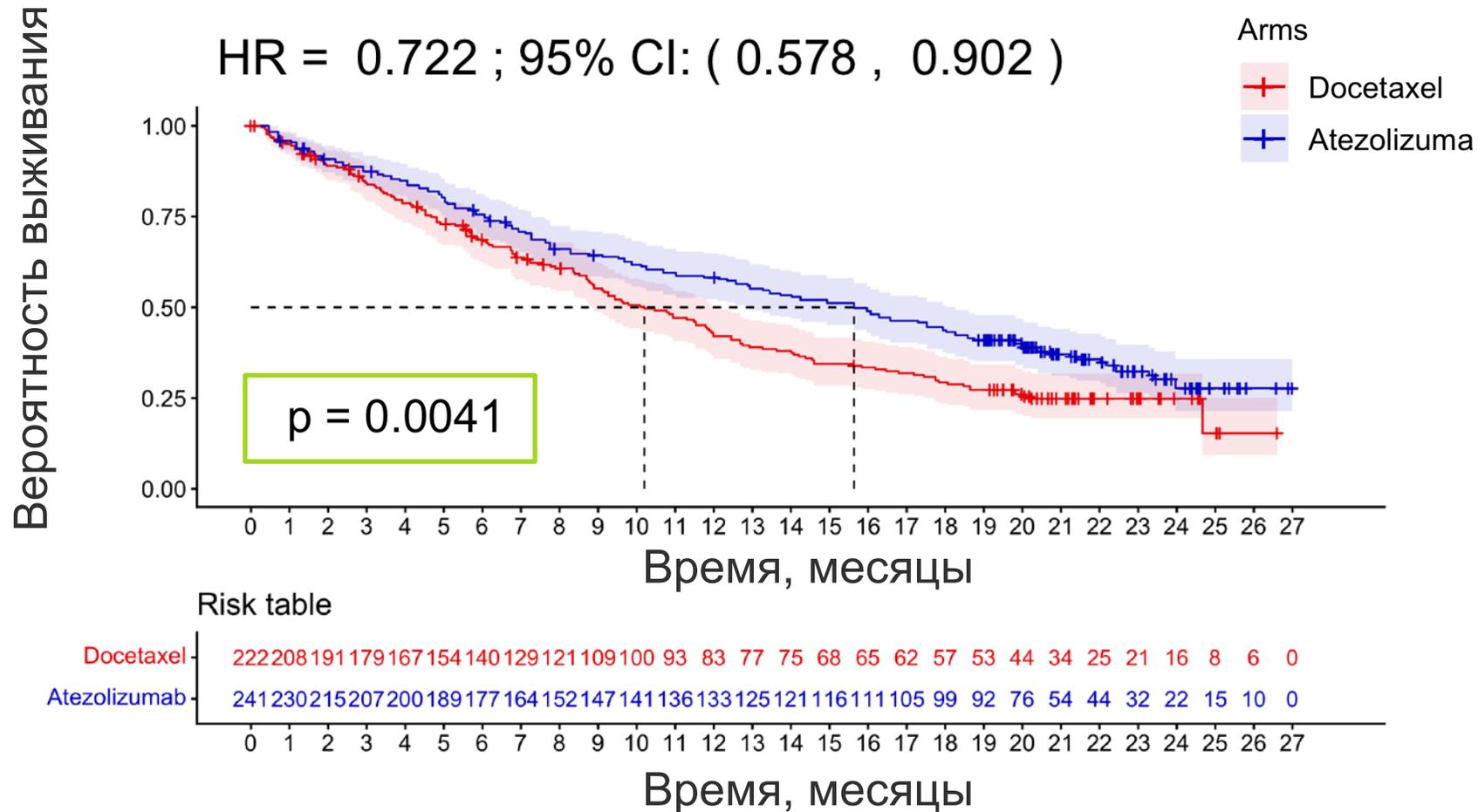
4 r_{i-} кол-во субъектов в группе риска к моменту наступления события (T_i^*), d_i - кол-во событий в



- Умение корректно восстановить индивидуальные данные по выживаемости является актуальной задачей для дальнейшего построения мета-анализа



Введение: Анализ выживаемости, оценочная функция КМ



6 Greenwood's formula: $\widehat{Var}[\hat{S}_{KM}(t)] = \hat{S}_{KM}(t)^2 \sum_{i: T_i^* \leq t} \frac{d_i}{r_i(r_i - d_i)}$



Введение: Модели выживаемости

$$S(t) = e^{-\int_0^t h(u)du}$$

Модель Кокса: $h(t|x) = h_0(t) \times e^{\beta^T x}$, $h_0(t)$ - базовая функция риска, x - вектор ковариат, β - вектор коэффициентов регрессии

Необходимое условие применимости модели Кокса

Пропорциональность рисков: $HR \stackrel{\text{def}}{=} \frac{h_i(t|x_i)}{h_k(t|x_k)} = e^{\beta^T(x_i - x_k)} = \text{const}(t)$



- Проверка соблюдения пропорциональности рисков диктует дальнейший выбор как статистических тестов сравнения, так и моделей выживаемости.
- Цензурирование может оказаться несбалансированным/информативным, и это далее может потребовать применения моделей с несколькими функциями риска одновременно (competing risk models) и др.



Введение: Актуальные вопросы

1. Каков наиболее корректный способ воспроизведения данных выживаемости из КМ графиков для дальнейшего построения моделей?
2. Как проверить соблюдение условий пропорциональности рисков? В случае невыполнения условия пропорциональности рисков какие альтернативы логрангового теста существуют?
3. Может ли неравномерность цензурирования существенно влиять на возможные итоги сравнения КМ кривых между собой?



Цель и задачи

Цель: разработка методики для реконструкции и анализа индивидуальных данных выживаемости

Задачи:

1. Сравнить эффективные алгоритмы реконструкции индивидуальных данных по КМ кривым
2. Проверить пропорциональность рисков для обоснования применимости ранговых тестов сравнения и оценки соотношения рисков (HR)
3. Протестировать данные выживаемости на несбалансированность цензурирования



Исследуемые данные

Репрезентативный выбор опубликованных данных для различных видов рака, различных конечных точек, различного количества субъектов:

- **Общая выживаемость (OS):**
advanced NSCLC,
immune therapy (atezolizumab, durvalumab, tremelimumab, combinations)
vs. chemotherapy (docetaxel, cisplatin, pemetrexed, combinations)
- **Выживаемость без прогрессии (PFS):**
pancreatic, ovarian, myeloma, thyroid, neuroendocrine

Исследование	Подгруппа	Total N
NCT02657434 [1], OS	ITT	578
	PD-L1 high	45
	PD-L1 low	136
	PD-L1 neg.	163
NCT03164616 [2], OS	T+D+CT & CT	675
	D+CT & CT	675
NCT02008227 [3], OS	ITT	850
	PD-L1 ≥ 1%	463
	PD-L1 ≥ 5%	265
	PD-L1 ≥ 50%	137
	PD-L1 < 1%	379
	Non-squam.	628
	Squamous	222
NCT02184195 [4], PFS	ITT	154
NCT00262847 [5], PFS	ITT	1248
ISRCTN49407852 [6], PFS	ITT	1971
NCT02655016 [7], PFS	ITT	733
NCT01321554 [8], PFS	ITT	392
NCT01524783 [9], PFS	ITT	302



1

Реконструкция данных

Процедура реконструкции индивидуальных данных

1. Оцифровка: извлечение координат точек КМ-графика, PlotDigitizer app
2. Реконструкция: итерационные алгоритмы расчета времени событий и цензурирования

Liu et al., 2021

Оцифровка:

- Максимальное количество точек КМ кривой

Дополнительные данные:

- **Таблица риска***
- **Общее количество событий***

Rogula et al., 2022

Оцифровка:

- Только нижние точки ступеней КМ кривой

Дополнительные данные:

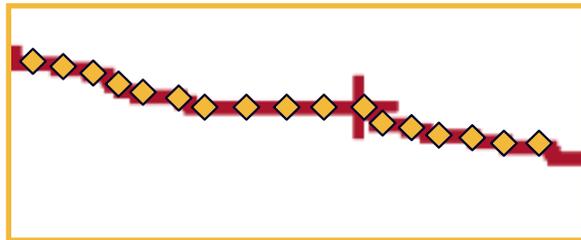
- **Точные цензурирования***
- **Общее количество пациентов**



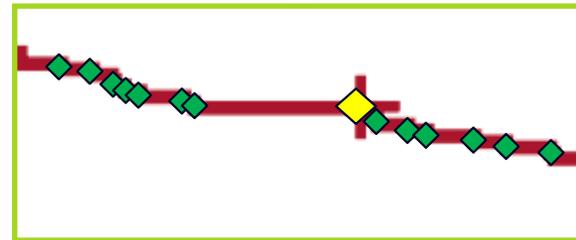
Процедура реконструкции индивидуальных данных

1. Оцифровка: извлечение координат точек КМ-графика, PlotDigitizer app
2. Реконструкция: итерационные алгоритмы расчета времени событий и цензурирования

Liu



Rogula



Time in months	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Docetaxel	222	208	191	179	167	154	140	129	121	109	100	93	83	77	75	68	65	62	57	53	44	34	25	21	16	8	6	0
Atezolizumab	241	230	215	207	200	189	177	164	152	147	141	136	133	125	121	116	111	105	99	92	76	54	44	32	22	15	10	0

Проверка качества реконструкции

- Медианная выживаемость
- HR (соотношение рисков)
- Таблица субъектов в группе риска
- Распределение цензурирования
- Наложение КМ графиков



Предварительные выводы о методах реконструкции

- Методы Liu и Rogula адекватно воспроизводят индивидуальные данные
- Метод Liu более удобен для больших выборок (>200), где крайне сложно детектировать отдельные события цензурирования
- Метод Rogula сохраняет точную информацию о цензурировании, но не требует максимально точного отслеживания траектории кривой КМ



2

Пропорциональность рисков

Проверка пропорциональности рисков

Пропорциональность рисков $HR = const(t)$:

1. Необходимое условие применимости модели Кокса (и вычисления HR)
2. Необходимое условие применимости логрангового теста сравнения КМ-кривых выживаемости

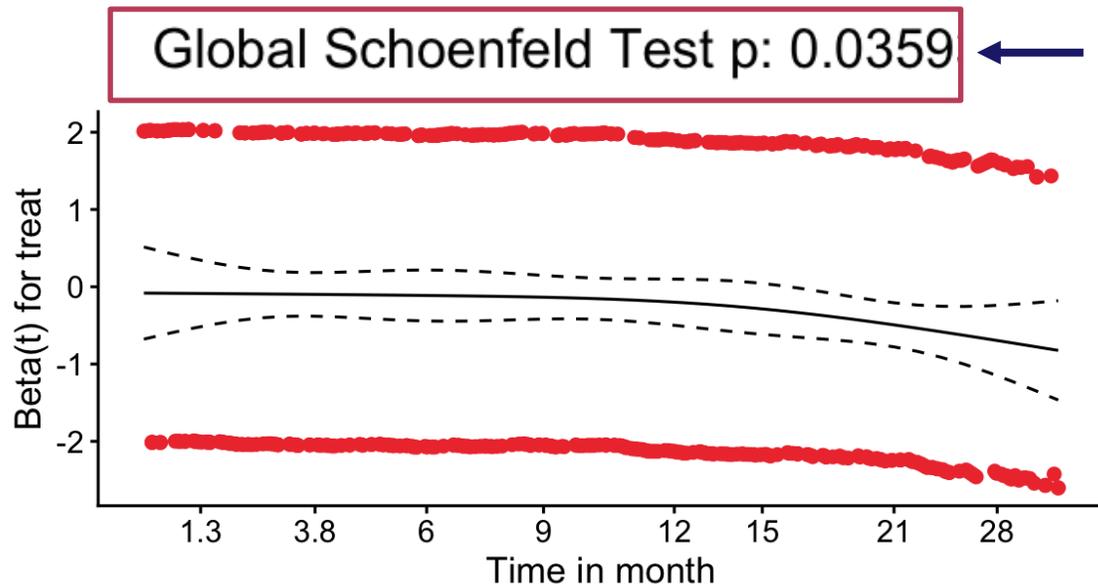
Проверка пропорциональности рисков:

Schoenfeld residuals $r_{Sch_k} = \sum_i \delta_{ik} (x_i - \bar{x}(\beta, T_k^*))$, k - номер события, i - субъект в группе риска в T_k^*

$p\text{-value}_{Sch} < 0.05 \Rightarrow$ наличие временной зависимости ковариат или их коэффициентов \Rightarrow
нарушение $HR = const(t)$



Проверка пропорциональности рисков



Метод Rogula: более частое нарушение гипотезы о пропорциональности рисков => более сложные модели выживаемости

Liu method

	Подгруппа	p(Sch)
[1]	ITT	> 0.05
	PD-L1 high	< 0.05
	PD-L1 low	> 0.05
	PD-L1 neg.	> 0.05
[2]	T+D+CT & CT	< 0.05
	D+CT & CT	> 0.05
[3]	ITT	> 0.05
	PD-L1 ≥ 1%	> 0.05
	PD-L1 ≥ 5%	> 0.05
	PD-L1 ≥ 50%	> 0.05
	PD-L1 < 1%	> 0.05
	Non-squam.	> 0.05
	Squamous	> 0.05
[4]	ITT	> 0.05
[5]	ITT	< 0.05
[6]	ITT	> 0.05
[7]	ITT	> 0.05
[8]	ITT	> 0.05
[9]	ITT	> 0.05

Rogula method

	Подгруппа	p(Sch)
[1]	ITT	> 0.05
	PD-L1 high	< 0.05
	PD-L1 low	> 0.05
	PD-L1 neg.	> 0.05
[2]	T+D+CT & CT	< 0.05
	D+CT & CT	> 0.05
[3]	ITT	< 0.05
	PD-L1 ≥ 1%	> 0.05
	PD-L1 ≥ 5%	> 0.05
	PD-L1 ≥ 50%	> 0.05
	PD-L1 < 1%	> 0.05
	Non-squam.	> 0.05
	Squamous	> 0.05
[4]	ITT	> 0.05
[5]	ITT	< 0.05
[6]	ITT	< 0.05
[7]	ITT	< 0.05
[8]	ITT	> 0.05
[9]	ITT	< 0.05

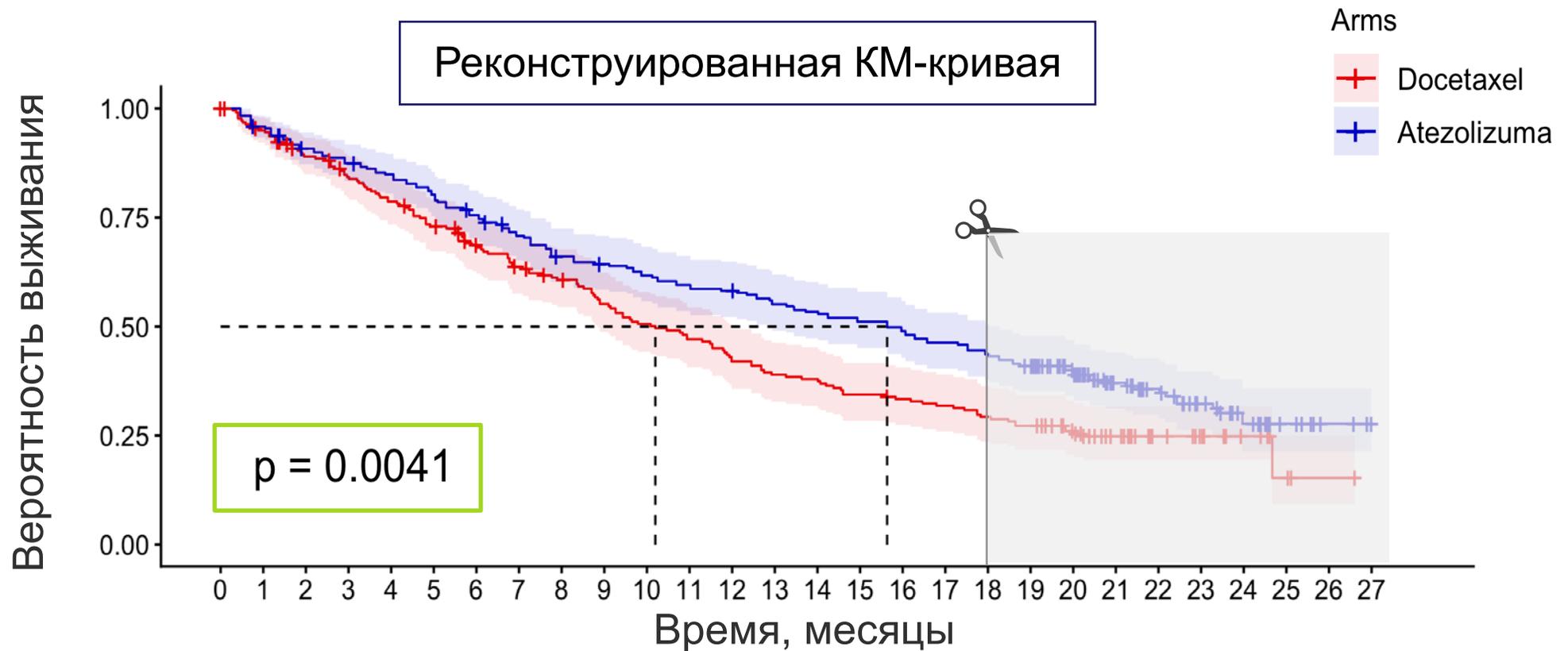


3

Сбалансированность цензурирования

Проверка сбалансированности цензурирования

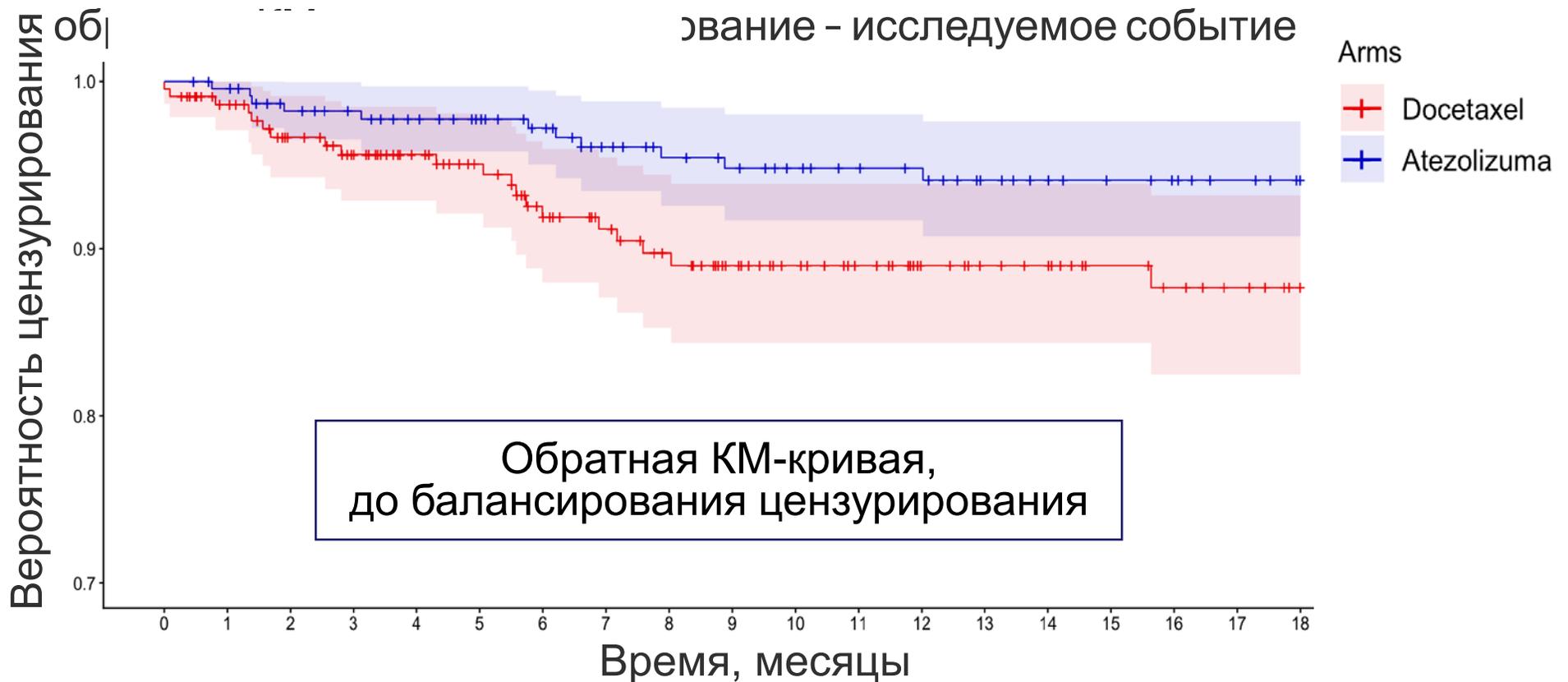
- I. Выбор тестируемого временного интервала для проверки сбалансированности цензурирования



Проверка сбалансированности цензурирования

1. Выбор тестируемого временного интервала для проверки сбалансированности цензурирования

2. Построен



Проверка сбалансированности цензурирования

Liu method

	Подгруппа	p(Sch)	Reversed HR
[1]	ITT	> 0.05	0.80 (0.38 - 1.68)
	PD-L1 high	< 0.05	0.70 (0.18 - 2.82)
	PD-L1 low	> 0.05	0.85 (0.20 - 3.56)
	PD-L1 neg.	> 0.05	0.88 (0.18 - 4.34)
[2]	T+D+CT & CT	< 0.05	0.26 (0.10 - 0.72)
	D+CT & CT	> 0.05	0.66 (0.32 - 1.39)
[3]	ITT	> 0.05	2.39 (1.73-3.32)
	PD-L1 ≥ 1%	> 0.05	0.41 (0.22 - 0.78)
	PD-L1 ≥ 5%	> 0.05	0.26 (0.11 - 0.64)
	PD-L1 ≥ 50%	> 0.05	0.50 (0.14 - 1.78)
	PD-L1 < 1%	> 0.05	0.61 (0.31 - 1.22)
	Non-squam.	> 0.05	0.40 (0.23 - 0.69)
	Squamous	> 0.05	0.90 (0.34 - 2.41)
[4]	ITT	> 0.05	0.59 (0.30 - 1.13)
[5]	ITT	< 0.05	0.74 (0.38 - 1.42)
[6]	ITT	> 0.05	1.26 (1.09 - 1.46)
[7]	ITT	> 0.05	0.99 (0.76 - 1.29)
[8]	ITT	> 0.05	—
[9]	ITT	> 0.05	—

Rogula method

	Подгруппа	p(Sch)	Reversed HR
[1]	ITT	> 0.05	0.60 (0.25 - 1.47)
	PD-L1 high	< 0.05	0.52 (0.12 - 2.34)
	PD-L1 low	> 0.05	1.21 (0.32 - 4.53)
	PD-L1 neg.	> 0.05	0.90 (0.13 - 6.37)
[2]	T+D+CT & CT	< 0.05	0.69 (0.24 - 2.00)
	D+CT & CT	> 0.05	0.59 (0.19 - 1.79)
[3]	ITT	< 0.05	0.68 (0.38 - 1.21)
	PD-L1 ≥ 1%	> 0.05	0.45 (0.22 - 0.95)
	PD-L1 ≥ 5%	> 0.05	0.36 (0.13 - 1.01)
	PD-L1 ≥ 50%	> 0.05	0.58 (0.16 - 2.18)
	PD-L1 < 1%	> 0.05	1.16 (0.51 - 2.63)
	Non-squam.	> 0.05	0.65 (0.34 - 1.23)
	Squamous	> 0.05	1.04 (0.38 - 2.87)
[4]	ITT	> 0.05	0.49 (0.23 - 1.06)
[5]	ITT	< 0.05	0.51 (0.22 - 1.18)
[6]	ITT	< 0.05	0.52 (0.43 - 0.62)
[7]	ITT	< 0.05	0.55 (0.39 - 0.79)
[8]	ITT	> 0.05	0.62 (0.28 - 1.38)
[9]	ITT	< 0.05	0.72 (0.44 - 1.18)



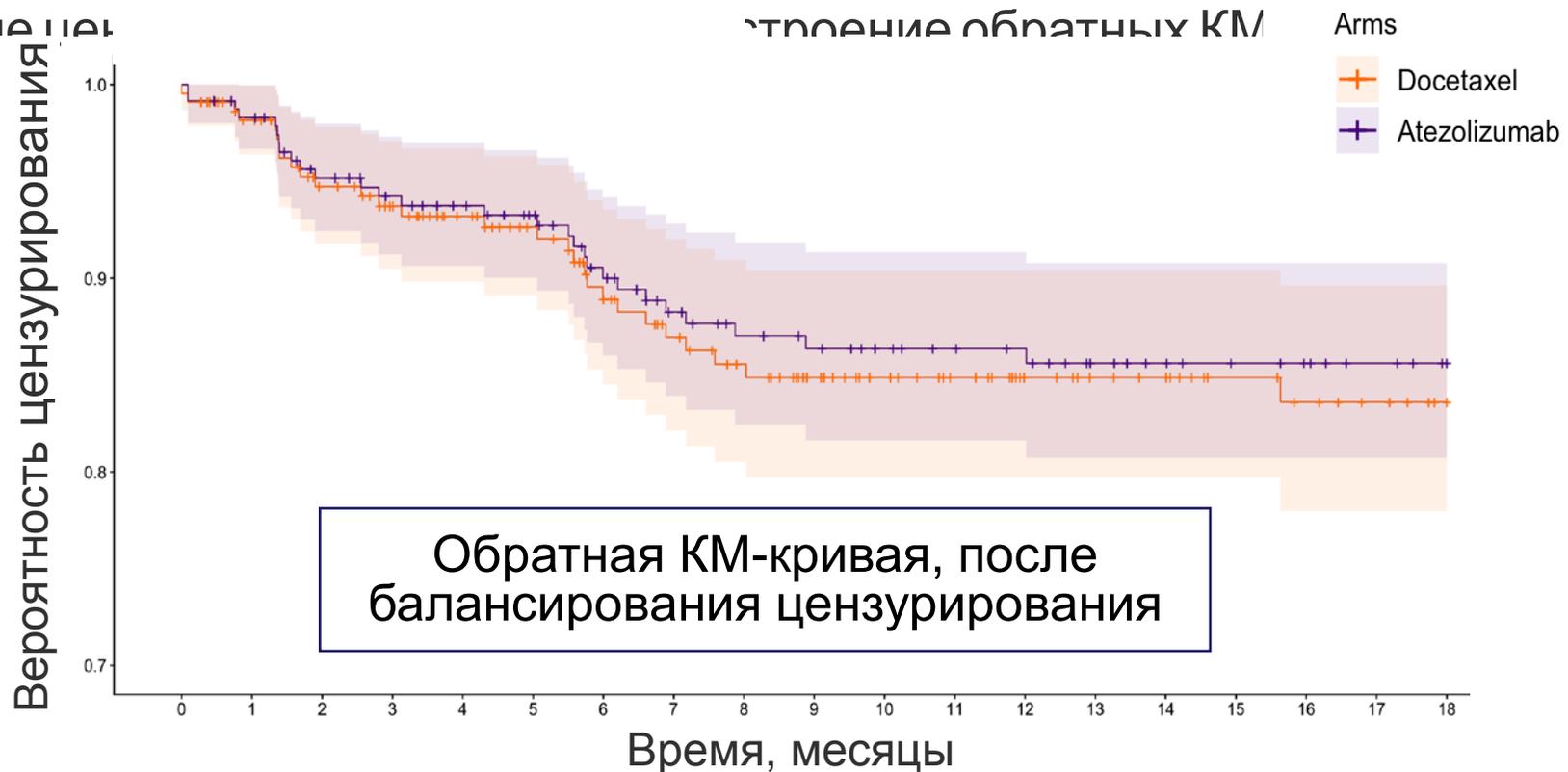
Проверка сбалансированности цензурирования

1. Выбор тестируемого временного интервала для проверки сбалансированности цензурирования
2. Построение обратных КМ кривых: цензурирование - исследуемое событие
3. Балансирование цензурирования в группах:
 - Цензурирование больше в экспериментальной группе => цензурирование в контрольной группе первого пациента, у которого после этого произошло событие.
 - Цензурирование больше в контрольной группе => цензурирование пациента, прожившего дольше всех и не испытавшего событие в экспериментальной группе.



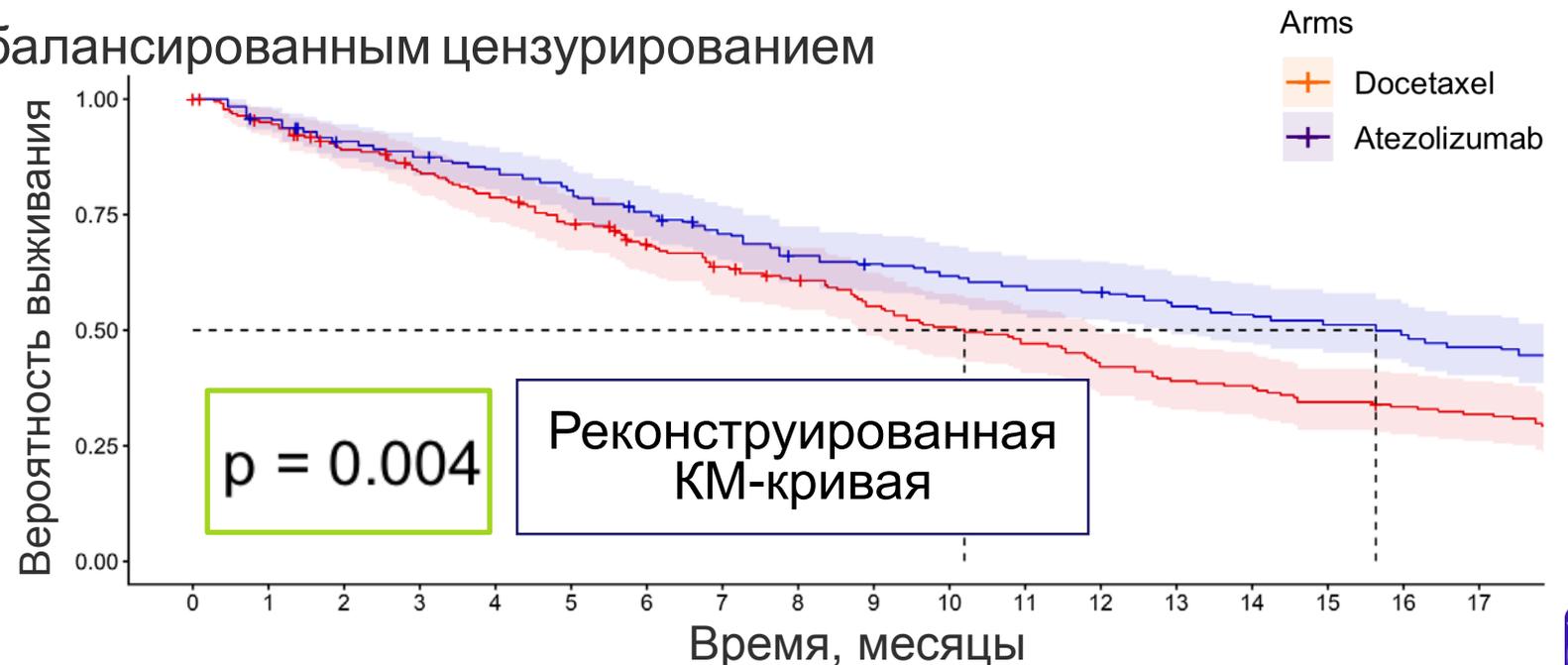
Проверка сбалансированности цензурирования

1. Выбор тестируемого временного интервала для проверки сбалансированности цензурирования
2. Построение обратных КМ кривых: цензурирование - исследуемое событие
3. Балансирование



Проверка сбалансированности цензурирования

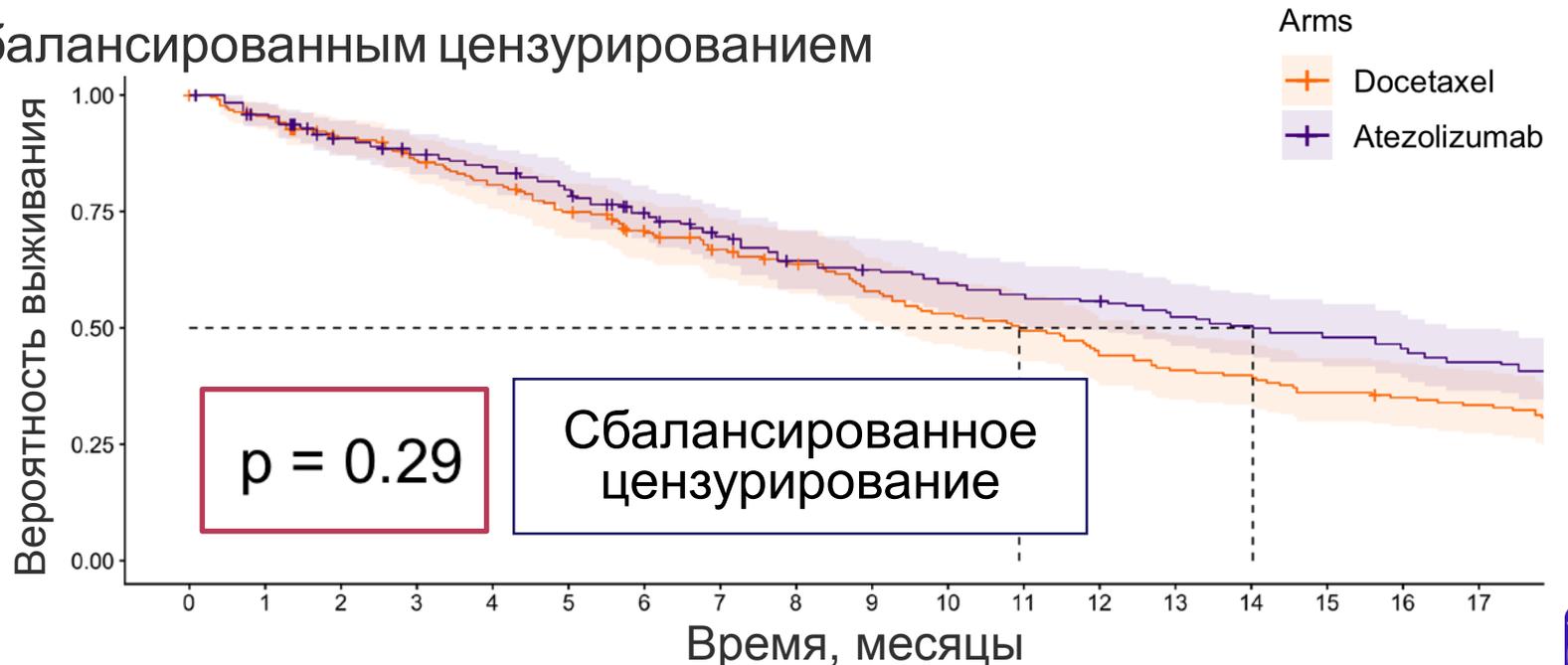
1. Выбор тестируемого временного интервала для проверки сбалансированности цензурирования
2. Построение обратных КМ кривых: цензурирование = интересующее событие
3. Балансирование цензурирования в группах, построение обратных КМ кривых
4. Сравнение КМ кривых со сбалансированным цензурированием



Проверка сбалансированности цензурирования

1. Выбор тестируемого временного интервала для проверки сбалансированности цензурирования
2. Построение обратных КМ кривых: цензурирование = интересующее событие
3. Балансирование цензурирования в группах, построение обратных КМ кривых
4. Сравнение КМ кривых со сбалансированным цензурированием

Балансировка цензурирования существенно влияет на итоги сравнения КМ кривых
=> Модели competing risks для событий и цензурирования



Проверка сбалансированности цензурирования

Liu method

	Подгруппа	Reversed HR	New HR
[1]	ITT	0.80 (0.38 - 1.68)	
	PD-L1 high	0.70 (0.18 - 2.82)	
	PD-L1 low	0.85 (0.20 - 3.56)	
	PD-L1 neg.	0.88 (0.18 - 4.34)	
[2]	T+D+CT & CT	0.26 (0.10 - 0.72)	0.87 (0.73 - 1.03)
	D+CT & CT	0.66 (0.32 - 1.39)	
[3]	ITT	2.39 (1.73-3.32)	1.30 (1.07 - 1.58)
	PD-L1 ≥ 1%	0.41 (0.22 - 0.78)	0.94 (0.73 - 1.15)
	PD-L1 ≥ 5%	0.26 (0.11 - 0.64)	1.03 (0.75 - 1.40)
	PD-L1 ≥ 50%	0.50 (0.14 - 1.78)	
	PD-L1 < 1%	0.61 (0.31 - 1.22)	
	Non-squam.	0.40 (0.23 - 0.69)	0.91 (0.74 - 1.11)
	Squamous	0.90 (0.34 - 2.41)	
[4]	ITT	0.59 (0.30 - 1.13)	
[5]	ITT	0.74 (0.38 - 1.42)	
[6]	ITT	1.26 (1.09 - 1.46)	1.07 (0.92 - 1.24)
[7]	ITT	0.99 (0.76 - 1.29)	
[8]	ITT	—	—
[9]	ITT	—	—

Rogula method

	Подгруппа	Reversed HR	New HR
[1]	ITT	0.60 (0.25 - 1.47)	
	PD-L1 high	0.52 (0.12 - 2.34)	
	PD-L1 low	1.21 (0.32 - 4.53)	
	PD-L1 neg.	0.90 (0.13 - 6.37)	
[2]	T+D+CT & CT	0.69 (0.24 - 2.00)	
	D+CT & CT	0.59 (0.19 - 1.79)	
[3]	ITT	0.68 (0.38 - 1.21)	
	PD-L1 ≥ 1%	0.45 (0.22 - 0.95)	0.89 (0.71 - 1.11)
	PD-L1 ≥ 5%	0.36 (0.13 - 1.01)	
	PD-L1 ≥ 50%	0.58 (0.16 - 2.18)	
	PD-L1 < 1%	1.16 (0.51 - 2.63)	
	Non-squam.	0.65 (0.34 - 1.23)	
	Squamous	1.04 (0.38 - 2.87)	
[4]	ITT	0.49 (0.23 - 1.06)	
[5]	ITT	0.51 (0.22 - 1.18)	
[6]	ITT	0.52 (0.43 - 0.62)	0.63 (0.56 - 0.71)
[7]	ITT	0.55 (0.39 - 0.79)	
[8]	ITT	0.62 (0.28 - 1.38)	
[9]	ITT	0.72 (0.44 - 1.18)	

Балансировка цензурирования в реконструкции Liu систематически приводит к потере значимости различия KM кривых



Выводы

1. Оба протестированных метода реконструкции индивидуальных данных показали эффективное воспроизведение формы KM кривых, однако метод Rogula более предпочтителен при работе с данными, т.к. максимально точно учитывает цензурирование.
2. Реконструированные данные требуют проверки предположения о пропорциональности рисков, что определяет последующий выбор модели выживаемости и ранговых тестов сравнения: анализ Schoenfeld residuals выявил 7 случаев нарушения пропорциональности при реконструкции Rogula и 3 нарушения в случае Liu => может потребоваться разработка более сложных моделей выживаемости.
3. Анализ цензурирования выявил систематическую несбалансированность при использовании метода Liu, что является недостатком метода в целом. Выявление несбалансированности при оцифровке методом Rogula требует дальнейшего рассмотрения информативного цензурирования в рамках competing risks и других соответствующих подходов.



Литература

1. Nishio M et al. (2021). *J Thorac Oncol*, 16(4), 653–664
2. Johnson M et al. (2023). *J Clin Oncol*, 41(6), 1213–1227
3. Rittmeyer A et al. (2017). *Lancet*, 389(10066), 255–265
4. Ilson D et al. (2019) *NEJM Journal Watch*, NA49254
5. Burger R et al.(2011). *NEJM*, 365(26), 2473-2483
6. Jackson G et al.(2019). *The Lancet Oncology*, 20(1), 57-73
7. González-Martín A et al. (2019). *NEJM*, 381(25), 2391-2402
8. Schlumberger M et al. (2015). *NEJM*, 372(7), 621-630
9. Yao J et al. (2016). *The Lancet*, 387(10022), 968-977
10. Liu N et al. (2021). *BMC Med Res Methodol*, 21(1), 111
11. Rogula B et al. (2022). *MDM Policy Pract*, 7(1), 23814683221077643
12. Rizopoulos, D. (2012). CRC press.
13. Sawyer, S. (2003). *Health (San Francisco)*. i, 2-5.
14. Karadeniz, P. G., & Ercan, I. (2017). *Stat Transit*, 18(2), 311-28.
15. Gilboa S et al. (2021). *Eur J Cancer*, 153, 190-202



0

BACKUP

Выбор рангового теста сравнения

	Test name, weights	PH assumption	Censoring distr. pattern	Sample size
1	Regular log-rank test, equal	HR = const.(t)	Different	Small, if surv. curves cross
2	Gehan-Breslow (generalized Wilcoxon), very heavy beginning	HR<1 or >1	Same, low	
3	Tarone-Ware, heavy beginning	Only equal pop. var. assumption		Not very small
4	Peto-Peto's modified, slightly heavy beginning	—	Same, tests more resistant than Gehan-Breslow	
5	modified Peto-Peto (by Andersen), slightly heavy beginning	—		
6	Fleming-Harrington (p=1, q=1), weight in the middle	Specific, choice of weights can impact the sensitivity to detect differences at different time points		



Пропорциональность рисков

Test name	NCT02657434 [1], High PD-L1; $p > 0.05$		NCT03164616 [2], T+D+CT & CT; $p = 3E-3$		NCT02008227 [3], ITT; $p = 3E-4$	
	Liu p	Rogula p	Liu p	Rogula p	Liu p	Rogula p
Log-rank test	0.4752	0.4766	0.0043	0.0031	0.0004	0E+00
Peto-Peto's modified	0.2757	0.2778	0.0540	0.0478	0.0023	2E-04

Невыполнение условия пропорциональности рисков: неприменимость HR, логрангового

теста

